

**The development of learners' listening comprehension,
self-efficacy, and anxiety within an informal digital learning
of English listening (IDLEL) context: Examining the role of
IDLEL engagement and self-regulation**

Huining Yang

Institute of Education
UNIVERSITY OF READING

Thesis submitted for the degree of
Doctor of Philosophy

February 2025

Declaration of Original Authorship

I confirm that this is my own work and the use of all material from other sources
has been properly and fully acknowledged.

Huining Yang

Abstract

Listening is a crucial language skill for L2 learners, as it is not only one of the most frequently used skills but also plays a vital role in the development of other language skills. However, beyond these benefits, English listening proficiency can also be crucial for Chinese university EFL learners' academic success, career prospects, and global mobility. For instance, it can impact learners' exam performance, is essential for meeting job market demands in an increasingly globalised economy, and is a key requirement for studying abroad, where strong listening skills are vital for both academic and daily life.

However, Chinese EFL university learners often face numerous challenges in developing their listening proficiency in formal classroom settings, such as limited instructional time, insufficient teaching resources, a lack of diverse teaching methods, and an exam-oriented curriculum. These factors may hinder adequate listening development, leading many learners to experience moderate to high levels of listening anxiety and low self-efficacy. One effective solution is for students to actively engage in informal L2 listening practice outside the classroom. With technological advancements, informal digital learning of English (IDLE) has provided learners with greater access to resources and opportunities. While IDLE has gained increasing research attention, little is known about Chinese EFL university learners' L2 listening development within informal learning contexts.

Moreover, unlike traditional teacher-centered language classrooms, where external regulation is dominant, informal language learning contexts can provide learners with greater autonomy and freedom. In the absence of external regulation, however, learners' self-regulated learning (SRL) abilities become particularly crucial. While numerous models

have been developed to illustrate the mechanisms of SRL, models specifically constructed for L2 listening remain absent. Furthermore, considering the significant impact of motivational and affective factors on listening, regulation of motivation and affect should be incorporated into SRL frameworks. However, few existing models simultaneously address both cognitive (i.e., listening) and motivational/affective regulation. Additionally, SRL, self-efficacy, and listening anxiety have a complex relationship; however, their joint predictive mechanism for listening remains unclear. Therefore, addressing the research gaps above serves as the primary aim of this study.

This study employed a mixed-method research design. Based on an analysis of the strengths and limitations of existing SRL models, the study proposed a five-phase, dual-level SRL model targeted L2 listening, referred to as the Self-Regulated L2 Listening Model. To validate the hypothesised structure of the model, a questionnaire was developed based on this theoretical framework, namely the *Self-Regulated L2 Listening Questionnaire (SRLLQ)*. A total of 582 EFL learners from five universities in China were invited to complete the SRLLQ, and 523 valid responses were analysed using Confirmatory Factor Analysis (CFA), which confirmed the five-phase, dual-level structure of the Self-Regulated L2 Listening Model.

Additionally, the study proposed two hypothesised joint predictive mechanisms of SRL, self-efficacy, and listening anxiety on L2 listening. To test these mechanisms and explore learners' L2 listening development in an informal digital learning of English listening (IDLEL) context, another 130 English majors from two of the five universities participated in the IDLEL study. They completed three listening tests (pre, post, and delayed post-test) and two questionnaires (pre, and post-test) measuring their self-regulation, self-efficacy, and

listening anxiety. Additionally, they participated in a four-week observational IDLEL study, during which they recorded their IDLEL engagement in their weekly E-logs.

The Structural Equation Modeling (SEM) results confirmed the two hypothesised predictive mechanisms: the first mechanism revealed the direct predictive effect of self-efficacy on listening, as well as its indirect effect through SRL and listening anxiety; the second mechanism focused on the direct predictive effect of SRL on listening and its indirect effect, with self-efficacy and listening anxiety being the mediators. Together, these two mechanisms illustrate a positive cycle that facilitates L2 learners' listening development. Additionally, the multiple predictive pathways of SRL suggest that its influence on listening is not only immediate but also potentially long-term.

Moreover, descriptive analysis, thematic analysis, and cluster analysis of participants' IDLEL E-logs provided insights into the quantity (i.e., frequency and duration of engagement) and quality (i.e., diversity of activities engaged and SRL strategy use) of their IDLEL engagement. Linear Mixed Models (LMMs) were then constructed to reveal the predictive effects of SRL and IDLEL engagement on listening, self-efficacy, and listening anxiety. The results showed that participants' SRL (post-test) and the duration of their IDLEL activity engagement significantly and positively predicted their listening improvement from pre-test to post-test, but they were not found to have a significant predictive effect on the listening post-test. Additionally, participants' SRL (post-test) and the duration of their IDLEL engagement significantly and positively predicted their listening self-efficacy at both the pre-test and post-test, while the frequency of IDLEL engagement had a significant negative predictive effect on their listening self-efficacy at both time points. Furthermore, participants'

SRL (post-test) was found to significantly negatively predict their listening anxiety at both time points. Finally, the moderation analysis showed that SRL did not moderate the relationship between IDLEL engagement and listening and self-efficacy, indicating that IDLEL engagement may be universally beneficial to learners, regardless of their SRL abilities.

Acknowledgements

I would first like to express my deepest gratitude to my two supervisors, Professor Suzanne Graham and Dr. Pengchong Anthony Zhang. From the moment they accepted me as their student, they have offered me unwavering support and the warmest care, not only throughout my PhD journey but also beyond. Suzanne has been an invaluable academic guide, especially in shaping my theoretical understanding. What touches me the most is how she frequently emails me relevant literature that could benefit my research. Her recommendations always come at the perfect time, sparking new ideas and inspiring my work. Moreover, Suzanne has always been incredibly patient and understanding, embracing my mistakes and guiding me with unwavering encouragement. Her support and kindness have been a crucial pillar in helping me complete my PhD. Whenever people mention Suzanne, I feel an immense sense of pride. She is such an outstanding scholar, and YES, she is my dear supervisor. From the bottom of my heart, thank you, Suzanne.

I am equally grateful to Anthony. Without his expertise in data analysis methods, I would not have been able to complete my research. However, the impact of his guidance goes far beyond this study, it can be long-lasting, even lifelong. Anthony, thank you for always answering my questions, listening to my concerns, and engaging with me as a friend. Thank you for believing in me from the very beginning. I feel incredibly fortunate to have gained not only an exceptional supervisor but also a great friend. You will always be my 好恩师 ! Whenever you need my help, I will give my all to support you in any way I can. Once again, my deepest and most heartfelt thanks to my two dearest supervisors!

I would also like to extend my heartfelt gratitude to my dear friends and fellow PhD companions, Jiarun Ye and Xiaobo Li. Over these years in a foreign land, we have supported and accompanied each other through every challenge and milestone. We are not just fellow researchers, we are the best teammates on this academic journey!

I would also like to express my gratitude to all the faculty members at the Institute of Education at Reading who have supported me in various ways. Special thanks to Rowena, Kari, and Allen. It has been a truly enjoyable experience working with you. Your support and kindness have made my academic journey all the more fulfilling, and I have cherished every moment of my time here.

I am deeply grateful to my beloved family. Thank you to my father, Fuzhong Yang, and my mother, Fengmei Yao, for your unwavering financial and emotional support. Without your sacrifices and trust, I would not have been able to reach where I am today. I love you both!

A heartfelt thank you to my partner, Junwei Zhu. Thank you for your trust, your constant support from afar, and your patience in waiting for me. I am also grateful to Junwei's parents for their understanding and support.

For me, these years of doctoral study have been both simple and challenging. Looking back on every single day of hard work over the past few years: studying for more than 14 hours a day, reading countless research papers, feeling the anxiety after receiving feedback, the preparation and nervousness before every meeting with my supervisors, the frustration of writer's block, pursuing a PhD seems like an incredibly difficult journey. But I persevered. So today, I also want to thank myself, for my hard work, dedication, and resilience, for never

giving up. I will forever cherish these years of struggles and joy, a journey that has shaped me into who I am.

Finally, thank you, the University of Reading! Thank you, the IoE! I will always love you, my alma mater!

Table of Contents

Chapter 1. INTRODUCTION	1
1.1 Identifying the Problem	2
1.1.1 The Significance of Listening for Chinese University EFL Learners	2
1.1.2 The Challenges Facing L2 Listening Development Among Chinese University EFL Learners	6
1.2 Listening in Informal Digital Learning of English (IDLE) Context: An Underexplored Area	11
1.3 Self-Regulated Learning and Listening	15
1.3.1 The SRL Model Targeting L2 Listening	16
1.3.2 Establishing the Joint Predictive Mechanisms of Multiple Factors for L2 Listening ...	17
1.4 Significance and Aims of the Study	18
1.5 Thesis Overview	21
Chapter 2. LITERATURE REVIEW	23
2.1 Introduction	23
2.2 Listening	23
2.2.1 What Is Listening?	23
2.2.2 How Does Listening Work?	25
2.2.3 Personal Factors Influencing L2 Listening Comprehension	29
2.3 Self-Regulated Learning (SRL)	38
2.3.1 What Is Self-Regulated Learning?	38
2.3.2 How Does SRL Work? - SRL Models	41
2.3.3 Self-Regulated L2 Listening Model	46
2.3.4 Mechanisms of Joint Prediction for Listening: A Social Cognitive Theory (SCT) Perspective	53
2.4 Listening in Informal Digital Learning of English (IDLE) Context	61
2.4.1 An Ecological Perspective on Contextual Factors	62
2.4.2 Defining Informal Digital Learning of English	64
2.4.3 Current Research Status of IDLE	65
2.4.4 Indicators for Evaluating IDLE Engagement	71
2.5 Identifying Research Gaps	75
2.6 Main Research Questions	77
Chapter 3. METHODOLOGY	79
3.1 Research Design	79
3.1.1 Ontology, Epistemology, and Methodology	79
3.1.2 Rationale for Adopting Mixed Method and Correlational Research Design	82
3.2 Participants	85
3.2.1 Selection of Participants	85
3.3 Data Collection: Research Instruments and Procedures	91
3.3.1 English Listening Questionnaire	92
3.3.2 IDLEL E-Logs	98
3.3.3 Listening Comprehension Tests	103
3.4 Pilot Study	112
3.4.1 Procedure of Pilot Study	113

3.4.2 Data Analysis and Results for the Pilot Study	114
3.4.3 Reliability and Validity	118
3.4.4 Lessons & Implications	122
3.5 Data Analysis	123
3.6 Ethical Considerations	145
Chapter 4. RESULTS	149
4.1 Introduction	149
4.2 RQ1: Validation of the Self-Regulated L2 Listening Model	150
4.3 RQ2: Validation of the Hypothesised Joint Predictive Mechanisms	153
4.3.1 CFA for Questionnaire Validation: LSEQ and LLAS	153
4.3.2 SEM for the Validation of the Hypothesised Joint Predictive Mechanisms	157
4.4 RQ3: The Characteristics of Participants' IDLEL Engagement	163
4.4.1 Diversity, Frequency and Duration: Descriptive Analysis	163
4.4.2 Strategy Use: Thematic Analysis and Descriptive Analysis	171
4.4.3 Cluster Analysis	182
4.5 RQ4: The Prediction of IDLEL Engagement and SRL on L2 Listening, Listening Anxiety and Self-Efficacy	189
4.5.1 The Prediction of IDLEL Engagement and SRL on L2 Listening	190
4.5.2 The Prediction of IDLEL Engagement and SRL on Self-Efficacy	196
4.5.3 The Prediction of IDLEL Engagement and SRL on Listening Anxiety	203
4.6 RQ5: The Moderating Effect of SRL on the IDLEL Engagement and L2 Listening, Listening Anxiety, and Self-Efficacy	206
Chapter 5. DISCUSSION	210
5.1 Introduction	210
5.2 Discussion for RQ1: Validation of the Construct of the Self-regulated L2 Listening Model	211
5.3 Discussion for RQ2: Validation of the Joint Predictive Mechanisms	213
5.3.1 The Mediating Role of SRL and Listening Anxiety	213
5.3.2 The Mediating Role of Self-Efficacy and Listening Anxiety	217
5.4 Discussion for RQ3: Characteristics of IDLEL Engagement	223
5.4.1 The Quantity of IDLEL Engagement	223
5.4.2 The Quality of IDLEL Engagement	225
5.4.3 Purposes and Benefits of IDLEL Engagement	231
5.4.4 IDLEL Engagement Patterns	232
5.5 Discussion for RQ4: The Prediction of IDLEL Engagement and SRL on L2 Listening, Self-Efficacy, and Listening Anxiety	234
5.5.1 The Predictive Effects of IDLEL Engagement and SRL on L2 Listening	234
5.5.2 The Predictive Effects of IDLEL Engagement and SRL on Self-Efficacy	236
5.5.3 The Predictive Effects of IDLEL Engagement and SRL on Listening Anxiety	238
5.6 Discussion for RQ5: The Moderating Effect of SRL on the IDLEL Engagement and L2 Listening, Listening Anxiety, and Self-Efficacy	239
Chapter 6. CONCLUSION	242
6.1 Introduction	242
6.2 Summary of the Study	242
6.3 Summary of the Findings	243

6.3.1 Validation of the Hypothesised Self-Regulated L2 Listening Model and the Joint Predictive Mechanisms	243
6.3.2 The Characteristics of IDLEL Engagement	245
6.3.3 The Predictive Effects of IDLEL Engagement and SRL on L2 listening, Self-Efficacy, and Listening Anxiety	248
6.3.4 The Moderating Role of Self-regulation Between IDLEL Engagement and L2 Listening, Self-efficacy, and Listening Anxiety	250
6.4 Contributions of the Study	251
6.4.1 Empirical Contributions	251
6.4.2 Theoretical Contributions	253
6.4.3 Methodological Contributions	255
6.4.4 Pedagogical Contributions and Implications	257
6.5 Limitations and Future Research	260
6.6 Conclusion	263
REFERENCES	265
APPENDICES	280
Appendix 1: PowerPoint for Participant Recruitment	280
Appendix 2: English Listening Questionnaire	281
Self-Regulated L2 Listening Questionnaire (SRLLQ)	281
L2 Listening Anxiety Scale (LLAS)	285
Listening Self-Efficacy Questionnaire (LSEQ)	287
Appendix 3: IDLEL E-Log	290
Appendix 4: English Listening Tests	301
Listening Test 1	301
Listening Test 2	305
Listening Test 3	309
Audioscripts For Test 1 & Test 3	313
Audioscripts For Test 2	317
Appendix 5: Main Ideas for the Free-Recall Task	321
Test 1	321
Test 2	323
Appendix 6: The Codebook For the Thematic Analysis Purpose	324
Codebook for the Task-Level Strategy	324
Codebook for the Motivation & Affect-Level Strategy	326
Appendix 7: Ethical Approval & Consent Forms	327
Ethical Approval Form	327
Information Sheet and Consent Form (Students)	337
Information Sheet and Consent Form (Teachers)	342
Information Sheet and Consent Form (Head of the faculty)	346

Chapter 1. INTRODUCTION

This study explores the role of informal digital learning of English listening (IDLEL) activity engagement (i.e. frequency, diversity, duration, and strategy-use) and self-regulation in the development of Chinese undergraduates' L2 listening comprehension, listening self-efficacy, and listening anxiety. In the context of China, research of this kind can be particularly important. This is because Chinese university EFL learners are facing various challenges in developing their L2 listening proficiency and self-regulated learning (SRL) abilities within traditional teacher-oriented and exam-oriented L2 classrooms, and study of this kind may offer insights into addressing these difficulties.

The present research also investigates the structure and mechanism of self-regulated L2 listening. Research like this can be beneficial for enriching and extending the theory of SRL in the field of language learning, as previous SRL studies have paid relatively little attention to L2 listening skill, and the mechanism of self-regulated L2 listening has remained unclear. Moreover, the current research also examines the joint predictive mechanisms of listening self-efficacy, listening anxiety, and self-regulation on listening comprehension, in which self-regulation plays different roles. A study of this kind can be essential for advancing theories of L2 listening, as the joint predictive mechanisms and pathways of motivational, affective, and cognitive factors on L2 listening have not been well established.

This introductory chapter will begin with highlighting the importance of listening for Chinese university EFL learners, followed by identifying the challenges they face in improving L2 listening within classroom settings. The chapter will then introduce a potential solution to these challenges: engaging in listening practice activities within the context of

informal digital learning of English. Subsequently, this chapter will focus on SRL, which is crucial for both L2 listening and informal language learning, explaining the necessity of proposing an SRL model target L2 listening and further demonstrating the significance of exploring the joint predictive mechanism of SRL, self-efficacy, and anxiety on listening comprehension. Finally, the chapter will conclude with a summary outlining the structure of the thesis.

1.1 Identifying the Problem

1.1.1 The Significance of Listening for Chinese University EFL Learners

For all L2 learners, the significance of listening lies in its role as a source of input for the development and even emergence of other language skills (Goh & Vandergrift, 2012; Indrasari, 2019; Oxford, 1993; Rost, 2013), such as speaking, which can be regarded as the main pathway to second language acquisition (SLA) (Gilakjani, 2016; Krashen et al., 1984). For EFL learners across various educational stages in China, where the education system is exam-oriented, however, the significance of listening may extend far beyond this.

In 2001, the Ministry of Education (MOE) mandated the nationwide implementation of English curriculum in primary schools as a crucial component of basic education reform (covering primary to secondary education) in the early 21st century. This policy established English as a compulsory subject for most students from Year Three onwards in basic education, with some economically advanced regions potentially introducing it earlier (Silver et al., 2002). According to the 2011 version of the Ministry of Education's curriculum standards, schools are required to ensure that the total weekly English instruction time for students in Grades Three to Six is no less than 80 to 90 minutes and guarantee 20 to 25

minutes of in-class audio-visual exposure to English per week (English Curriculum Standard in MOE, 2011). The revised 2022 version of the curriculum standards not only recommend that compulsory education students (primary and junior secondary) engage in no less than 30 minutes of extracurricular English audio-visual activities weekly but also advocate the creation of conditions for organising listening exams, and that the proportion of listening test items is reasonably adjusted and gradually increased (English Curriculum Standard in MOE, 2022). This highlights the growing importance of listening performance in achieving high English test results.

Indeed, based on the researcher's personal experience as an L2 learner, English listening is a mandatory tested skill in various English examinations taken by most basic education stage students. It is particularly crucial for senior high school students preparing for the National College Entrance Examination (NCEE), as achieving a good score in the English section may depend largely on their listening proficiency. Before the State Council issued the *Implementation Opinions on Deepening the Reform of the Examination and Enrollment System* in 2014, the exam policies and formats for university entrance exams differed across provinces. In some provinces, such as Henan, Shanxi, and Liaoning, listening scores were not included in the total NCEE English score but were provided to universities as a reference during the application process. From the introduction of the NCEE reform till now, however, 29 provinces, autonomous regions, and municipalities, including those that previously excluded English listening scores from the total, have successively announced the use of a unified national English test format, incorporating the listening comprehension scores (30 points) into the total English score (150 points). This demonstrates that the importance of

listening for English learning among basic education stage students is becoming increasingly prominent.

For EFL learners in higher education, English listening comprehension represents an especially crucial L2 skill. At the undergraduate level in China, English courses are primarily categorised into two types: General College English for non-English major undergraduates and Specialist College English for English major undergraduates (Silver et al., 2002). General College English is typically divided into compulsory and elective stages. Students are generally required to complete the compulsory English courses during their first academic year. These compulsory English courses are similar to those in basic education, often taking the form of integrated English courses that combine multiple language skills, including listening, within a single class. Upon meeting specific English proficiency requirements, students may opt to participate in elective English courses. The primary criterion for determining their eligibility for these elective courses is usually students' performance in the College English Test (CET) Band Four and Band Six written examinations. These nationwide standardised tests hold particular significance, as it is common practice for institutions to link the attainment of bachelor's degrees to students' CET Band Four scores (Silver et al., 2002). Notably, the listening comprehension section accounts for as much as 35% of the total score (Wu et al., 2022), highlighting the pivotal role of listening proficiency in achieving a strong English performance for non-English major undergraduates.

Different from non-English major students, English majors from foreign studies universities (focusing on language teaching and research), comprehensive universities (covering various majors, such as arts, sciences, engineering, and medicine), normal

universities (focusing on teacher training), and teacher training colleges (specialising in primary and secondary teacher education), typically undergo specialised training in various English skills. Consequently, English listening is a compulsory module for English major students (Silver et al., 2002). In addition to the CET Band Four and Band Six, English majors are usually required to participate in subject-specific assessment tests, namely the Test for English Majors (TEM) Band Four and Band Eight. Although there is no uniform policy across universities regarding whether passing the TEM exams is a prerequisite for obtaining a bachelor's degree in English, the TEM exams serve as an important measure of English majors' professional knowledge and competence, holding significant value for both of their academic pursuits and career opportunities. In the TEM-4 examination, the listening component constitutes 30% of the total score, comprising 10% for dictation and 20% for listening comprehension; for TEM-8, listening comprehension accounts for 25% of the total marks (Jin & Fan, 2011), making it a crucial factor for English majors to successfully pass these exams.

Additionally, the adoption of English as a Medium of Instruction (EMI) in Chinese higher education has expanded at an unparalleled rate over the last twenty years, with an increasing number of Chinese students enrolling in various types of EMI programmes (Zhou & Rose, 2022). For example, there are fully English-taught programmes primarily for domestic students, bilingual courses for domestic students, and specialised courses for English majors, conducted entirely in English or bilingually (Zhou & Rose, 2022). EMI courses place higher demands on the English listening proficiency of Chinese university students.

Beyond assessment and course requirements, listening proficiency also holds substantial importance for Chinese university EFL learners in terms of further education and professional development. For instance, an increasing number of Chinese undergraduates choose to pursue postgraduate degrees abroad. English listening is therefore not only essential for achieving the required scores in international language proficiency tests such as IELTS and TOEFL, which are prerequisites for overseas study, but also serves as an indispensable skill for daily life and academic activities abroad. Moreover, in addition to some English-related professions (e.g., translators and teachers), the expansion of Chinese enterprises into overseas markets and the growing presence of foreign companies in China have also created a demand for job applicants with competent English listening proficiency. Taken together, English listening proficiency plays a vital role in the learning and career prospects of Chinese EFL learners particularly at the university level.

1.1.2 The Challenges Facing L2 Listening Development Among Chinese University EFL Learners

The importance of listening for second language (L2) learners is intertwined with the difficulty it presents, making it one of the most critical language skills, especially for Chinese university EFL learners. The difficulty of listening stems from various factors, including the inherent complexity of listening itself, personal factors, and the contextual factors.

1.1.2.1 The Complexity of Listening

Listening is an active dynamic process of receiving, attending, perceiving, interpreting, and responding to the listening input, encompassing a series of stages (Oxford, 1993; Purdy, 1997; Rost, 2005). Anderson (1995) described listening comprehension as a three-stage

process, including perception, parsing, and utilization stages. From the perception stage through to the utilisation stage, listeners construct the meaning of the speech input by employing both top-down and bottom-up processing (Mendelsohn, 1994; O’Malley et al.; 1989; Oxford, 1993). Briefly stated, in bottom-up processing, listeners focus on the linguistic input itself at the level of sounds, words and phrases, while in top-down processing they draw on non-linguistic knowledge such as world knowledge to comprehend (see Section 2.2.2). One source of listening difficulty, thus, lies in the need to orchestrate, in real time, a range of linguistic and non-linguistic factors (Rost, 2013). Additionally, the transience and uncontrollability of auditory input, where listeners “have no control over what is going to be said, how it is going to be said, and how quickly it is going to be said” (Mendelsohn, 1994, p. 9), as well as the unobservable nature of the listening comprehension process, make listening appear more “uncontrollable” than other language skills (Graham, 2011), thereby increasing its difficulty.

1.1.2.2 Personal Factors Influencing Listening

Regarding the factors influencing listening comprehension, insufficient vocabulary knowledge and limited working memory capacity may cause listening difficulties at various stages of listening comprehension (Goh, 2000). Beyond these two, personal factors such as motivational beliefs (e.g., self-efficacy), and affective responses (e.g., listening anxiety), may also exert significant influences on L2 listening (Canaran et al., 2024; Xu & Huang, 2018) (see Section 2.2.3).

Self-efficacy refers to learners’ judgment of their capability to perform the courses of action necessary to achieve a designated level of performance (Bandura, 1986). As an

important component of the construct of motivation, self-efficacy beliefs can create expectations regarding the outcome of learners' efforts and in turn affect their choice of tasks, levels of effort and persistence exerted, as well as their thought patterns and affective responses, including anxiety (Bandura, 1986). For example, learners may avoid listening tasks they perceive as beyond their capabilities while actively participating in those they feel confident in handling. Listening seems more difficult to control than other language skills (Graham, 2011), due to the inherent complexity and uncontrollable nature of auditory input, the transience of information, and the unobservable nature of the listening comprehension process. Self-efficacy, therefore, can be particularly important for L2 listening, with its direct and indirect predictive effects on listening being confirmed by existing studies (Canaran et al., 2024; Du & Man, 2023; Payaprom, 2023; Zhang & Xu, 2024).

According to Social Cognitive Theory (SCT), perceived self-efficacy in controlling potential threatening events plays a crucial role in anxiety arousal (Bandura, 1990). Therefore, a sense of lack of control over listening may lead to learners' anxiety, which relates to "negative feelings such as uneasiness, frustration, self-doubt, apprehension and tension" (Xu, 2011, p. 1709), in the face of L2 listening difficulties. The aroused anxiety may cause learners to shift their focus from problem-solving to self-doubt and concerns about potential failure, thereby reducing their willingness to invest effort in overcoming difficulties and completing tasks (Bandura, 1990), which may in turn directly or indirectly affect listening comprehension (Golchi, 2012; Xu, 2017; Xu & Huang, 2018).

While the importance of vocabulary size and working memory capacity for listening have been acknowledged (Du et al., 2022; Du & Man, 2022; Goh, 2000; Kim et al., 2022;

Masrai, 2020; Satori, 2021; Wallace & Lee, 2020), personal factors are believed to demonstrate stronger explanatory power across different listening contexts (Horwitz et al., 1986; Oxford, 2017) and hold greater potential for improvement through interventions, making them worthy of greater attention (Graham & Macaro, 2008; Fathi et al., 2020). Additionally, the combination of all the factors mentioned above has resulted in L2 listening being perceived as a challenging language skill to improve (Graham, 2018). This issue can be even more pronounced for Chinese university EFL learners (Lei & Hu, 2014; Wang, 2011; Zeng & Goh, 2018), as, in addition to the influence of these factors, their L2 listening development also faces numerous challenges arising from contextual factors.

1.1.2.3 Contextual Factors Influencing Listening: Limitations of Listening Instruction in Chinese Universities

Firstly, many Chinese tertiary-level EFL learners receive limited in-class listening instruction (Zeng & Goh, 2018). This issue can be particularly severe for non-English major students compared to their English-major counterparts, as they do not have courses specifically designed to improve their English listening, as mentioned earlier. Based on my personal experience as a non-English major, teachers usually devote a significant amount of time to training students in reading and writing rather than listening in integrated English courses.

Secondly, for both non-English majors and English majors, listening courses are mainly taught in a single format: students listen to listening materials and complete the corresponding listening exercises, then the teacher gives the answers (Zhu, 2021). Influenced by an examination-oriented education system, listening instruction tends to emphasise the outcome

rather than the process of listening, with the primary goal of preparing students for the CET and/or the TEM exams (Renandya & Hu, 2018). Teachers may teach specific listening strategies for tests or question-answering strategies based on the listening materials students have just listened to. However, the lack of systematic instruction in listening strategies may make it difficult for students to achieve significant improvement in their listening proficiency within the repetitive cycle of listening to recordings, answering comprehension questions, and checking answers.

Thirdly, inauthentic listening materials are used extensively in English textbooks, classes, and examinations in Chinese universities (Zhu, 2021). Although some listening textbooks, such as *New Horizons College English: Listening and Speaking Course*, incorporate original BBC audio and video materials, these resources still present issues such as a mismatch between difficulty levels and students' listening proficiency (Xu, 2024; Zhang, 2009), as well as outdated topics. Moreover, in teacher-centred English classrooms in Chinese universities (Shen & Bai, 2024), learners typically lack opportunities to select and control listening materials (e.g., pausing, replaying, or adjusting playback speed), according to my personal learning experience. A lack of control over listening and insufficient knowledge of listening strategies may undermine learners' confidence in English listening, leading to listening anxiety and affecting their listening performance. In this case, Chinese university EFL learners, whether English majors (Gu, 2019) or non-English majors (Wang, 2023), have been found to experience relatively low levels of listening confidence and moderate to high levels of English listening anxiety. Moreover, non-English majors have been reported to exhibit higher levels of general listening anxiety (i.e., anxiety arising from classroom, daily usage,

and media learning) compared to English majors (Li et al., 2023). One potential reason for this could be that non-English majors receive less listening instruction than their English major counterparts.

Taken together, listening, as a critical L2 skill for Chinese university EFL learners, has not received sufficient attention and support in formal L2 listening instructional settings. Understanding language acquisition through the lens of ecology, which examines the interactions between organisms and their surrounding elements within an ecosystem, it relies not only on learners' efforts but is also influenced by environmental factors, resulting from the continuous interaction between learners and their learning contexts (Menezes, 2011). Different learning contexts provide different types of affordances, including interaction opportunities, resources, demands, constraints, and obstacles (Shotter & Newson, 1982). Therefore, when the affordances provided by a single context, such as the traditional EFL classroom, fail to meet learners' language development needs, exploring and utilising new affordances in other contexts, such as out-of-class informal language learning contexts, may serve as a potential solution (see Section 2.4).

1.2 Listening in Informal Digital Learning of English (IDLE) Context: An Underexplored Area

From a lifelong learning perspective, the majority of people's learning is achieved in informal learning that can provide abundant learning resources and ample practice opportunities (Boileau, 2018). The advancement and popularisation of internet and digital technologies have further strengthened the advantages of informal learning, making it even more significant for learners (Lee, 2019a). Informal Digital Learning of English (IDLE) has

emerged as a result of this trend (Lee, 2019a), becoming an essential component of technology-assisted language learning.

IDLE activities can be categorised into two main types. The first type, referred to as extracurricular activities, is characterised by being “self-directed and self-instructed” (Lee, 2019b, p. 115) but “still linked to a formal language program” (p. 115). The second type, which has been the focus of much existing research and is also the focus of this study, indicates extramural activities (Lee, 2019b), defined as “self-directed English activities in informal digital settings, motivated by personal interests, undertaken independently, and not assessed by a teacher” (Lee & Lee, 2021, p. 359). Accordingly, listening practice conducted in IDLE contexts can be referred to as Informal Digital Learning of English Listening (**IDLEL**). It refers to self-directed English listening activities carried out in unstructured, out-of-class digital environments, driven by personal interests, independent of formal language learning programs, and without being assessed by a teacher (Lee & Lee, 2021) (see Section 2.4.2).

Learners’ engagement in IDLE has attracted increasing attention from researchers (Lee, 2019a, b; Lee & Dressman, 2018). Engagement refers to “the student’s psychological investment in and effort directed toward learning, understanding, or mastering the knowledge, skills, or crafts that academic work is intended to promote” (Newmann, 1992, p. 12). Engagement is considered a prerequisite for language processing and a key factor in L2 development; therefore, it plays a crucial role in L2 learning (Reinders & Nakamura, 2021). Engagement is a multidimensional construct, encompassing behavioral (i.e., the quantity and quality of task engagement), cognitive (i.e., deliberate allocation and maintenance of attention,

intellectual effort, use of learning strategies, and self-regulation), affective (i.e., learners' affective responses and subjective experiences during learning activities), and social dimensions (i.e., interaction and connection among learners) (Reinders & Nakamura, 2021; Zhou et al., 2021). Given its complexity, engagement may influence L2 learners in multiple ways.

Current research tends to focus on learners' behavioral engagement in IDLE. For instance, extant research typically analyses learners' IDLE engagement in terms of its quantity and quality (Lee, 2019a, b; Lee & Dressman, 2018). Specifically, the duration of learners' IDLE engagement over a period (e.g., daily, weekly, or monthly) is often used as an indicator to assess the quantity of IDLE engagement, directly reflecting the extent of learners' investment in informal language learning; the variety of activities engaged in is commonly employed to evaluate the quality of learners' IDLE engagement, revealing the breadth of their informal language learning experiences (see Section 2.4.4).

Understanding the development of L2 learners' listening in the context of IDLE can provide insights for their long-term listening development in the digital era. However, existing studies predominantly focus on the relationship between learners' IDLE engagement attributes (i.e., quality and quantity) and vocabulary as well as between IDLE engagement and speaking (Lee, 2019a, b; Lee & Dressman, 2018), with limited attention paid to listening (Sylvén & Sundqvist, 2012) (see Section 2.4.3). A few studies have highlighted the importance of IDLE engagement (i.e., playing English digital games) for the L2 listening development of young Swedish EFL learners (i.e., fifth grade) (Sylvén & Sundqvist, 2012). Given differences in learners' learning contexts, age characteristics, language aptitude, as

well as the types and ways of engaging in IDLE activities, however, the findings of existing studies may offer limited insights into Chinese university students' IDLEL engagement and their L2 listening development within the IDLEL context. Therefore, further research is needed to investigate the relationship between IDLE engagement and L2 listening among Chinese university EFL learners.

In addition to L2 proficiency, extant studies have also shed light on the influence of IDLE activities on participants' affective factors, such as foreign language anxiety, language learning enjoyment, and willingness to communicate, as well as motivational factors, such as L2 confidence (Lee, 2019b; Lee & Dressman, 2018). To our knowledge, however, almost no studies have investigated the relationship between university EFL learners' IDLEL engagement and their self-efficacy and listening anxiety, which are important factors influencing L2 listening, as discussed earlier. Moreover, almost no research has explored learners' affective engagement in IDLEL, namely their learners' affective responses and subjective experiences during IDLEL activities.

Another research direction that has been largely overlooked in extant studies is EFL learners' self-regulated learning (SRL) in IDLEL engagement, namely learners' cognitive IDLEL engagement. However, examining SRL in the context of informal language learning is particularly relevant for contexts such as China, which serves as the setting for the present study. In China's traditional teacher-centred, exam-oriented language classrooms (Shen & Bai, 2024; Zhao et al., 2013), external regulation by teachers usually predominates (Boekaerts & Minnaert, 1999). This means that teachers typically control the learning process, such as goal setting, selection of learning resources, learning pace, immediate feedback, and standardised

assessment methods. In such a context, although learners may experience low autonomy and freedom in their learning process (Lee et al., 2009), they may also have a low demand for self-directed planning and strategy adjustment, as most learning decisions are made by the teacher. By contrast, in informal learning contexts, learners often lack clear external regulation from authoritative figures such as teachers. While they enjoy greater autonomy and freedom in learning, they must also take on greater responsibility for their own learning, and the importance of SRL may become even more pronounced in these contexts.

1.3 Self-Regulated Learning and Listening

As aforementioned, beyond the inherent complexity of listening, motivational factors (e.g., self-efficacy) and affective responses (e.g., listening anxiety) may also influence learners' L2 listening comprehension, and the ability to self-regulate multiple influencing factors can thus be considered one of the essential skills for successful L2 listening (Goh & Vandergrift, 2012; Pintrich, 2000).

Defined as “self-generated thoughts, feelings, and actions for attaining one’s learning goals” (Zimmerman & Moylan, 2009, p. 299), self-regulated learning (SRL) indicates the process by which “students exercise agency by consciously controlling and intervening in their learning” (Winne & Hadwin, 2012, p. 297). Successful L2 listeners, namely self-regulated listeners, therefore, are those who “set goals for their learning and then attempt to monitor, regulate, and control” (Pintrich, 2000, p. 453) their cognition, metacognition, motivational beliefs, affective responses, and behaviours during listening, until their learning goals are achieved. Compared to other L2 skills such as reading, writing, and vocabulary, listening has received less attention in self-regulation research (Zhou et al., 2024). However, a

substantial body of empirical research has confirmed the positive relationship between SRL and L2 listening comprehension (Fatemi et al., 2014; Nasrollahi-Mouziraji & Birjandi, 2016; Xu & Luo, 2024; Yabukoshi, 2024; Zeng & Goh, 2018), indicating that SRL is another crucial factor affecting L2 listening (see *Section 2.3.1*).

1.3.1 The SRL Model Targeting L2 Listening

SRL, as a metacognitive activity at the highest level (Borkowski, 1996), involves multiple dimensions and phases, with its process elaborated in various SRL models. Most SRL models have been designed to explain its mechanisms in general learning contexts rather than focusing on specific domains, such as L2 learning. This means that they may fail to fully capture the uniqueness of self-regulated L2 learning, even though some of these models have been applied in L2 research, either to guide the design of SRL interventions and the development of questionnaires, such as Zimmerman's (2000) Cyclical Phases model (Çakiroğlu, 2021; Morsedian et al., 2017; Teng & Zhang, 2016; Zhou et al., 2024), or to demonstrate how cognitive conditions (e.g., motivational factors and domain knowledge) influence the entire SRL process by affecting the accuracy of task analysis, such as Winne and Hadwin's (1998) COPES model (Ranalli, 2012). A few models target L2 learning, such as Oxford's (2017) S²R model. However, they fail to capture the complexity and characteristics of self-regulated L2 listening specifically. Additionally, while some existing models stress the importance of motivation and affect in SRL, they do not adequately address how these motivational and affective factors are regulated (e.g., Winne & Hadwin's (1998) COPES model). Some models, such as Efklides' (2011) MASRL model and Oxford's (2017) S²R model, acknowledge the conditions under which motivation and affect regulation may

occur but fail to detail the specific processes involved. Others, such as Pintrich's (2000) SRL framework, have considered motivation/affect as an independent aspect of self-regulation but show it as overlapping with the regulation of other aspects, increasing model complexity (see *Section 2.3.2*).

The limitations of existing SRL models have resulted in an unclear understanding of the structure and mechanism of self-regulated L2 listening. Clarifying this mechanism may hold potential benefits for addressing the challenges of L2 listening development, addressing the limitations of current SRL models, enriching and expanding SRL theory by considering its application within a specific language skill domain (i.e., listening), and guiding self-regulated L2 listening educational practices. Therefore, a new model specifically designed to depict the mechanism of self-regulated L2 listening needs to be constructed and validated to fill this research gap.

1.3.2 Establishing the Joint Predictive Mechanisms of Multiple Factors for L2 Listening

SRL and other key factors influencing L2 listening, such as self-efficacy and listening anxiety, share a close and complex relationship. According to SCT, self-efficacy plays a critical role in triggering anxiety and also impacts SRL sub-functions (e.g., goal setting and attribution) (Bandura, 1990, 1991). Moreover, self-efficacy and anxiety are both contributors to and outcomes of SRL (Winne & Hadwin, 1998). While the individual effects of these factors on L2 listening have been extensively confirmed by empirical research, as discussed earlier, the mechanisms and pathways through which these three factors jointly predict L2 listening comprehension remain unclear. Some studies have made an attempt to address this gap (Zhang & Xu, 2024), exploring the mediating role of metacognitive awareness (a key

component of SRL) between self-efficacy and listening, as well as between listening anxiety and listening. However, while metacognitive awareness is an essential component of SRL (Efklides, 2011), it is not synonymous with SRL. Additionally, existing research has only considered the predictive effects of motivational and affective factors on listening when cognitive factors (i.e., metacognitive awareness) act as mediators; it has yet to examine the predictive effects of cognitive factors, such as SRL, on listening when motivational and affective factors serve as mediators. Therefore, to comprehensively understand the joint predictive mechanisms of self-efficacy, listening anxiety, and SRL on L2 listening, and to provide theoretical and practical guidance for self-regulated L2 listening, it is essential to explore these mechanisms within a single empirical framework where these factors play different roles.

1.4 Significance and Aims of the Study

Firstly, no research, so far, has proposed and validated an SRL model targeting L2 listening. From a theoretical perspective, the construct of a self-regulated L2 listening model can expand and enrich existing SRL theories. Current SRL models have several limitations that require improvement. For instance, while these models have broad applicability, they lack specificity and provide insufficient explanations for motivational and affective self-regulation. Thus, the development and validation of a new model can not only advance existing SRL models but also extend SRL theory into the field of L2 learning. From a practical perspective, the development and validation of a self-regulated L2 listening model can provide theoretical foundations, guidance, and resources for language educators and learners in their self-regulated L2 listening educational practices.

Secondly, to the researcher's knowledge, no study, within a single research framework, has demonstrated a comprehensive understanding of the mechanisms through which L2 listening proficiency is jointly predicted by self-efficacy, anxiety, and self-regulation, with SRL plays different roles in. Exploring this issue can not only offer new perspectives and empirical evidence for understanding how motivational, affective, and cognitive factors jointly predict L2 listening, but also contribute to understanding the importance of SRL for L2 learners' listening development and underscores the necessity of integrating SRL theory into the domain of L2 listening. Moreover, the establishment and validation of the joint predictive mechanisms for listening can help language educators and learners better understand the interplay between self-efficacy, listening anxiety, and SRL in influencing L2 listening. This understanding can in turn guide the design of targeted teaching and learning strategies to effectively enhance learners' immediate and long-term listening performance.

Thirdly, to date, no such research has examined the relationship between the quantity (i.e., duration and frequency) and quality (i.e., diversity and strategy use) of university EFL learners' IDLEL engagement (including that of Chinese university EFL learners), and the changes in their L2 listening comprehension, listening self-efficacy, and listening anxiety within a single empirical framework. Additionally, no research has, to this researcher's knowledge, explored the self-regulation behaviours of university EFL learners (including Chinese university EFL learners), in their IDLEL engagement, or the role of SRL in the changes of their L2 listening comprehension, listening self-efficacy, and listening anxiety. Theoretically, exploring these issues can both enrich the theoretical framework of IDLEL and extend the theories of L2 listening, SRL, self-efficacy, and anxiety into the informal learning

domain, thereby refining the theoretical landscape of SLA. Practically, exploring learners' L2 listening, SRL, self-efficacy, and anxiety in the context of IDLEL can help educators design interventions to encourage learners to strategically engage in IDLEL activities, using digital learning resources outside the classroom to improve their L2 listening proficiency and SRL and to boost their listening self-efficacy, while alleviating their listening anxiety.

Taken together, the present study aims to fill the research gaps stated above, contributing to the understanding of the role of IDLEL activity engagement (i.e., quantity and quality), and self-regulation in the development of learners' listening comprehension, self-efficacy, and listening anxiety within the IDLEL context. There are six primary aims of the present study:

- 1) To understand the structure and mechanism of self-regulated L2 listening;
- 2) To understand the mechanisms through which L2 listening is jointly predicted by listening self-efficacy, listening anxiety, and self-regulation;
- 3) To explore the characteristics of Chinese undergraduates' IDLEL engagement, in terms of its quantity (i.e., frequency and duration) and quality (i.e., diversity and strategy-use);
- 4) To gain insights into the relationship between learners' IDLEL engagement (i.e., quantity and quality) and their English listening comprehension, listening anxiety and listening self-efficacy;
- 5) To gain insights into the relationship between learners' SRL and their English listening comprehension, listening anxiety and listening self-efficacy;
- 6) To investigate the extent to which learners' level of SRL moderates the relationship between IDLEL engagement (i.e., quantity and quality) and their L2 listening comprehension, listening anxiety, and listening self-efficacy.

1.5 Thesis Overview

Firstly, the *Introduction* Chapter above delved into the conflict between the importance of L2 listening for Chinese university EFL learners and the limited support provided by traditional L2 classrooms for their listening development, thereby highlighting the necessity and urgency of improving their L2 listening in informal learning contexts, such as the informal digital learning of English (IDLE) context. Since learners may not receive as much external regulation in informal language learning settings as they can in the classroom, the importance of self-regulation thus becomes more prominent in such context. Therefore, this chapter subsequently focused on self-regulated learning (SRL), revealing the significance of constructing an SRL model targeting L2 listening and clarifying the joint predictive mechanisms of self-regulation, self-efficacy, and listening anxiety on L2 listening.

Secondly, the *Literature Review* Chapter will provide a comprehensive introduction to L2 listening, including what listening is, the process of listening, and the personal factors that can influence L2 listening, with a particular focus on self-efficacy and listening anxiety. Following this, the chapter will offer a detailed discussion of another personal factor that plays a crucial role in learners becoming successful L2 listeners, namely SRL. Specifically, this chapter will focus on introducing the newly proposed SRL model targeting L2 listening and the joint predictive mechanisms of SRL, self-efficacy, and listening anxiety on L2 listening, developed within the framework of Social Cognitive Theory. Finally, this chapter will analyse the potential impact of contextual factors on L2 listening, addressing the necessity of improving L2 listening in informal language learning contexts, and provide an introduction to the basic concepts and current research status of IDLE. Research questions

will be proposed at the end of this chapter.

Thirdly, the *Methodology* Chapter will provide a detailed introduction to the research design, participants, data collection and analysis methods, as well as ethical considerations involved in the current study. Additionally, the pilot study conducted prior to the main study and its impact on the main research will be reviewed in this chapter.

Fourthly, the *Results* Chapter will present the detailed findings for each research question, including the validated structure of the self-regulated L2 listening model, the joint predictive mechanisms of SRL, self-efficacy, and listening anxiety on listening, the characteristics (i.e., quantity and quality) of participants' engagement in IDLEL activities, the prediction of IDLEL engagement and SRL on L2 listening, listening self-efficacy and listening anxiety, as well as the moderating role of SRL in the relationships between IDLEL engagement and L2 listening, listening self-efficacy and listening anxiety.

Fifthly, the *Discussion* Chapter will provide an in-depth analysis of the research findings presented in the previous chapter, examining whether these results align with the expectations and with the findings of existing research. If discrepancies arise, this chapter will further analyse and discuss possible reasons.

Finally, the *Conclusion* Chapter will make a summary of the whole study as well as the main findings of the study, discussing the potential contributions of the research from empirical, theoretical, methodological, and pedagogical perspectives. Moreover, the limitations of the study as well as the implications for future research will also be presented.

Chapter 2. LITERATURE REVIEW

2.1 Introduction

This chapter will delve into the theoretical and empirical evidence related to several key concepts of this study (i.e., listening, self-efficacy, listening anxiety, SRL, and IDLE) in the context of this study, which focuses on Chinese university EFL learners' L2 listening development. By reviewing relevant theory and empirical studies, this chapter will reveal gaps in existing research which will in turn lead to the formulation of the research questions. Specifically, the chapter will begin by introducing the concept of listening, discussing what listening is, how it works, and the factors influencing listening. Among these factors, self-efficacy and listening anxiety, which are also key concepts of this study, will be introduced in this section. Following this, SRL, another critical personal factor that affects listening and a focus of this study, will be introduced in a separate section. This section will discuss how SRL has been defined and the characteristics and limitations of previous SRL models. It will also present the study's newly proposed SRL model targeting L2 listening, as well as the joint predictive mechanisms of SRL, self-efficacy, and listening anxiety on L2 listening. Finally, the chapter will highlight the necessity of developing listening in informal learning contexts by explaining the potential impact of contextual factors on L2 listening, ending by defining IDLE and establishing its research status.

2.2 Listening

2.2.1 What Is Listening?

Listening, or listening comprehension, refers to the dynamic process of receiving, attending to, perceiving, interpreting, and responding to spoken input (Purdy, 1997; Rost,

2005). The purpose of listening is to construct meaning from the spoken input (Mendelsohn, 1994; O’Malley et al., 1989). Listeners use their linguistic knowledge (e.g., phonological knowledge, lexical knowledge, and syntactic knowledge), non-linguistic knowledge (e.g., prior knowledge, contextual knowledge, cultural knowledge, pragmatic knowledge, discourse knowledge), and strategic resources to transform the input they hear into internalised meaningful intake (O’Malley et al., 1989; Oxford, 1993, p. 205). When listeners use the above-mentioned knowledge and resources to construct meaning from the second language (L2) auditory input and transform it into internalised meaningful intake, they are involving in the process of L2 listening comprehension.

It has been argued that listening can have significant influence on the development or even emergence of other skills such as speaking (Goh & Vandergrift, 2012; Oxford, 1993; Rost, 2013). This is mainly because listening is an important way of offering L2 learners comprehensible input, which can be considered as the essential pathway to second language acquisition (Gilakjani, 2016). In language classrooms and daily conversations, listeners may have the opportunity to ask for clarification from or negotiate meaning with the speaker (Rost, 2013), ensuring that the input is comprehensible. When engaging in listening activities using digital devices, such as watching movies or listening to songs, learners may pause, replay, or adjust the speed of audio and video to aid their listening comprehension. However, in many other situations, such as academic lectures and listening tests, learners “have no control over what is going to be said, how it is going to be said, and how quickly it is going to be said” (Mendelsohn, 1994, p. 9). Therefore, listening is often regarded as a difficult language skill for L2 learners to improve (Graham, 2018), which is also the case for Chinese tertiary-level

EFL learners (Zeng & Goh, 2018).

2.2.2 How Does Listening Work?

There are two fundamental types of processing that listeners can employ for listening comprehension, namely bottom-up processing and top-down processing. Bottom-up processing refers to the way in which listeners construct meaning from the acoustic input in an incremental manner, namely from phonemes to words, phrases, sentences, and discourse (Goh & Vandergrift, 2012). If speech is considered as a code, bottom-up processing is the decoding process that requires listeners to draw on different forms of linguistic knowledge, including phonological knowledge, lexical knowledge, and syntactic knowledge. Bottom-up processing can help listeners to grasp the literal meaning of the speech input. To understand the implied meaning of the speaker, and when there are gaps in their linguistic knowledge or they cannot apply it, however, listeners need to employ top-down processing.

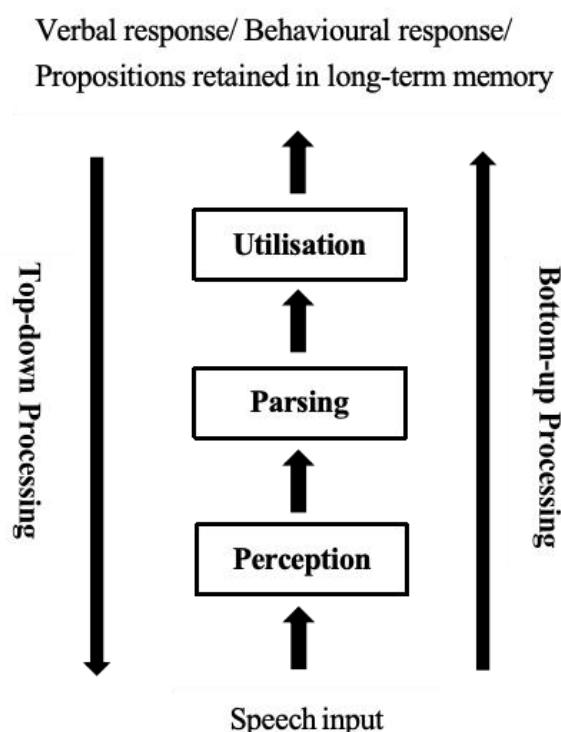
In contrast to bottom-up processing, top-down processing involves listeners inferring and interpreting the meaning of the speech stream by applying their non-linguistic knowledge stored in long-term memory. Such non-linguistic knowledge includes prior knowledge of the topic being addressed, contextual knowledge, cultural knowledge, pragmatic knowledge, and discourse knowledge (Goh & Vandergrift, 2012). Top-down processing requires listeners to initially establish expectations of the speech input and then activate relevant non-linguistic knowledge stored in memory. If listeners are unfamiliar with the topic or lack the relevant knowledge, listening comprehension may break down (Goh & Vandergrift, 2012).

Scholars have proposed different models to describe the process of listening comprehension, such as Anderson's (1995) three-stage language comprehension model, and

the listening comprehension models of Cutler and Clifton (1999) and Field (2013). Although the different models are distinctive, Anderson's (1995) model (see Figure 1) seems to contain key information found in various models and therefore can be a representative model for understanding the listening comprehension process.

Figure 1

The Three-Stage Listening Comprehension Model. Adapted from Anderson (1995)



Anderson's (1995) model consists of three listening comprehension phases, namely the perception, parsing, and utilisation stages. In the perception phase, listeners encode the acoustic/spoken message by processing it in memory (Gilakjani & Ahmadi, 2011; Goh, 2000). During this stage, the acoustic message is initially stored in the auditory sensory memory (also known as echoic memory) for a very short time. Listeners initially identify and segment the phonemes, which are the smallest units of speech that can distinguish between different spoken messages, from the continuous speech. They then recognize the speech signals as

words or other meaningful chunks of language in the form of phonetic representations while storing them in working memory for further processing (Gilakjani & Ahmadi, 2011; Goh & Vandergrift, 2012; Yeldham, 2019). At this phase, lexical knowledge is essential. Lack of phonological and lexical knowledge may result in listeners' inability to effectively identify the segmented phonemes as words and transfer them into working memory for further analysis (Goh, 2000). As working memory has limited capacity (Baddeley, 2003), information that cannot be transferred into working memory will quickly be forgotten, thus affecting listeners' listening comprehension (Goh, 2000).

In the parsing phase, the utterance retained in working memory is segmented according to syntactic and semantic cues as chunks. These chunks are then regrouped as propositions, which are the fundamental units of listening comprehension. Therefore, linguistic knowledge, such as syntactic and semantic knowledge are essential at this stage. Meanwhile, a meaning-based mental representation which is "an abstraction of the original word sequences" (Gilakjani & Ahmadi, 2011, p.980) is generated and held in working memory (Goh, 2000; Goh & Vandergrift, 2012; Yeldham, 2019). The limited capacity of working memory and the constant input of new information mean that to prevent loss of information, listeners have to transfer this processed information to long-term memory as soon as possible, which can also provide space for new information to be processed. Otherwise, listeners may experience difficulty because they quickly forget what they have heard (Goh, 2000). The storing of the information in long-term memory then occurs in the third stage, the utilisation stage.

To better interpret the intended or implied meanings of the input speech, listeners are expected to make inferences that go beyond the processed input (Anderson, 1995). In the

utilisation phase, listeners make bridging inferences to enhance their comprehension by retrieving non-linguistic knowledge (e.g., prior knowledge, contextual knowledge, cultural knowledge, pragmatic knowledge, and discourse knowledge) from long-term memory in the form of propositions and schemas, associating and matching them with the mental representations stored in working memory. The comprehension of the speech input can be achieved when the two sources of information match each other well. The representation of the speech in this case serves as the product of the listening comprehension process and listeners then can utilize this mental representation in practice. For example, if the input is an assertion, listeners can store the propositions in long-term memory for future use; if it is a question or an instruction, they may respond to the interlocutor verbally or in action (Gilakjani & Ahmadi, 2011; Goh & Vandergrift, 2012). A typical listening difficulty that learners may encounter at this stage is that they can understand words but cannot comprehend the intended message (Goh, 2000).

It should be noted that from the perception stage to the utilisation stage, top-down processing plays an increasingly prominent role. However, both types of processing can occur simultaneously at any one stage (Yeldham, 2019). Additionally, the three phases do not operate in an entirely linear fashion. In fully fluent, automated listening, the three stages can be operated concurrently in one listening event (Anderson, 1995; Gilakjani & Ahmadi, 2011).

Listening is considered a challenging L2 skill not only because, as discussed above, it is a complex process involving multiple stages, various ways of processing spoken input, and several different knowledge sources, but also because it can be influenced by a variety of factors, as discussed next.

2.2.3 Personal Factors Influencing L2 Listening Comprehension

As stated above, insufficient vocabulary knowledge and limited working memory capacity may lead to listening difficulties at various stages of listening comprehension (Goh, 2000). Their significant impact on listening comprehension has been confirmed among EFL learners from various L1 backgrounds (i.e., Arabic, Brazilian, Chinese, Iranian, Japanese, Korean) (Du et al., 2022; Du & Man, 2022; Kim et al., 2022; Masrai, 2020; Satori, 2021; Wallace & Lee, 2020). However, many other factors may influence listening comprehension, such as personal factors (e.g., motivational factors, affective responses, and self-regulation). Indeed, while acknowledging the role of vocabulary knowledge and working memory capacity, this study argues that understanding the influence of under-explored personal factors on learners' L2 listening comprehension may hold greater significance. Two reasons supporting this claim are discussed in detail below.

Firstly, it is believed that humans process visual information (e.g., images, videos) and verbal information (e.g., text, speech) through two independent channels (Mayer, 2014). The two information processing channels complement and support each other, thereby achieving effective comprehension of the input (Guichon & McLornan, 2008; Salmani & Rahimi, 2024). Most existing studies investigating the relationship between vocabulary knowledge and listening comprehension as well as working memory and listening comprehension have used listening comprehension tests that involve audio only or audio plus text input (Du et al., 2022; Du & Man, 2022; Kim et al., 2022; Masrai, 2020; Satori, 2021; Wallace & Lee, 2020). In the context of listening comprehension tests that lack the support of non-verbal cues, learners can consequently only rely on the verbal information processing channel to comprehend the input,

and the importance of vocabulary knowledge and working memory capacity for listening comprehension becomes particularly evident. However, in most real-life situations, learners are typically exposed to multimodal input, which can provide rich non-verbal information, such as images, gestures, and facial expressions, offering critical contextual cues for listening comprehension (Guichon & McLornan, 2008). The reliance on verbal information (i.e., vocabulary) may thus decrease to some extent, although it still plays a crucial role. Additionally, information is processed simultaneously through the listener's ears and eyes, which can help alleviate the cognitive load on a single channel (Salmani & Rahimi, 2024), and thus reduce the limitations imposed by working memory capacity on listening comprehension.

By contrast, motivational beliefs, affective responses, and self-regulation, function as intrinsic psychological traits of learners. These factors seem to have a more universal impact on listening comprehension, regardless of contexts and modality, as they can significantly influence learners' performance across different listening contexts (Boekaerts & Cascallar, 2006; Horwitz et al., 1986; In'nami, 2006; Lee et al., 2009; Oxford, 2017; Serraj, 2015) and are thus of equal if not greater relevance for research.

Secondly, it has been found that performance on working memory tasks improves with brain development from childhood to early adulthood but declines in older adults, and it is closely associated with measures of intelligence (Ma et al., 2014). This suggests that working memory capacity is more likely constrained by biology, and the effects of various forms of external interventions may thus be limited (Melby-Lervåg & Hulme, 2013; Shipstead et al., 2012). Consequently, its practical application in enhancing listening comprehension becomes

uncertain and of less interest, in this study's view.

Additionally, it has been found that vocabulary size can be improved through either direct interventions (e.g., explicit vocabulary instruction or vocabulary-focused tasks) (Subon, 2016; Yeung et al., 2020) or indirect interventions (e.g., through reading, listening, or oral interaction tasks) (Yaghoubi & Seyyedi, 2017). However, specific interventions may only have significant positive effects on vocabulary and language skills directly related to the intervention content, while their impact on overall vocabulary size and relevant language proficiency might be limited. For instance, Gellert et al. (2021) found that morphological interventions significantly improved the ability of 332 fifth-grade learners of Danish as a second language to segment and explain vocabulary and slightly enhanced their reading comprehension related to the trained content. However, the intervention had no significant impact on standardised reading comprehension tests or overall vocabulary size. It can thus be inferred that improving vocabulary size through specific vocabulary interventions to enhance overall listening comprehension might be limited in its effectiveness. The improvement of overall vocabulary size is often a gradual and long-term process, suggesting that enhancing listening comprehension through vocabulary growth is also likely to be a long-term process, and consequently, it may not be the most effective or efficient way of addressing immediate challenges in listening comprehension. Therefore, similar to the argument made in respect of working memory, vocabulary size is not a focus of interest in the present study.

By contrast, learners' personal factors, such as their motivational beliefs, affective responses, and self-regulation, appear to be more malleable. They are more likely to be effectively enhanced within a shorter period through various interventions (Fathi et al., 2020;

Graham & Macaro, 2008; Milliner & Dimoski, 2024; Xu et al., 2021), exerting direct and/or indirect effects on learners' listening comprehension (Canaran et al., 2024; Xu & Huang, 2018; Yabukoshi, 2024; Zhang & Xu, 2024). Therefore, from a practical perspective, exploring the influence of personal factors on L2 listening comprehension can provide more direct and actionable guidance for instructional interventions than investigating vocabulary knowledge and working memory, and hence arguably deserves greater attention in L2 listening research.

Taken together, while extant studies focus on the effects of vocabulary knowledge and working memory on listening comprehension, this study aims to explore the impact of personal factors on learners' L2 listening performance. Specifically, personal factors that have received extensive attention, including listening anxiety and self-efficacy, will first be discussed below, followed by a separate chapter (Chapter 3) dedicated to a detailed explanation of self-regulation.

2.2.3.1 Self-Efficacy

Self-efficacy refers to learners' beliefs in their capabilities to complete a specific task or their judgment of their capability to perform the courses of action necessary to achieve designated performances (Bandura, 1986; Graham, 2022). As defined, self-efficacy beliefs exhibit contextual specificity (Oxford, 2017), and task specificity. As an important component of the construct of motivation, self-efficacy beliefs have motivational power (Efklides, 2011). They affect "how people feel, think, motivate themselves, and behave" (Bandura, 1993, p. 118), which in turn may exert influences on learning results, in this case, listening comprehension. For instance, learners with a strong sense of self-efficacy tend to

choose tasks that are challenging and are willing to put more effort into the task, set higher goals, expect positive outcomes and are able to maintain confidence in completing the task when facing difficulties and failures and can actively adapt the use of strategies (Bandura, 1986). Learners who lack self-efficacy, on the contrary, may exhibit decisions and behaviours that are contrary to the above. Successful experiences, such as achieving good results in tests, are considered an important source of self-efficacy (Bandura, 1997). This may be because successful experiences allow learners to see that their actions can lead to positive outcomes, thereby gaining a sense of control over their learning, which may, in turn, have a positive impact on their self-efficacy (Graham, 2011).

The complexity and uncontrollability of listening input, the transience of information, and the unobservable nature of the listening comprehension process make listening appear more “uncontrollable” compared to other language skills (Graham, 2011). Therefore, perceived self-efficacy to exercise control over the listening process and potential listening difficulties can be particularly important for L2 listeners. The relationship between L2 learners’ self-efficacy and listening comprehension has been demonstrated in numerous empirical studies. For instance, the direct positive relationship between self-efficacy and L2 listening proficiency has been confirmed among university EFL learners from Turkey (Canaran et al., 2024), Thailand (Payaprom, 2023), and China (Du & Man, 2023). Rather differently, Du and Man (2022) found no predictive effect of self-efficacy on L2 listening comprehension in their study involving 367 Chinese EFL undergraduate students. They thus speculated that self-efficacy might indirectly predict listening comprehension through other factors, such as metacognition, and their speculation was confirmed in other subsequent

studies by Du and Man (2023) as well as Zhang and Xu (2024) on Chinese university EFL learners. These findings suggest that the higher L2 listeners' self-efficacy, the better their L2 listening proficiency tends to be. Conversely, when L2 listeners have low self-efficacy, that is, they do not believe they can exercise control over potential threats (e.g., listening comprehension difficulties), their anxiety, which is an emotion of fear, may be aroused (Bandura, 1988, 1990), thereby negatively impacting their listening performance.

2.2.3.2 Listening Anxiety

Anxiety can be conceptualised as “an emotion of fright indexed by physiological arousal or subjective feelings” (Bandura, 1988, p. 90). It can manifest as negative emotions such as “uneasiness, frustration, self-doubt, apprehension, and tension” (Xu, 2011, p. 1709). For language learning specifically, anxiety may arise in various specific situations, such as in the foreign language classroom and in tests, or it may be associated with particular language skills, such as listening (Horwitz et al., 1986; In’nami, 2006; Serraj, 2015). Therefore, like self-efficacy, anxiety also has a context-specific and task-specific nature. Additionally, learners may feel anxious when listening to the target language for various reasons, not only from the characteristics of the listening materials and accompanying tasks but also from concerns about tests, negative evaluations and low levels of L2 proficiency, as well as environmental factors, such as classroom atmosphere (Canaran et al., 2024).

Theoretically, anxiety is considered to have a negative impact on learning outcomes, in this case, listening comprehension. For instance, from a neurobiological perspective, anxiety “can create neural static, sabotaging the ability of the prefrontal lobe to maintain working

memory" (Goleman, 2020, p. 64). The affected working memory may, in turn, undermine critical tasks in listening comprehension, such as parsing, meaning construction, and discourse construction. Additionally, as Social Cognitive Theory (SCT) posits, anxiety arousal may shift learners' attention away from addressing learning difficulties (e.g., listening comprehension challenges) to focusing on their perceived inadequacies and worrying about undesirable outcomes, which may further reduce the effort they invest in overcoming these difficulties, ultimately affecting learning outcomes.

From an empirical perspective, the influence of anxiety on L2 listening has also been explored in extant research, but the results are mixed. For example, Golchi (2012) found a significant direct negative correlation between listening anxiety and L2 listening comprehension among Iranian EFL learners. The higher the learners' level of listening anxiety, the lower their scores on listening comprehension tests. In terms of an indirect effect, studies by Xu (2017) as well as Xu and Huang (2018) both identified an indirect negative effect of listening anxiety on L2 listening test scores through metacognitive awareness among Chinese university EFL learners. These findings together suggest that anxiety may influence L2 listening through multiple pathways and mechanisms. In contrast to studies that identified a negative relationship between anxiety and L2 listening, In'nami (2006) found no significant correlation between test anxiety and L2 listening test scores among Japanese university EFL learners. In'nami (2006) attributed this lack of correlation to several factors, such as the low-stakes nature of the listening test used in the study and learners' personal characteristics (e.g., high English proficiency, previous successful test-taking experiences, and high self-esteem). Among them, learners' strategic competence was also considered a key factor.

Specifically, learners' effective use of strategies may have enabled them to compensate for the difficulties they encountered during listening, enhanced their sense of control over the listening process, and thus helped them control their anxiety (Bandura, 1990; Graham & Macaro, 2008; In'nami, 2006).

In addition to managing anxiety, effective deployment of strategies may also contribute to the development of learners' self-efficacy (Graham, 2011; Graham & Macaro, 2008; Vandergrift, 2002) and L2 listening comprehension (Xu & Luo, 2024; Yabukoshi, 2024; Zeng & Goh, 2018). Language learning strategies (LLS) is a concept developed within the field of applied linguistics. It began to focus on second language acquisition (SLA) in the late 1970s, entered a period of rapid growth in the 1980s with the emergence of various strategy taxonomies, and reached a watershed moment in the early 21st century due to increasing criticism (Rose et al., 2017). One of the focal points of criticism from scholars was the theoretical ambiguity of LLS (Tseng et al., 2006). Scholars not only failed to reach a consensus on whether learning strategies should be regarded as observable behaviours, mental activities, or a combination of both (Rose et al., 2017), but they also disagreed on how to differentiate between learners' strategic and non-strategic behaviours (Tseng et al., 2006). For example, Weinstein et al. (2000) argued that strategic learning was characterised by being goal-directed, intentionally triggered, and effortful. However, Tseng et al. (2006) contended that these three characteristics were not unique to strategic learning and were applicable to almost all motivated or effortful learning, making them insufficient to serve as defining features of strategic learning. Additionally, Cohen (1998) emphasised choice as a key characteristic of strategic learning, arguing that the voluntary selection of strategies by the

learner is an essential feature. However, mere choice does not equate to strategic learning, as students often make various choices during their learning process, such as choosing the time, place, or content of their studies, that are not necessarily strategic (Tseng et al., 2006) and may even be random.

Riding and Rayner (1998) offered a more personalised and dynamic perspective on strategic learning. They proposed that the key to strategic learning lies in appropriateness, namely, the alignment between the strategy and the learner, rather than the universal effectiveness of the strategy. The process by which learners purposefully select and effortfully implement learning behaviours they believe to be suitable and beneficial for their learning constitutes strategic learning (Tseng et al., 2006). This suggests that a specific strategy may be highly effective for a particular learner at a certain stage or in a specific task because it aligns with the learner's needs and characteristics at that stage or for that task. However, the same strategy might be ineffective for the same learner at a different stage or task, or for other learners, as it may not match their needs and characteristics. It can thus be inferred that what makes learners strategic learners is not the specific strategies they use, but rather the fact that they decide which strategies to select and how to use them based on their own learning needs and characteristics to improve their learning effectiveness (Tseng et al., 2006). Traditional LLS research primarily focused on describing learners' specific learning behaviours. Consequently, some scholars called for a shift in research emphasis, from "focusing on the product-the actual techniques employed-to the self-regulatory process itself and the specific learner capacity underlying it". (Tseng et al., 2006, p. 81).

Self-regulation has a long-standing tradition in psychology, particularly in educational

psychology (Rose et al., 2017). Although self-regulation was not theoretically designed specifically for studying SLA, there is a close relationship between the constructs of LLS and self-regulation. Specifically, learners' strategy use is considered an integral component of the self-regulated learning framework (Gao, 2007), while self-regulation is regarded as one of the key purposes of learning strategy use (Oxford, 2017).

2.3 Self-Regulated Learning (SRL)

2.3.1 What Is Self-Regulated Learning?

Self-regulated learning (SRL) can be described as “self-generated thoughts, feelings, and actions for attaining one's learning goals” (Zimmerman & Moylan, 2009, p. 299). SRL is neither a mental ability nor an academic performance skill. Rather, it is a “self-directive process by which learners transform their mental abilities into academic skills” (Zimmerman, 2002, p.65), and a process by which “students exercise agency by consciously controlling and intervening in their learning” (Winne & Hadwin, 2012, p. 297). Self-regulated learners are proactive in their learning (Zimmerman, 2002). They are aware of their strengths and limitations, and they “set goals for their learning and then attempt to monitor, regulate, and control” their cognition, metacognition, motivational beliefs, affective responses, and behaviours until their goals are achieved (Pintrich, 2000, p. 453).

Despite the growing attention to L2 self-regulation in recent years, listening remains a language skill that has been largely overlooked in self-regulation research (Zhou et al., 2024). However, successful L2 listeners are often regarded as self-regulated listeners (Goh & Vandergrift, 2012; Pintrich, 2000), and the relationship between SRL and L2 listening has gained increasing attention over the past decade. For example, studies by Fatemi et al. (2014)

as well as Nasrollahi-Mouziraji and Birjandi (2016) demonstrated that self-regulation had a significant positive impact on L2 listening comprehension among Iranian EFL learners. SRL listening strategy instruction has been found to benefit Chinese university EFL learners' L2 listening, both within (Xu & Luo, 2024) and beyond the classroom (Zeng & Goh, 2018). Similarly, Yabukoshi (2024) explored the impact of 135 Japanese undergraduate EFL learners' perceived SRL strategy use outside the classroom on their listening proficiency, measured by the TOEIC listening test. The results indicated that L2 learners with higher levels of self-regulation, as shown in strategy use, were more likely to have better listening performance.

Like self-efficacy and anxiety, self-regulation is not a stable individual trait, instead, it is closely related to learning contexts (Boekaerts & Cascallar, 2006; Lee et al., 2009), which refer to various aspects of the task environment (Efklides, 2011; Pintrich, 2000). Broadly speaking, the context includes external variables such as "cultural or linguistic setting, type of learning environment (e.g., formal versus informal, home versus study abroad), or input and instructional conditions (e.g., focus-on-form, task-based learning)" (Ushioda, 2015, p. 64). It also encompasses internal variables of the learner, such as "beliefs about the nature of knowledge and knowing, motivational factors, and knowledge" (Winne & Hadwin, 1998, p. 301). Narrowly speaking, and as most scholars often refer to it in their research, context mainly relates to external conditions independent of the individual. Learning takes place in various contexts, and contextual factors may thus exert influences on learners' SRL (Oxford, 2017; Ushioda, 2015). For example, in a cross-cultural study, Wang et al. (2013) found differences in the use of SRL strategies for learning English between Chinese and German

college students. Specifically, German students tended to prefer strategies involving self-reflection (e.g., self-evaluation, persistence) and external interaction (e.g., seeking opportunities), while Chinese students were more inclined to use strategies focusing on internal structure and discipline (e.g., rehearsing and memorizing, organizing and transforming), as well as strategies based on their native language and self-driven consequences. Wang et al. (2013) concluded that the difference in SRL strategy use was mainly related to the different teaching methods of the two countries and attributed it to the influence of contextual factors on SRL.

The existing literature predominantly focuses on learners' self-regulation in formal learning contexts (Fatemi et al., 2014; Nasrollahi-Mouziraji & Birjandi, 2016; Xu & Luo, 2024). However, it has been argued that a traditional teacher-centred learning environment may not be the most conducive learning environment for the improvement of learners' SRL ability (Boekaerts & Minnaert, 1999; Pintrich, 2000). Learners in such contexts seem to receive much external regulation from their teachers (Boekaerts & Minnaert, 1999), which may lead to a lack of autonomy, responsibility, and freedom in their learning, ultimately hindering the development of their abilities to self-regulate their learning (Lee et al., 2009). By contrast, informal learning contexts may offer learners a high degree of autonomy and freedom to independently select learning goals and strategies. Thus, SRL in informal learning contexts deserves further investigation. However, research on L2 learners' self-regulated listening in informal learning contexts is limited and has notable limitations (Yabukoshi, 2024; Zeng & Goh, 2018). For instance, the study by Zeng and Goh (2018) investigated the impact of strategies used by Chinese university EFL learners in self-regulated extensive listening

activities on their listening performance and examined the learners' engagement in different phases of self-regulated listening. However, their research included only four participants, making the generalisability of the findings limited. Moreover, although Yabukoshi's (2024) study revealed the significant predictive role of SRL strategy use on EFL learners' listening proficiency, the questionnaire used in the study was specifically designed to measure learners' use of SRL strategies in a particular out-of-class listening task, namely TOEIC listening practice, rather than other common out-of-class English listening tasks. Therefore, the generalisability of the study's findings remains questionable. Taken together, further studies are needed to understand L2 learners' self-regulated listening in informal language learning contexts.

2.3.2 How Does SRL Work? - SRL Models

As a metacognitive activity at the highest level (Borkowski, 1996), SRL encompasses various dimensions and multiple phases. Many models have been constructed to depict the structure and mechanisms of SRL, such as Zimmerman's (2000) Cyclical Phases model, Winne and Hadwin's (1998) COPES model, Pintrich's (2000) SRL framework, Oxford's (2017) S²R model, and Efklides' (2011) MASRL model. Despite the many differences among existing SRL models (e.g., different theoretical frameworks and the specific phases included), they also exhibit notable similarities. Firstly, most SRL models encompass three essential phases: the preparatory phase, the performance phase, and the appraisal phase (Puustinen & Pulkkinen, 2001). During the preparatory phase, learners analyse the task, become aware of their motivational beliefs and affect, and set learning goals. Next, in the performance phase, learners employ strategies to accomplish the task while monitoring their progress. Finally,

during the appraisal phase, learners reflect on the outcomes of the task and, if necessary, adjust their strategies, motivations, and/or affective responses for future tasks. Secondly, most SRL models are characterised by their cyclical nature, meaning that the regulation of the current task can influence learners' future learning. Attribution plays a critical role in achieving the cyclicity of SRL, acting as a bridge between the current task and future tasks. This is because explanations for the outcomes of the current task may affect learners' motivational beliefs (e.g., self-efficacy), affective responses (e.g., anxiety), and strategy choices in similar future tasks (Panadero, 2017; Puustinen & Pulkkinen, 2001).

Some SRL models have been applied in the field of L2 research. For example, Morshedian et al. (2017) found that the self-regulation instruction designed based on the Cyclical Phases model (Zimmerman, 2000) contributed to fostering self-regulatory reading skills among Iranian EFL learners. Similarly, Öztürk and Çakiroğlu (2021) conducted research among university EFL students in Turkey to explore the relationship between SRL strategy training based on the Cyclical Phases model (Zimmerman, 1998) and EFL students' language skills in a flipped classroom context, where "students are preparing for the lesson with the materials out of class" followed by "hands-on activities in the in-class sessions" (Öztürk & Çakiroğlu, 2021, p.1). The study found that when there was no significant difference in pre-test language proficiency test scores between learners who received SRL strategy training (Experimental Group/EG) and those who did not (Control Group/CG), the EG significantly outperformed the CG in the overall average score of the post-test language achievement test, as well as in the speaking, reading, writing, and grammar sections. However, no significant between-group difference was observed in post-test listening test scores.

Learners in the EG received training in three types of SRL strategies during the study, including *organising*, *feedback*, and *time management*. The findings may thus indicate that these three SRL strategies may have limited effectiveness in enhancing learners' L2 listening proficiency, while other types of SRL strategies might be more beneficial for their L2 listening.

Additionally, the Cyclical Phases Model has also guided the development of several self-regulated language learning questionnaires, such as the *Writing Strategies for Self-Regulated Learning Questionnaire (WSSRLQ)* developed by Teng and Zhang (2016) and the *Mobile-Assisted Self-Regulated Listening Strategy Questionnaire (MSRLS-Q)* developed by Zhou and colleagues (2024).

Focusing on how task and cognitive conditions influence individual L2 learners' task definitions and how their task definitions, in turn, affect other SRL phases, Ranalli (2012) used a case study approach of involving four university ESL learners. They were drawn from a larger research project aimed at evaluating the effectiveness of a technology-based strategy instruction for L2 vocabulary learning, guided by the COPES model (Winne & Hadwin, 1998). The findings demonstrated the influence of motivational factors (particularly goal orientation, language-learning beliefs, and experiences) and domain knowledge, as cognitive conditions, on task definition, while highlighting the impact of task definition, as the initial phase of SRL, on the accuracy of other SRL phases, such as strategy selection, self-monitoring, and evaluation.

Although existing SRL models have been applied in L2 research, they may not be the optimal theoretical framework for exploring learners' self-regulated L2 listening. Firstly,

most SRL models have been constructed to illustrate its mechanisms for learning in general, rather than targeting specific learning domains such as L2 learning. While these models can provide broad SRL theoretical guidance when applied to L2 research, they may fail to capture the unique characteristics of self-regulation in the language learning domain. A few models target L2 learning, attempting to depict the mechanisms of self-regulated L2 learning, such as Oxford's (2017) S²R model. However, these models fail to capture the complexity and distinct features of self-regulated L2 listening specifically, thus offering limited theoretical guidance for L2 listening research and educational practice.

Secondly, SRL is a multi-dimensional and multi-stage framework. The multi-dimensional aspect of SRL indicates that it involves regulation across multiple domains, such as cognition, metacognition, motivation, affect, and behavior. The multi-stage feature emphasises that SRL is a dynamic process composed of multiple phases, such as the preparatory phase, the performance phase, and the appraisal phase. It can be concluded that the multi-dimensional aspect focuses on what learners should regulate, whereas the multi-stage aspect concerns how learners regulate. Together, these two features define the SRL process and are equally important, meaning that neither of them can be overlooked. If SRL models focus only on the dimensions to be regulated but overlook how these factors function before, during, and after learning, then SRL may be perceived as a static ability rather than a dynamic process skill. Furthermore, if models emphasise only the stages of regulation without analysing which specific factors are involved, they may fail to explain why regulation breaks down at certain points.

Previous models have generally placed greater emphasis on the multi-stage nature of

SRL while often neglecting its multi-dimensional aspect. For instance, extant SRL models have not sufficiently explained the processes and mechanisms of learners' motivation and affect self-regulation (e.g., Efklides' MASRL model, 2011; Oxford's S²R, 2017; Pintrich's SRL framework, 2000; Winne and Hadwin's COPES model, 1998; Zimmerman's Cyclical Phases model, 2000). The regulation of motivational and affective factors matters because learners' motivational states and affective responses are not static; instead, they may change as tasks progress, driven by ongoing metacognitive monitoring (Efklides, 2011; Winne & Hadwin, 2012). If these changes, especially negative ones, cannot be effectively controlled, they are likely to adversely impact learners' SRL. Therefore, motivational and affective factors should be one of the primary objects of self-regulation.

However, existing models either fail to address the regulation of motivation and affect (e.g., Winne and Hadwin's COPES model, 1998; Zimmerman's Cyclical Phases model, 2000), only detail the conditions that trigger motivation and affect regulation without explaining the specific processes (e.g., Efklides' MASRL model, 2011), or list strategies for regulating motivation and affect without presenting the comprehensive regulation mechanism (e.g., Oxford's S²R, 2017). Pintrich's (2000) SRL framework has considered motivation/affect as an independent aspect of self-regulation and presented the key phases but shows it as overlapping with the regulation of other aspects, namely behaviour and context. This overlap occurs because the awareness, monitoring, and control of behaviour and context cannot be fully dissociated from the regulation of motivation and affect (Efklides, 2011). The regulation of behaviour and context should be viewed more as a means of regulating motivation and affect rather than as an end in itself (Efklides, 2011). Consequently, treating behavioural and

contextual regulation as separate from the regulation of motivation and affect may increase the complexity of the model, thereby hindering its application as a theoretical guide in language learning research and educational practice. Furthermore, it is important to emphasise that motivational and affective factors within SRL should also be conceived of rather differently than how they might be viewed as separate entities; within SRL, the emphasis needs to be on how learners regulate these factors, rather than on, say, their antecedents.

Given the specific challenges and demands of L2 listening, the limitations of existing SRL models, as well as the potential benefits of understanding the mechanisms of self-regulated L2 listening for SRL theory and instructional practice, the construct of self-regulated L2 listening needs to be systematically modelled. However, the aforementioned limitations of extant SRL models have resulted in the structure and mechanisms of self-regulated L2 listening remaining unclear, which is a gap that the current study aims to fill by proposing a self-regulated L2 listening model.

2.3.3 Self-Regulated L2 Listening Model

The newly proposed self-regulated L2 listening model considered in this study comprises five phases, respectively covering *Task Representation, Goal Setting & Strategy Planning, Performance, Attribution & Adjustment, and Monitoring & Control* (Figure 2), inspired by existing SRL models (Efklides, 2011; Pintrich, 2000; Winne & Hadwin, 1998). The initial four phases progress in a logical chronological order (i.e., before, during, and after task processing), with monitoring and control processes integrated throughout all these phases, influencing the entire task execution. In practice, the outcomes of monitoring and control may interrupt the task processing at any phase. As a result, learners may cycle between these four

phases multiple times during a listening task until their learning goals are successfully met.

The new SRL model is applicable to various listening contexts. In other words, whether in formal listening contexts, such as listening classes or tests, or informal ones, such as extracurricular listening activities, the model can offer theoretical and practical guidance for L2 learners' listening development, L2 educators' listening instruction, and L2 researchers' listening studies.

During the *Task Representation* phase, L2 listeners develop an initial subjective understanding of the task, including identifying the type of comprehension required, leveraging prior knowledge to anticipate task demands, and recognising potential challenges, anxieties, or motivations related to the task. This task representation informs the subsequent *Goal Setting & Strategy Planning* phase, where learners set specific goals and plan strategies at both the task and motivational/affective levels. For instance, they may aim to grasp the gist of the material, sustain focus and effort, and manage anxiety about unfamiliar vocabulary. They consider the likely outcomes of employing particular strategies (Schunk, 2012a) such as focusing on discourse markers and self-talk, they determine which strategies can effectively help them achieve their goals at both the cognitive and motivational/affective-level, and formulate a feasible strategy plan, such as plans for aiding listening comprehension and plans for managing motivation and affect.

To achieve the cognitive and motivational/affective goals, such as grasping the main idea of the listening input, understanding details, or alleviating listening anxiety, set at the *Goal Setting & Strategy Planning* phase, learners implement the strategy plans during the *Performance* phase. Although both phases centre around "strategies" (e.g., SRL strategies and

listening strategies), the main task during the Goal Setting & Strategy Planning phase is to plan the use of strategies, whereas the primary task at the Performance phase is to execute the strategy plan, turning the blueprint into reality.

At the *Attribution & Adjustment* phase, learners reflect on their performance in task processing and motivation/affect management, as well as the knowledge or skills gained, and the effectiveness of the strategies employed during the task. Based on these reflections and any external feedback (e.g., the transcript of listening material, scores, written or verbal feedback from teachers and peers), learners attribute their success or failure to various factors. These attributions may have implications for how they listen and how they manage their motivation and affect in future similar listening tasks (Graham, 2011). Learners then make decisions to adapt, namely adjusting strategies at both the task and the motivation/affect level to enhance future learning, or to defend, namely avoiding further learning to prevent unfavourable outcomes (Zimmerman & Moylan, 2009). As a result, the proposed model exhibits a cyclical nature, similar to other SRL models mentioned above.

Self-regulated learners continuously monitor the fluency of their cognitive processing across all four phases. For instance, during the *Performance* phase, they not only track their task processing (e.g., strategy use) but also assess their listening comprehension using various information sources by checking, verifying, and confirming (Goh, 2002; Goh & Vandergrift, 2012). For instance, listeners can monitor their comprehension by checking for any words or concepts they fail to understand (difficulty-based monitoring). They can also monitor comprehension by verifying the consistency of the current listening input with their predictions (prediction-based monitoring) or with the ongoing interpretation of the context

(understanding-based monitoring). Additionally, they can confirm whether their current understanding aligns with linguistic knowledge or common sense (linguistic knowledge-based monitoring and common sense-based monitoring). When disruptions or conflicts in cognitive processing are identified, metacognitive control (e.g., retrieving relevant information from memory) is activated in response to restore fluency in task execution (Efklides, 2011).

Additionally, during the *Task Representation, Goal Setting & Strategy Planning*, and *Attribution & Adjustment* phases, listeners are not required to monitor their listening comprehension, yet they still need to continuously monitor the fluency of their cognitive processing to ensure smooth progression through each phase and to produce the intended outcomes, including accurately understanding the task requirements while recognising their motivational and affective states; setting goals and developing strategic plans at both cognitive and motivational/affective levels; and making attributions for their performance and determining whether adjustments are necessary. If listeners detect interruptions or conflicts in their cognitive processing during any of these phases, such as failing to understand task requirements clearly, lacking awareness of their listening anxiety, struggling to set clear goals and strategies, or failing to find reasonable explanations for their performance, they need to activate metacognitive control, taking immediate actions to help themselves successfully complete each phase's tasks.

Therefore, unlike previous SRL models (e.g., Zimmerman's Cyclical Phases model, 2000) which subsume metacognitive monitoring and control under the *Performance* stage, the new model posits that metacognitive monitoring and control are applicable across all major SRL phases, namely *Task Representation, Goal Setting & Strategy Planning, Performance*,

and *Attribution & Adjustment*.

The fluency of cognitive processing can impact learners' motivational and affective states. Consequently, the monitoring and regulation of motivational and affective factors begin at the start of the task and persist throughout its completion (Efklides, 2011). Control at the motivational and affective levels involves deliberate strategic adjustments by learners to manage and improve their affective and motivational states, particularly negative ones. In Figure Two, which depicts the model, dashed arrows symbolise the conditional nature of metacognitive control, which is triggered only when cognitive disruptions (e.g., listening difficulties) or negative changes in motivation/affect are detected.

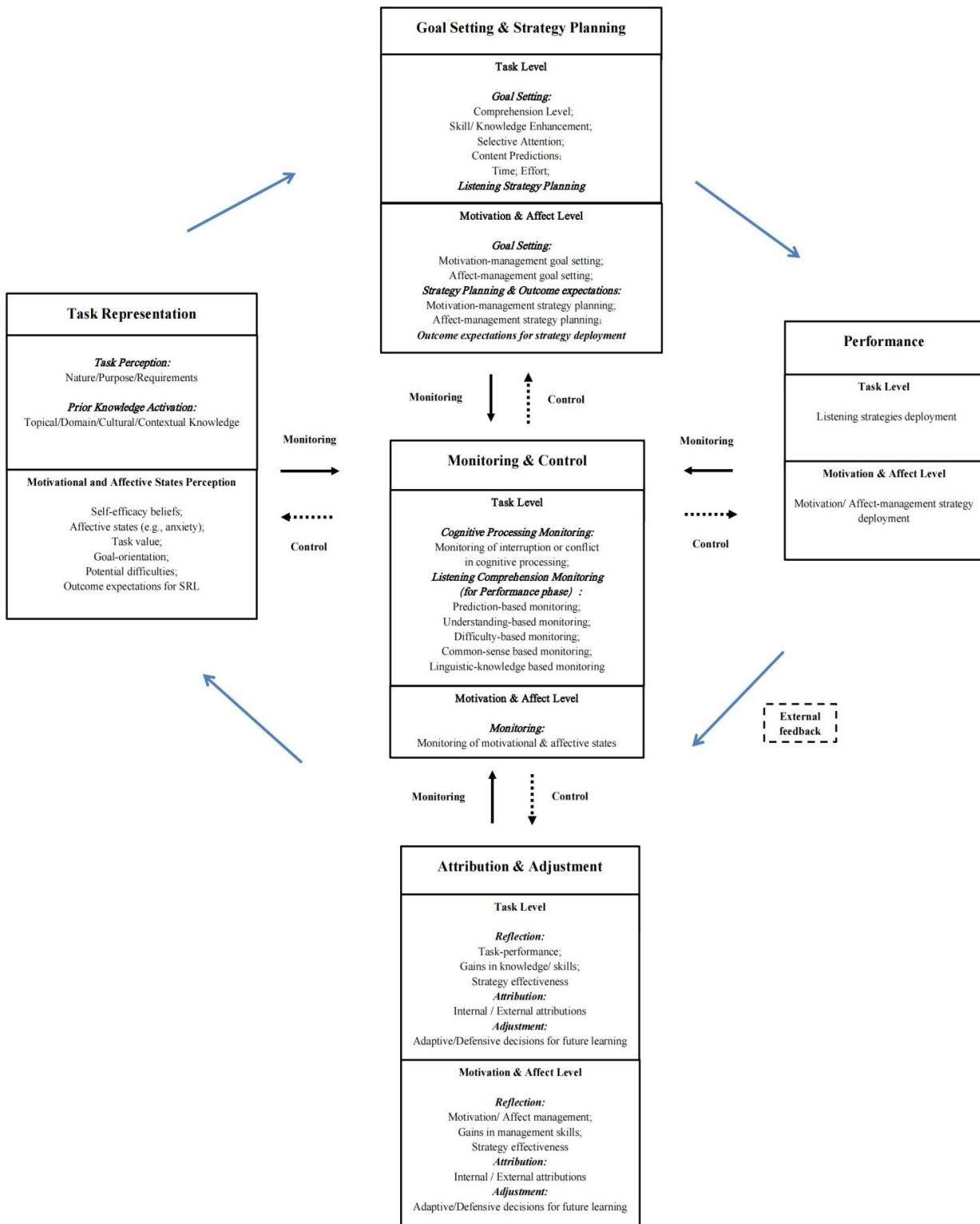
The self-regulated L2 listening model addresses some of the limitations of the extant SRL models discussed above. Firstly, rather than being a general model applicable to various learning domains, the newly proposed model is specifically targeted at the L2 listening skill. By integrating SRL theory with the processes and characteristics of L2 listening, the new model extends and develops SRL theory within the L2 listening domain, providing a more relevant theoretical framework for self-regulated L2 listening instruction, learning, and research. Secondly, the new model not only encompasses regulation at the cognitive level (i.e., task-level) but also elaborates on the mechanisms of self-regulation concerning L2 listeners' motivation and affect, highlighting the critical role motivational and affective factors play in L2 listening while reflecting an equal emphasis on both the multi-dimensional and multi-stage features of SRL.

One of the aims of this study is to validate empirically the five-phase dual-level structure of the self-regulated L2 listening model; using it alongside other important factors influencing

listening discussed above, namely self-efficacy and anxiety, to predict learners' L2 listening comprehension is another.

Figure 2

Hypothesised Structure of the Self-regulated L2 Listening Model



2.3.4 Mechanisms of Joint Prediction for Listening: A Social Cognitive Theory (SCT)

Perspective

As three key factors influencing listening, self-regulation, self-efficacy, and listening anxiety are closely and intricately connected, as depicted in Bandura's (1990,1991) Social Cognitive Theory (SCT). SCT underscores the role of self-regulation as a human capability that distinguishes them from other species and enables them to pursue a sense of agency; it also emphasises the significant role of self-efficacy within the framework of triadic reciprocal causation (i.e., personal, behavioural, and social/environmental factors) and in exercising personal agency (Schunk, 2012b). As demonstrated in *Section 2.2.3.1*, *Section 2.2.3.2* and *Section 2.3.1*, existing research has extensively examined the individual effects of these three factors on L2 listening, whereas limited studies have investigated their interplay as concurrent predictors, which constitutes another research gap this study seeks to address. As one of the few studies exploring the joint predictive effects of different variables on L2 listening, Zhang and Xu (2024) found that listening anxiety and self-efficacy in listening could not only directly predict L2 listening performance but also indirectly predict it through the mediation of metacognitive awareness. However, while metacognitive awareness is a key component of SRL (Efklides, 2011), it is not synonymous with SRL. Additionally, their study examined the predictive effects of motivational and affective factors on listening when a metacognitive factor (metacognitive awareness) served as the mediator, but it did not investigate the predictive effects of metacognitive factors on listening when motivational and affective factors acted as mediators. To gain a comprehensive understanding of the joint predictive mechanisms of self-efficacy, listening anxiety, and SRL on L2 listening, and to provide

theoretical and practical guidance for self-regulated L2 listening, it is necessary to explore these mechanisms within an empirical framework, where these factors play different roles. Therefore, the hypothesised joint predictive mechanisms of self-efficacy, listening anxiety, and SRL on L2 listening were proposed for this study under the framework of SCT.

2.3.4.1 Self-Efficacy Plays a Key Role in Anxiety Arousal

SCT posits that “perceived self-efficacy to exercise control over potentially threatening events plays a central role in anxiety arousal” (Bandura, 1990, p. 413). Threat is not a fixed attribute but rather a relational property, shaped by the alignment between an individual’s perceived ability to cope and the potentially harmful elements of their environment. Those who believe they can control potential threats do not generate apprehensive cognitions and thus exhibit little to no anxiety arousal. By contrast, those who perceive themselves as unable to manage potential threats are likely to experience heightened anxiety. They tend to dwell on their perceived inability to cope, and these unproductive thoughts may cause distress, restricting and diminishing their ability to function effectively (Bandura, 1990). Personal control can be attained through either behavioural or cognitive means (Bandura, 1988). Behavioural control involves taking actions to prevent or reduce the impact of adverse events, while cognitive control relies on the belief that one can effectively manage potential threats if they occur; and perceived self-efficacy plays a proactive role in regulating anxiety arousal in both forms of control (Bandura, 1988).

According to this perspective within SCT, when learners do not perceive themselves as capable of handling potential threats, such as comprehension difficulties, during the listening process, their listening anxiety is likely to be triggered. The aroused anxiety may cause them

to focus on their perceived inadequacies, such as low listening proficiency, rather than on the listening task at hand, thereby negatively affecting their listening performance. Thus, it can be hypothesised that:

H1: Self-efficacy can predict listening indirectly through listening anxiety.

The indirect effect of self-efficacy on linguistic performance through anxiety has been observed in the L2 field, especially in speaking (e.g., Passiatore et al., 2019), but it has not yet been explored in L2 listening.

2.3.4.2 Self-Efficacy as a Proximal Determinant of SRL

Self-efficacy plays a pivotal role in influencing SRL (Bandura, 1991). According to SCT, self-efficacy acts as a proximal determinant of SRL, shaping various sub-functions such as goal setting, attribution, and task interest (Bandura, 1991). For instance, learners with high self-efficacy are more likely to establish challenging goals, which encourage the use of effective analytical strategies, ultimately improving their performance (Bandura, 1991). These learners typically attribute failures to insufficient effort rather than a lack of ability (Bandura, 1991). Additionally, self-efficacy can predict learners' intrinsic interest, as they tend to maintain lasting engagement in activities where they perceive themselves as capable (Bandura, 1991). Hence, it can be hypothesised that:

H2: Self-efficacy can predict listening indirectly through SRL.

Existing research (e.g., Du & Man, 2023) has confirmed the mediating role of metacognitive awareness, a core component of SRL, between self-efficacy and listening. However, as mentioned earlier, metacognitive awareness is not equivalent to SRL, and the mediating role of SRL between self-efficacy and L2 listening has not yet been investigated.

2.3.4.3 SRL as a Positive Influence on Self-Efficacy

Self-efficacy serves as both a crucial contributor to and an important outcome of SRL (Pintrich, 2000; Winne & Hadwin, 1998). As SCT posits, when individuals perceive their environment as controllable, they are more likely to be motivated to fully utilise their personal efficacy to drive change, thereby increasing their chances of success (Bandura, 1991). For instance, attributions, the reasons learners give for their success or failure, are considered a vital sub-function of SRL, and are thought to influence performance attainment primarily through their impact on self-efficacy (Bandura, 1991; Graham & Macaro, 2008). Different types of attributions can lead learners to feel varying degrees of control over their learning, which can directly influence their self-efficacy (Graham & Macaro, 2008; Pintrich, 2000). For instance, internal attributions, namely “those relating to the degree of effort exerted or the strategies employed on a task” (Graham, 2011, p. 114) can make learners aware that they can influence learning outcomes by adapting the effort they put in and their use of strategies, having a sense of control over these internal factors and thus contributing to positive self-efficacy beliefs (Graham & Macaro, 2008). Effective self-regulation, whether at the cognitive level or the motivational and affective level, can enable learners to develop a sense of control over listening, thus enhancing their self-efficacy to exercise personal agency and positively impacting their listening comprehension (Graham, 2011). Consequently, it can be hypothesised that:

H3: SRL can predict listening indirectly through self-efficacy.

Similar to the two hypotheses discussed earlier, this hypothesis has not been tested yet in previous research.

2.3.4.4 Hypothesised Mechanisms of Joint Prediction for Listening

Rooted in SCT, the three aforementioned hypotheses reflect the interaction between self-efficacy and listening anxiety as joint predictors of listening (*H1*) as well as the interplay between self-efficacy and SRL as concurrent predictors of listening (*H2* & *H3*). To understand the joint predictive mechanism of these three factors on listening, two hypothesised models were further proposed.

By combining *H1* and *H2*, Model One (Figure 3) hypothesises that listening self-efficacy can not only directly predict listening comprehension but also indirectly, mediated by listening anxiety and SRL. By combining *H1* and *H3*, Model Two (Figure 4) hypothesises that SRL can not only directly predict listening comprehension but also indirectly, mediated by listening self-efficacy and listening anxiety. Meanwhile, the predictive effect of self-efficacy on listening can be both direct and indirect, mediated by listening anxiety.

Figure 3

Hypothesised Model One: The Mediating Role of Self-Regulation and Anxiety

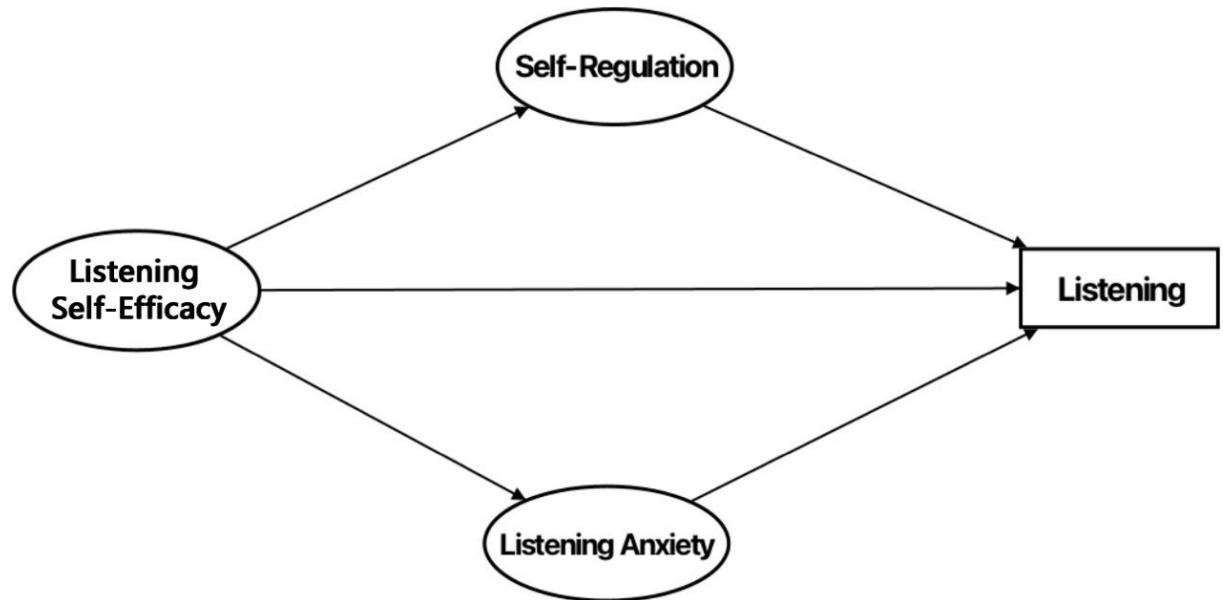
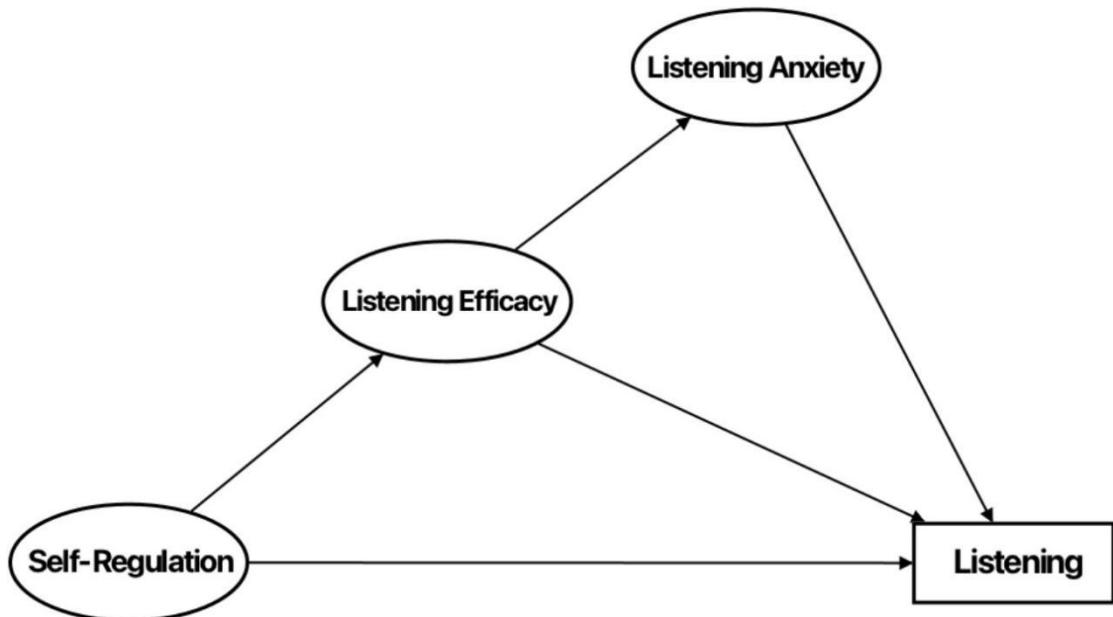


Figure 4

Hypothesised Model Two: The Mediating Role of Self-Efficacy and Anxiety



Theoretically, the two models form a positive cycle that may facilitate the development of L2 listening comprehension. On the one hand, Model Two, grounded in *H3*, suggests that effective SRL could boost learners' self-efficacy in tackling current and/or future listening tasks. On the other hand, Model One, rooted in *H2*, implies that strong self-efficacy may motivate learners to better regulate their listening processes and suppress anxiety during current and/or future tasks, ultimately enhancing their listening performance.

Additionally, these two models demonstrate the distinct roles SRL may serve within the joint prediction mechanism. Specifically, the mediating role of SRL between self-efficacy and listening comprehension depicted in Model One highlights its function as a vital tool for transforming learners' belief in their ability to complete listening tasks into their actual performance. Meanwhile, the indirect effect of SRL on listening comprehension through its influence on learners' self-efficacy, as reflected in Model Two, underscores its role as a motivation booster.

Although the primary distinction between the two models lies in the relationship between motivational and cognitive factors, namely, the direction of the arrow between them, this research argues that these two models should not be integrated into a single model for analysis. Instead, they should be treated as two independent models. Firstly, when exploring the joint predictive mechanisms of multiple factors on listening comprehension, it is necessary to consider the different roles that each factor plays within the mechanisms. To do that, an ideal approach is to take motivational, cognitive, and affective factors as independent variables separately, examining how each interacts with the other two (as mediating variables) to jointly predict learners' L2 listening. Rooted in Social Cognitive Theory, the two models

proposed in this study respectively take the motivational factor (Model One) and the cognitive factor (Model Two) as independent variables and explore their joint predictive effects on listening, along with the other two factors (i.e., affective factor, and cognitive factor for Model One or motivational factor for Model Two) as mediators. These two models, therefore, provide two key pieces of the puzzle for constructing a comprehensive joint predictive framework of listening. Integrating them into a single model would arguably weaken the central role of motivational or cognitive factors in their respective predictive mechanisms, increase the difficulty of interpreting the model, and ultimately hinder the construction of a clear and complete mechanism.

Secondly, the full predictive mechanism is still missing a third piece, namely, a model that takes the affective factor (i.e., anxiety) as the independent variable and explores its joint predictive effects on listening with motivation and cognition as mediators. This would then sit alongside the two hypothesised models proposed in this study. Up to now, however, whether within the framework of Social Cognitive Theory or other theoretical frameworks, researchers have not identified a suitable theory to model the direct and indirect predictive effects of affective factors on listening comprehension, nor the mediating roles played by motivational and cognitive factors in these indirect relationships. Therefore, this study argues that there is not strong theoretical support for this third pathway and so it was not examined in the present study, which is acknowledged as one of its limitations. However, treating Model One and Model Two as separate can allow space for a future exploration of the predictive mechanism that is centered on affective factors.

Thirdly, constructing two predictive mechanisms that are both distinct and

complementary may hold greater practical value, as researchers may in the future choose to focus on one of these mechanisms depending on their specific research goals, focal areas, or contextual needs for further validation or exploration.

Furthermore, from a theoretical perspective, this study aims to depict a comprehensive joint predictive mechanism of multiple factors on listening comprehension, rather than selecting a single mechanism from many potential mechanisms to represent the so-called optimal predictive model. Therefore, the two hypothesised models in this study are positioned as complementary rather than competing.

The two mechanisms have not been validated in previous research, either in formal listening contexts or informal listening contexts. This study will test two hypothesised joint predictive mechanisms, in order to address key issues not considered previously. First, it will seek to comprehensively illustrate the joint predictive effects of three closely interconnected factors (i.e., self-efficacy, listening anxiety, and SRL) on listening within the framework of a single empirical framework; second, it will uncover the significance of SRL in the joint predictive mechanism; and third it will provide insights for L2 listening and SRL theory and practice.

2.4 Listening in Informal Digital Learning of English (IDLE) Context

As previously stated, self-regulation, self-efficacy, and listening anxiety, as key personal factors influencing L2 listening development, are context-specific and even task-specific (Bandura, 1986; Horwitz et al., 1986; Oxford, 2017). In other words, contextual factors may have an impact on these factors. Additionally, social and listening contexts, shaped by listening environments (e.g., formal and informal listening) as well as the characteristics of

tasks, texts, and speakers, can exert direct influences on L2 listening (Goh & Vandergrift, 2012). For instance, Yu and colleagues (2021) compared the L2 listening development of 149 Chinese postgraduates studying in either China or the United Kingdom from the perspectives of linguistic knowledge and language processing. Their findings revealed that study-abroad learners made greater progress in processing speed than at-home learners. This disparity could stem from differential input exposure, where study-abroad learners benefit from extensive target-language (English) listening input, while at-home learners primarily operate in a non-target language (L1 Chinese) environment. Taken together contextual factors can be key factors influencing learners' L2 listening development.

2.4.1 An Ecological Perspective on Contextual Factors

Ecology, which explores “the interrelation between an organism and other elements within an ecosystem” (Menezes, 2011, p.60), can serve as an appropriate metaphor for understanding language acquisition. From an ecological perspective, language acquisition can be seen as a dynamic and interactive process that occurs through the continuous interplay between the learner and their learning contexts (Menezes, 2011). Specifically, learners first perceive the affordances, that is, the interaction opportunities, resources, demands, constraints, and obstacles (Shotter & Newson, 1982), in their contexts through observation and/or experience. Subsequently, learners interpret the perceived affordances by analysing and attributing meaning to them (van Lier, 2004) and then act upon the interpreted affordances by engaging in concrete learning behaviours (van Lier, 2004). The outcomes of these actions, such as feedback received during a conversation, may in turn provide new perceptions for learners (van Lier, 2004). For instance, a learner may perceive a new word during a

conversation with a native English speaker. They try to deduce the meaning and usage of the new word, then attempt to use the new word in communication or mimicking its pronunciation. They may then notice that the word has multiple variations, initiating a new cycle of interpretation and action. This “perception-interpretation-action” cycle can enable learners to continuously extract linguistic resources from their learning contexts, validate and adjust their language use, and ultimately acquire the language through dynamic interaction with contexts.

From the above process, it becomes evident that language acquisition relies not only on learners’ agency, initiative, and effort but also on contextual factors. In ecology, a niche refers to an organism’s functional position and resource requirements within its environment (Menezes, 2011). From an ecological learning perspective, learning contexts can be viewed as learners’ niches, defining how they interact with and utilise surrounding learning resources (Menezes, 2011). Different learning contexts can offer different types of affordances. For example, classrooms provide structured learning resources with clear goals and feedback mechanisms but may also offer limited opportunities for interaction and exposure to authentic language input. By contrast, informal learning contexts outside the classroom are more likely to offer diverse, authentic language resources, ample opportunities for input and output, and personalised learning experiences, although they may require higher levels of self-regulation from learners. When the affordances provided by a single context, such as classroom learning, may not be sufficient to meet learners’ language development needs (Menezes, 2011), it becomes necessary for learners to make adaptive changes, expand their niches, and perceive and utilise affordances from other contexts to facilitate language learning. This is precisely

the situation faced by Chinese university EFL learners in developing their L2 listening skills.

As detailed in the Introduction chapter (see Section 1.1.2.3), the importance of listening for learners seems to contrast sharply with the limited guidance and practice opportunities they receive within formal L2 listening contexts. Affordances beyond the classroom are crucial in the language learning process (Menezes, 2011). Therefore, engaging in extracurricular listening practice in informal contexts, where practice opportunities are abundant, materials are authentic and diverse, and learners have the freedom to select and control listening resources, appears to be a crucial way for Chinese university EFL learners to improve their L2 listening. However, limited research has shed light on the L2 listening development of Chinese university EFL learners within informal learning contexts (Zeng & Goh, 2018).

2.4.2 Defining Informal Digital Learning of English

Dressman (2020) described informal language learning (ILL) as “any activities taken consciously or unconsciously by a learner outside of formal instruction that led to an increase in the learner’s ability to communicate in a second (or other, non-native) language” (p. 4). Therefore, all activities undertaken by learners outside a formally organised program of language instruction can be regarded as a form of ILL. In the digital era, the spread of digital technology has expanded the range of possibilities for people’s informal learning activities. For example, English learners can now be exposed to and learn English outside of class, anytime and anywhere, by viewing and commenting on the videos posted by English speakers or watching English Talk shows. This fast-emerging phenomenon is known as informal digital learning of English (IDLE) (Lee, 2019a) which has been considered as a widely

studied subset of the technology-assisted language learning (Lee, 2019a, b; Lee, J. & Lee, K., 2021).

There are two main types of IDLE activities: IDLE activities in extracurricular settings and activities in extramural contexts (Lee, 2019a). The former refers to IDLE activities that are “self-directed, self-instructed” (Lee, 2019b, p. 115) but “are still linked to a formal language program” (p. 115), which is inconsistent with Dressman’s (2020) definition of informal language learning as discussed above. The second type of IDLE activities indicates “self-directed English activities in informal digital settings, motivated by personal interests and undertaken independently” (Lee, J. & Lee, K., 2021, p. 359) outside the academic requirements of the student’s teacher. Aligned with other existing research (Hamat & Hassan, 2019; Indrayani et al., 2024; Lee, 2019a, b; Lee & Dressman, 2018; Mohammed & Ali, 2021; Sylvén & Sundqvist, 2012; Yurieva et al., 2021), this study focuses on examining IDLE in the extramural context. Specifically, the current study investigates L2 learners’ listening development within IDLE contexts, and this specific learning context is thus termed as “informal digital learning of English listening (**IDLEL**)” in the present research. It refers to self-directed English listening activities conducted in unstructured, out-of-class digital environments, driven by personal interests, independent of formal language learning programs, and without being assessed by a teacher.

2.4.3 Current Research Status of IDLE

Participants’ level of engagement in IDLE activities, and their perceptions of and attitudes towards IDLE activities, have been examined both quantitatively and qualitatively (Hamat & Hassan, 2019; Mohammed & Ali, 2021; Yurieva et al., 2021).

Originating in educational psychology, engagement can be understood as “the student’s psychological investment in and effort directed toward learning, understanding, or mastering the knowledge, skills, or crafts that academic work is intended to promote” (Newmann, 1992, p. 12). While it is a psychological construct, engagement also involves action-related elements in the learning process. In other words, engagement is a multidimensional construct, which can be reflected not only cognitively, but also behaviorally, affectively and socially (Reinders & Nakamura, 2021; Zhou et al., 2021). Specifically, behavioral engagement refers to the quantity and quality of learners’ participation in learning activities; cognitive engagement involves mental processes such as the “the deliberate allocation and maintenance of attention and intellectual effort” (Zhou et al., 2021, p. 77), as well as the use of strategies and self-regulation that support these processes; affective engagement concerns learners’ affective responses and subjective experiences during learning; and social engagement focuses on the relationships established among learners through interaction and connection (Reinders & Nakamura, 2021; Zhou et al., 2021).

Engagement is seen as an essential condition for language processing and a crucial contributor to L2 development, making it fundamental to successful L2 learning (Reinders & Nakamura, 2021). Among the four dimensions mentioned above, learners’ behavioral engagement in IDLE activities has received considerable attention in existing L2 research. For example, extant studies have attempted to explore the relationship between the attributes of IDLE activity engagement, namely diversity (i.e., the variety of IDLE activity type) and frequency (i.e., the time that learners spend on IDLE activities), and different English skills, with vocabulary and speaking being the most frequently explored language skills (Lee, 2019a,

b; Lee & Dressman, 2018). By contrast, listening and reading have received limited attention (Sylvén & Sundqvist, 2012).

Regarding the relationship between IDLE engagement and L2 vocabulary and/or speaking, Lee and Dressman (2018) found that the diversity of learners' IDLE engagement had a significant positive predictive effect on their English speaking and vocabulary, using the data collected from 126 Korean university EFL learners via the Productive Vocabulary Level Test (PVLT), a speaking test, and follow-up interviews. Lee and Dressman (2018) concluded that the more diverse the IDLE activities learners engaged in, for instance, activities that balance form-focused and meaning-focused approaches, the more they could contribute to improving their L2 vocabulary and speaking proficiency. In addition to focusing on the predictive effect of IDLE engagement diversity on L2 vocabulary and/or speaking, some scholars have also examined the relationship between the time learners spend on IDLE activities and their L2 vocabulary and/or speaking. For example, Lee (2019b), after analysing data from 71 Korean university EFL learners on IDLE engagement time, English speaking test scores, and vocabulary test scores (PVLT), found that the amount of time learners spent on IDLE activities was not related to their speaking and vocabulary test scores; however, the diversity of their IDLE engagement was significantly and positively correlated with their L2 speaking and vocabulary. Similar findings can be found in Lee's (2019a) research, which investigated the relationship between IDLE engagement (duration and diversity) and L2 vocabulary among 77 Korean university EFL learners. These results suggest that the amount of time spent on IDLE activities may not be the key to ensuring L2 learners' acquisition of L2 vocabulary and speaking; rather, engagement in diverse types of IDLE activities can be

critical.

However, these findings contradict those of Sylvén and Sundqvist (2012). They conducted research among 86 Swedish fifth-grade English learners using a questionnaire, a one-week language diary, and English vocabulary tests. The questionnaire was used to gather information about learners' extracurricular English activity habits, the language diary recorded the time learners spent on seven predetermined extracurricular English activities over a week, and the vocabulary tests (measuring receptive and productive vocabulary knowledge) assessed the learners' English vocabulary proficiency. Sylvén and Sundqvist (2012) found that frequent gamers (i.e., those who played English digital games for at least five hours per week) significantly outperformed moderate gamers (i.e., those who played for approximately one and a half hours per week) in vocabulary tests, while moderate gamers outperformed non-gamers (i.e., those who did not engage in any English digital games). These results indicate a positive correlation between learners' IDLE engagement time and their L2 vocabulary proficiency. Lee (2019b) argued that the differing results regarding the correlation between IDLE engagement time and L2 proficiency across studies might be attributed to differences in learners' learning contexts (e.g., Korean EFL context versus other EFL context), highlighting this as an area for further investigation. Additionally, the researcher noted that differences in the tools used to measure language proficiency, learners' age characteristics and language aptitude (e.g., adult learners versus young learners) and learning styles (e.g., the types and ways of IDLE engagement) (Sylvén & Sundqvist, 2012) could also explain the variations in study outcomes.

In the same study discussed above, Sylvén and Sundqvist (2012) also provided empirical

evidence for the relationship between IDLE engagement time and L2 listening and reading, as they collected data on learners' English listening comprehension and reading comprehension scores from a national mandatory English test. Similar to the relationship between IDLE engagement time and L2 vocabulary, Sylvén and Sundqvist (2012) found that learners who spent more time playing English digital games achieved higher scores in the English listening and reading comprehension tests. As one of the few empirical studies focusing on the relationship between IDLE engagement and L2 listening, Sylvén and Sundqvist's (2012) study highlights the significant role of IDLE engagement at an early age in L2 listening development. However, the relationship between IDLE engagement and L2 listening among university-level EFL learners remains to be investigated, as does the relationship between **IDLEL** and L2 listening. Additionally, existing studies exploring the relationship between learners' IDLE engagement and L2 proficiency have only established a correlation between the two, without clarifying the direction of the relationship. However, correlation does not imply causation (Lee, 2019b). This means that it may not be IDLE engagement that leads to improved L2 proficiency, but rather that higher language proficiency encourages learners to engage in a greater variety of IDLE activities or spend more time on them. Therefore, the directional relationship between IDLE engagement and L2 proficiency warrants further investigation.

Apart from L2 proficiency, the influence of IDLE activities on participants' affective variables, such as foreign language anxiety, language learning enjoyment, and willingness to communicate, as well as motivational factors, such as L2 confidence, has also been investigated by many scholars (Lee, 2019b; Lee & Dressman, 2018). For instance, the time

that learners spend on IDLE activities was found to be a significant positive predictor of EFL learners' language learning confidence and enjoyment, indicating that learners who spent more time on IDLE activities tended to have higher levels of L2 confidence and enjoyment (Lee, 2019b). It is important to note that existing research on L2 self-confidence may overlap with the focus of this study on L2 self-efficacy to some extent. However, even though the two constructs share similarities, they also exhibit significant differences. L2 self-confidence can be defined as "a trait made up of cognitive (self-perceived competence) and affective (lack of anxiety) elements that contribute (with other motivational propensities) to an intention: a willingness to communicate" (Wyatt, 2022, p. 208). L2 self-confidence can be further divided into trait-like self-confidence, believed to stem from enjoyable interactive experiences related to using L2, and state self-confidence, which exists transiently in specific situations (Wyatt, 2022). Although similar to self-efficacy in its context-specific nature, state self-confidence does not exhibit task specificity. Additionally, the two constructs differ in focus. Self-efficacy influences learners' willingness to participate in specific tasks, the level of effort they invest, and the persistence they demonstrate (Bandura, 1986), while self-confidence concerns why certain L2 learners take advantage of opportunities to interact in the L2, reflecting an overall willingness to communicate but others avoid such opportunities (Wyatt, 2022). Thus, self-confidence and self-efficacy are distinct concepts, and previous research has rarely explored the impact of IDLE engagement on self-efficacy, nor the impact of **IDLE** engagement on listening self-efficacy. Moreover, learners' listening self-efficacy during IDLE engagement has been scarcely explored.

IDLE diversity, that is, the variety of IDLE activity type, has been established as a

significant positive predictor of willingness to communicate in the L2 (Lee & Dressman, 2018), indicating that learners who engage in a greater variety of IDLE activities tend to show a stronger willingness to communicate in the L2. Additionally, IDLE diversity has also been identified as a significant negative predictor of language learning anxiety (Lee, 2019b), suggesting that learners who participate in a broader range of IDLE activities are likely to experience lower levels of L2 learning anxiety. However, the relationship between **IDLEL** engagement and listening anxiety has rarely been empirically examined, nor has learners' listening anxiety in the context of IDLE engagement, suggesting that learners' affective engagement in IDLE has received little attention.

Furthermore, as mentioned earlier, studies on learners' self-regulation in informal language learning contexts are very limited, and so far, no research has been located on learners' self-regulated L2 listening within **IDLEL** contexts, indicating that learners' cognitive engagement in IDLE remains to be explored.

Taken together, the relationship between learners' IDLEL engagement and their L2 listening proficiency, as well as the relationship between IDLEL engagement and the three key factors influencing listening, namely listening self-efficacy, listening anxiety, and SRL, remain to be explored. These represent another research gaps that this study aims to address.

2.4.4 Indicators for Evaluating IDLE Engagement

Existing studies have tended to analyse learners' IDLE engagement characteristics from two aspects: quantity and quality (Lee, 2019a, b; Lee & Dressman, 2018). Specifically, the duration of learners' IDLE engagement over a period (e.g., every day, week, or month) is typically used as an indicator of the **quantity** of IDLE engagement, while the variety of

activities engaged is commonly employed to assess the **quality** of learners' IDLE engagement.

The quantity of IDLE engagement provides a straightforward reflection of learners' investment in informal language learning, indicating the depth and intensity of their engagement, while the quality of IDLE engagement reveals the breadth of learners' informal language learning experiences.

However, in addition to the two indicators discussed above, the present study proposes that other variables, such as the number of times participants engaged in IDLEL activities within a period and the strategies used during activities, can also serve as important indicators for assessing the quantity and quality of learners' IDLEL engagement, respectively. These additional indicators can further offer further insights into the depth and breadth of learners' engagement in informal language learning.

2.4.4.1 Indicators for Assessing IDLE Engagement Quantity

As mentioned above, existing studies prefer using "frequency" to represent the amount of time learners spend on IDLE activities (Lee, 2019a, b; Sylvén & Sundqvist, 2012). In the present study, however, "duration" refers to the total time learners spend on IDLEL activities, while "frequency" is defined as the number of separate occasions learners participate in IDLEL activities over a period, which has rarely been examined in previous research but deserves to be considered as one of the key indicators of the quantity of learners' IDLEL engagement. On the one hand, engagement frequency can serve as a key indicator of L2 learners' motivation and interest in L2 learning. For instance, frequent engagement in IDLEL activities may indicate a high level of interest in English listening, which could also suggest a stronger motivation for learning. On the other hand, the frequency of IDLEL engagement can

reflect learners' habits in extracurricular L2 learning. Specifically, high-frequency engagement in IDDEL activities may suggest that learners have integrated English listening practice into their daily lives, rather than relying on sporadic or irregular practice. Frequent, and hence regular exposure to authentic English listening materials outside the classroom may help learners maintain a high sensitivity to English spoken input, as well as sustain interest and motivation in language learning. Moreover, repeated exposure may help reinforce and internalise learners' acquired L2 knowledge, L2 listening strategies, and self-regulation strategies, thereby supporting the continuity and stability of their L2 acquisition. Taken together, the duration and frequency of IDDEL engagement deserve to be used as two key indicators in the present study to measure the quantity of learners' IDDEL engagement.

2.4.4.2 Indicators for Measuring IDLE Engagement Quality

From an ecological perspective, as noted earlier, different types of IDLE activities may provide learners with diverse learning resources and interaction opportunities for their extracurricular English learning, while also presenting varying constraints and challenges (Menezes, 2011; Shotter & Newson, 1982). Consequently, a single type of IDLE activity may not suffice to meet all learners' language learning needs beyond the classroom. Leveraging the resources and opportunities offered by various IDLE activities, so that they "work together in a complementary manner", with a diversity that is "crucial in sustaining the adaptability and well-being of a learning ecology" (Lai et al., 2015, p. 7). Therefore, the quality of learners' IDLE engagement can be evaluated based on the diversity of the activity types they engage in.

When discussing the diversity of informal language learning activities, existing research

(Lee, 2019a, b; Lee & Dressman, 2018) typically adopts a learner-perspective framework for categorizing informal language learning activities, such as the classification proposed by Lai et al. (2015). Based on Lai et al.'s (2015) classification, IDLEL activities that focus more on the target language system, formal elements of the language, and linguistic accuracy can be considered as form-oriented activities. In such activities, learners' attention is primarily given to learning or memorising linguistic forms and structures (Lee, 2019b). Meaning-oriented activities, by contrast, emphasise the provision of naturalistic target language exposure in a real, informal digital environment, with a primary focus on meaning and communication. It is suggested that in language learning, diversity should manifest as a balanced focus on form and meaning (Lai et al., 2015). Therefore, the diversity of IDLEL engagement should reflect a balanced engagement in both form-oriented and meaning-oriented activities.

In addition to the diversity of activity engagement, the strategies learners employ during activities should also be regarded as an important indicator of the quality of IDLEL engagement. According to Oxford (2017), the primary goal of using L2 learning strategies is to support self-regulation, and learners' SRL is achieved through the use of learning strategies. SRL emphasises learners' ability to adjust their learning behaviours across different tasks and contexts (Bandura, 1991), and the diversity of learning strategies can therefore reflect their adaptability to varying task demands and learning contexts to some extent, helping researchers evaluate learners' self-regulation in IDLEL activities. Additionally, investigating the diversity of strategy use can assist researchers in assessing whether learners have achieved a balance in their self-regulation across cognitive, motivational and affective aspects. However, to the researcher's knowledge, so far, almost no research has examined learners'

strategy-use during their IDLE engagement. In summary, the diversity of activity engagement and the diversity of strategy use deserve to be employed as two indicators assessing the quality of learners' IDLEL engagement in the current research.

2.5 Identifying Research Gaps

By reviewing the literature on listening comprehension, self-efficacy, anxiety, SRL and IDLE, several research gaps can be identified:

- 1) No previous studies have proposed and validated the self-regulated L2 listening model;
- 2) No previous studies have proposed and validated two complementary hypothesised mechanisms for the joint prediction of listening by self-efficacy, listening anxiety, and SRL. In this study, one mechanism depicts the direct prediction of listening by self-efficacy and its indirect prediction mediated by SRL and listening anxiety, while the other illustrates the direct prediction of listening by SRL and its indirect prediction mediated by self-efficacy and listening anxiety.
- 3) There is limited research examining university L2 learners' listening comprehension in the context of IDLEL;
- 4) There is limited research examining university L2 learners' SRL, listening self-efficacy, and listening anxiety in the context of IDLEL;
- 5) No previous studies have simultaneously investigated the relationship between IDLEL engagement (i.e., frequency, diversity, duration, and strategy-use) and learners' L2 listening comprehension, listening anxiety, and self-efficacy, as well as the moderating role that self-regulation plays within the relationship.

Addressing the aforementioned gaps could make significant theoretical and pedagogical

contributions. From a theoretical perspective, firstly, the proposal and validation of the self-regulated L2 listening model can expand and enrich existing SRL theories. Current SRL models have several limitations that require improvement. For instance, while these models have broad applicability, they lack specificity and provide insufficient explanations for motivational and affective self-regulation. Thus, the development and validation of a new model can not only advance existing SRL models but also extend SRL theory into the field of L2 learning. Secondly, the proposal and validation of the joint predictive mechanisms for L2 listening can offer new perspectives and empirical evidence for understanding how motivational, affective, and metacognitive factors jointly predict L2 listening. It can also contribute to understanding the importance of SRL for L2 learners' listening development and underscore the necessity of integrating SRL theory into the domain of L2 listening. Thirdly, exploring learners' L2 listening, SRL, self-efficacy, and anxiety, as well as their interactions within the context of IDLEL, can both enrich the theoretical framework of IDLEL and extend the theories of L2 listening, SRL, self-efficacy, and anxiety into the informal learning domain, thereby refining the theoretical landscape of SLA.

From a practical perspective, firstly, the development and validation of the self-regulated L2 listening model can provide theoretical foundations, guidance, and resources for language educators and learners in their self-regulated L2 listening educational practices. Secondly, the establishment and validation of the joint predictive mechanisms for listening can help language educators and learners better understand the interplay between self-efficacy, listening anxiety, and SRL in influencing L2 listening. This understanding can in turn guide the design of targeted teaching and learning strategies to effectively enhance learners'

immediate and long-term listening performance. Thirdly, exploring learners' L2 listening, SRL, self-efficacy, and anxiety in the context of IDLEL can help educators design interventions to encourage learners to use digital learning resources more effectively outside the classroom to improve their L2 listening proficiency. It can also guide learners on how to strategically engage in IDLEL activities to achieve immediate improvements in listening proficiency and sustainable long-term listening development.

In summary, the present study aims to fill the research gaps stated above and make contributions to the understanding of the role of frequency, diversity, duration, and strategy-use of IDLEL activity engagement, as well as self-regulation in the development of learners' listening comprehension, self-efficacy, and listening anxiety within the IDLEL context.

2.6 Main Research Questions

To fill the research gaps listed above, the present research will investigate five main questions:

- 1) Can the hypothesised five-phase dual-level construct of the self-regulated L2 listening model be validated?
- 2) Can the two hypothesised mechanisms through which L2 listening is jointly predicted by listening self-efficacy, listening anxiety, and self-regulation be validated?
- 3) What are the characteristics of Chinese undergraduates' IDLEL engagement, in terms of its quantity (i.e., frequency and duration) and quality (i.e., diversity and strategy-use)?
- 4) How do learners' IDLEL engagement (i.e., quantity and quality) and SRL predict their English listening comprehension, listening anxiety and listening self-efficacy respectively?

5) To what extent does learners' level of SRL moderate the relationship between the IDLEL engagement (i.e., diversity, frequency, duration, and strategy-use) and their L2 listening comprehension, listening anxiety, and listening self-efficacy?

Chapter 3. METHODOLOGY

This chapter elaborates on the methodology of this study in terms of paradigm rationale, research design, participants, data collection, data analysis, considerations of reliability and validity, ethical issues, lessons from pilot study as well as implications for main study.

3.1 Research Design

3.1.1 Ontology, Epistemology, and Methodology

Paradigms refer to collections of fundamental beliefs that researchers use as frameworks to guide their research in practice (Bhattacherjee, 2012; Guba & Lincoln, 1994; Rahi, 2017; Schnelker, 2006). Different paradigms are distinguished by different philosophical assumptions, including 1) ontology, 2) epistemology, and 3) methodology (Guba & Lincoln, 1994; Kivunja & Kuyini, 2017). Ontology deals with “the form and nature of reality” and “what is there that can be known about it” (Guba & Lincoln, 1994, p. 108). Simply speaking, it concerns “how we see the world” (Bhattacherjee, 2012, p. 18). Epistemological assumptions deal with “how we study the world” (Bhattacherjee, 2012, p. 18) or “how we come to know something” (Kivunja & Kuyini, 2017, p. 27). Ontology and epistemology are fundamental and central to a paradigm and furthermore influence the choice of methodology, which refers to well-planned data collection and analysis methods and procedures that can be employed to understand something (Kivunja & Kuyini, 2017).

The interweaving of the basic assumptions above has generated a variety of research paradigms, of which three are widely used, including the positivist paradigm, the interpretive paradigm, and the pragmatism paradigm (Rahi, 2017). Positivists believe that apprehendable realities that are “driven by immutable natural laws and mechanisms” (Guba & Lincoln, 1994,

p. 109) exist independently of researchers, and can be investigated without being influenced by researchers, values, and bias (Dammak, 2015; Kivunja & Kuyini, 2017). The positivist paradigm is grounded in and leads to a scientific method of research, emphasising the essential role of observation and experiment in knowledge attainment and the establishment of cause-effect relationships (Kivunja & Kuyini, 2017). Therefore, the value of objectivity is a key characteristic of the positivist paradigm. Studies influenced by the positivist paradigm tend to apply deductive logic to research by proposing hypotheses and then using experimental data, especially quantitative data to test these hypotheses. Quantitative research methods, such as laboratory experiments, tests, and survey research, thus, are considered to be closely related to the positivist paradigm (Bhattacherjee, 2012; Rahi, 2017). In contrast to the positivist paradigm which values objectivity, the interpretive paradigm holds the belief that socially and experientially based apprehendable realities exist “in the form of multiple, intangible mental constructions” (Guba & Lincoln, 1994, p. 110). Their forms and contents are closely associated with the knowers who hold these constructions. Hence realities may vary for different individuals or groups. Knowledge can be obtained through the interaction between the researcher and the research subject for the purpose of investigating their subjective perceptions and interpretations of the world around them (Dammak, 2015; Guba & Lincoln, 1994; Kivunja & Kuyini, 2017). Studies in the interpretive paradigm usually employ an inductive approach, aiming to “derive a theory about the phenomenon of interest from the observed data” (Bhattacherjee, 2012, p. 35). Since the main task of the interpretive paradigm involves a deeper understanding of the subjective world of the participants being observed,

humanistic qualitative research falls primarily under this paradigm, which is different from the positivist paradigm that prefers scientific quantitative research methods.

The emergence of the pragmatism paradigm was an attempt to remedy the weakness of a single paradigm that might not allow for a thorough and comprehensive study of the phenomenon (Kivunja & Kuyini, 2017; Rahi, 2017). As Kivunja and Kuyini (2017) conclude, the pragmatism paradigm assumes a:

relational epistemology (i.e., relationships in research are best determined by what the researcher deems appropriate to that particular study), a non-singular reality ontology (that there is no single reality and all individuals have their own and unique interpretations of reality), a mixed methods methodology (a combination of quantitative and qualitative research methods) (p. 35).

Pragmatism advocates that research questions are at the core of the research, the choice and application of philosophical and/or methodological approaches therefore should serve the investigation of research questions (Kaushik & Walsh, 2019). Researchers should seek to use both qualitative and quantitative research methods to answer the research questions (Kivunja & Kuyini, 2017; Rahi, 2017). Research that uses both qualitative and quantitative methods in a single study to collect and analyse research data is known as mixed-method research, whereby quantitative research methods are used to collect numerical information, while qualitative research methods are responsible for the collection of textual and visual information (Creswell, 1999). It is believed that the joint use of two methods can help researchers to gain a more comprehensive understanding of a complex social phenomenon

and obtain results that cannot be derived from one research method alone (Bhattacherjee, 2012). For instance, a mixed-method study can allow researchers to not only observe the behaviour of participants, but also to explore the reasons and logic behind the behaviour, as well as the consequences of the behaviour (Kivunja & Kuyini, 2017).

3.1.2 Rationale for Adopting Mixed Method and Correlational Research Design

Although ontological and epistemological assumptions are essential to paradigms and they are the basis of the methodology, methodology seems to be more practically meaningful to the researcher when it comes to choosing a suitable paradigm for the study since it is directly related to how the research questions can be answered. As can be seen from the previous discussion, positivism prefers quantitative research methods, interpretivism favours qualitative methods, whereas pragmatism incorporates both qualitative and quantitative methods. To answer the five research questions listed in the previous chapter, both quantitative data (e.g., participants' listening proficiency, level of self-efficacy, listening anxiety and self-regulation, the duration, frequency, and diversity of IDLEL activity engagement) and qualitative data (e.g., participants' SRL strategy-use in IDLEL activities) were needed. Thus, this study employed a mixed-method approach, drawing on both quantitative and qualitative research methods. This also means that the pragmatism paradigm was chosen for the present research.

Research design is a strategy or blueprint for answering research questions in the empirical research (Bhattacherjee, 2012; Rahi, 2017). There are various widely used research designs, such as experimental design, case study, causal-comparative design, correlational design, phenomenological design, and ethnographic design. To address the research questions

for the current study, a correlational research design was employed. Correlational studies are those in which the researcher collects and analyses numerical quantitative data to establish statistical relationships between variables without imposing manipulation or control (Asamoah, 2014). There are two distinctive features of correlational studies. First, they can only be conducted to explore the strength and direction (positive, negative or zero correlation) of the relationship between multiple variables, but they cannot be used to demonstrate the cause-and-effect relationships between variables, unlike experimental studies. Second, in correlational research, the researcher will not influence, manipulate, or control variables, which is also very different from experimental research that requires manipulation and influence on variables.

Regarding the present study, in addition to validating the hypothesised self-regulated L2 listening model and the joint predictive mechanisms of listening comprehension, the main objectives was to explore the relationships between several variables, including L2 listening, listening self-efficacy, listening anxiety, self-regulation, and IDLEL engagement among Chinese university EFL learners. Therefore, a correlational study design was chosen for the present research.

The research questions proposed in the current study mainly revolve around two themes. The first theme concerns the validation of the hypothesised structure of the self-regulated L2 listening model. The second theme explores the relationships among variables, including listening self-regulation, listening self-efficacy, listening anxiety, IDLEL engagement, and L2 listening. Specifically, it relates to the confirmation of the joint predictive mechanisms of self-regulation, self-efficacy, and listening anxiety on L2 listening, as well as the examination

of the predictive roles of self-regulation and IDLEL engagement in listening, self-efficacy, and listening anxiety.

Given that the measurement of learners' L2 listening self-regulation relied on the validated structure of the self-regulated L2 listening model when exploring relationships between variables, the study was conducted in two consecutive sections. The first section focused on validating the hypothesised self-regulated L2 listening model structure, while the second investigated the relationships among the variables.

Table 1
Research Design

Part 1: Questionnaire Study						
Instruments	English listening questionnaire (self-regulation, listening anxiety, listening self-efficacy)					
	two-week gap					
Part 2: IDLEL Study						
	Phase 1		Phase 2		Phase 3	Phase 4
Week	1	2	3	4	5	6
	19					
Instruments	First English listening comprehension test	E-log	E-log	E-log	Second English listening comprehension test	Third English listening comprehension test
	First English listening questionnaire (self-regulation, anxiety, self-efficacy)				Second English listening questionnaire (self-regulation, anxiety, self-efficacy)	

Table 1 shows all aspects of the research design. According to Table 1, in the first part of the study, participants were required to complete electronic questionnaires investigating their self-regulation, self-efficacy, and anxiety in listening. Two weeks later, the second part of the study was conducted in four phases. In the first phase (week 1), a different cohort of participants from those in the first part of the study were first asked to complete a listening test in their English listening classes. Immediately after the listening test, participants were required to complete questionnaires capturing their immediate responses to factors (i.e., SRL strategies, self-efficacy, and anxiety) related to English listening. The questionnaires used in the second part of the study were the updated version based on the confirmatory factor analysis (CFA) results of the questionnaire data collected in the first part of the study. This is because during CFA, items that failed to effectively measure the target constructs in the initial version were removed, resulting in the revised questionnaires (see *Sections 4.1* for details). In the second phase (week 2 to week 5), participants were asked to complete as many IDLEL activities as they wished in addition to the work set by their teachers and keep one English E-log a week in which they briefly described the way they engaged in IDLEL activities. In the third phase (Week 6), participants completed a second English listening test as the post-test and filled in the same questionnaire they had completed in Week One. In the fourth phase (Week 19), participants took a third English listening test as the delayed post-test. All the above instruments will be discussed in detail in *Section 3.3*.

3.2 Participants

3.2.1 Selection of Participants

The participants of this study comprised two samples of Chinese university EFL learners: 1) For Part 1, 582 questionnaire respondents, whose questionnaire data were used to validate the structure of the hypothesised self-regulated L2 listening model; 2) For Part 2, 130 IDLEL study participants, whose listening test scores, questionnaire responses, and E-log data were used to explore the relationships among self-regulation, self-efficacy, listening anxiety, L2 listening proficiency, and IDLEL engagement.

Table 2 displays the key demographic information of the participants. The Part 1 582 questionnaire-only participants were drawn from five Chinese universities, with four located in a middle-China province and one in a southeastern coastal province. Research has shown that English education resources are unevenly distributed across different regions in China, particularly between coastal and inland areas (Hu, 2003). Additionally, students from coastal regions are also more likely to have higher confidence in their English listening than those from inland areas (Xu & Qiu, 2022). Therefore, recruiting participants from different regions was undertaken in order to enhance the representativeness of the sample, thereby improving the robustness of the study findings.

Additionally, the unequal distribution of English education resources in China is also evident between key and non-key universities. Key universities refer to those supported by initiatives such as *985 Project*, *211 Project*, or the “*Double First-Class*” initiative, which aim to develop domestic “world-class” universities. Compared to non-key universities, these key universities are more likely to receive greater financial support, have more specialised faculty, and offer higher-quality teaching resources (Han et al., 2023; Li & Zhou, 2017). These differences may lead to variations in students’ language proficiency, self-regulation skills,

language learning motivation, and affective factors between key and non-key university students. To further enhance sample representativeness of the present study, two key universities and three non-key universities were selected.

Moreover, considering that different academic disciplines typically have their distinctive learning practices, disciplinary differences may influence learners' L2 learning styles and preferences (Lau & Gardner, 2019), autonomy and sense of language competence (Xu & Qiu, 2022), strategy use (Rao, 2005), language development levels (e.g., vocabulary acquisition) (Gu, 2002), as well as motivation and affective factors. Therefore, learners from various academic disciplines were invited to participate in the questionnaire survey. Furthermore, academic year has been found to significantly influence students' confidence in their listening proficiency (Xu & Qiu, 2022) and listening anxiety (Wang, 2023). Hence, undergraduate students from different academic years were invited to participate in the study.

To recruit this participant sample, the study adopted a snowball sampling strategy, whereby "existing study subjects recruit future subjects among their acquaintances" and "sampling continues until data saturation" (Naderifar et al., 2017, p.2). This method can be particularly suitable for situations where direct access to the target population is challenging, which was the case for the Part 1 questionnaire-survey in the current study. Specifically, the researcher firstly asked university English teachers with whom they were familiar to distribute an electronic questionnaire to students who had attended their English courses during the second semester of 2022. These teachers thus served as initial contacts for the participants. Subsequently, these teachers forwarded the questionnaire to other English teachers, who then distributed it to their own students. This process continued until the dataset

met the predetermined sample size criteria. The number of participants required for Confirmatory Factor Analysis (CFA), which was used to validate the structure of the self-regulated L2 listening model in the present study, should be no less than five times the number of questionnaire items (Sun et al., 2021). Given that the questionnaire developed based on the self-regulated L2 listening model consisted of 46 items (*Section 3.3.1.1*), the student sample size for CFA was set at no less than 230. Ultimately, 582 Chinese university EFL learners aged 18-21 ($M = 20.01$, $SD = 1.13$) from 20 classes across five universities were recruited as the Part 1 student sample for this study.

The 130 participants in the Part 2 IDLEL study were from two universities located in middle-China, selected from the five universities mentioned earlier: one key university (University A, $N = 89$) and one non-key university (University B, $N = 41$). All 130 participants were English major undergraduates from the Faculty of Foreign Languages at their respective universities. These 130 students had not participated in the Part 1, first wave of the questionnaire survey. Since the English listening test and questionnaire in the IDLEL study needed to be conducted during participants' regular English classes, students were selected directly from intact classes as research participants, to avoid disrupting the original teaching schedule and enhance data collection efficiency. As this study was conducted during the COVID-19 pandemic, many Chinese universities had not yet resumed face-to-face classes. Therefore, only two of the five universities that had resumed face-to-face teaching were eligible for the study. The researcher contacted English listening course instructors with whom they were familiar at these two universities. The instructors identified students enrolled in their English listening classes during the second semester of 2022 as potential participants.

There are several advantages to selecting the Part 2 IDEL study in the way outlined. Firstly, the sample was drawn from different types of universities (key and non-key) and different academic years, as discussed earlier, ensuring sample representativeness and enhancing the robustness of the study findings. Secondly, these students had received English education throughout primary, secondary, and tertiary education and had at least eight years of English learning experience, equipping them with sufficient language proficiency to participate in the study. Thirdly, as young adults who had recently reached adulthood, they had developed independent thinking skills and mature cognitive abilities, enabling them to make autonomous decisions regarding their engagement and to fully comprehend the research content and requirements.

The questionnaires used in each part of the study included a language background section. This indicated that all 712 participants had Chinese as their first language and English as their second language. Additionally, participants were asked to self-assess their English listening proficiency on a 0-10 scale, where 0 represents no proficiency and 10 indicates high proficiency. The results showed that the self-assessed listening proficiency of 582 participants ranged from 1 to 10 ($M = 6.40$, $SD = 1.83$), while that of the 130 English majors ranged from 2 to 10 ($M = 6.50$, $SD = 1.71$).

All participant recruitment activities were approved in advance by the relevant English teachers at the target universities. Two weeks before the Part 1 questionnaire was administered, the researcher sent a *PowerPoint* file (see Appendix 1) via the instant messaging software *WeChat* to the English teachers at each university. This *PowerPoint* contained essential information about the study, including its purpose, duration, procedures,

privacy protection principles, potential risks, and benefits. The teachers then introduced the study to their students in class and answered their questions regarding the research. All students were given one week to decide whether to participate in the study. Additionally, two weeks before the Part 2 IDLEL study began, the researcher personally introduced the study and answered questions in the first ten minutes of each class's listening course. Similarly, students were given ample time (i.e., one week) to decide whether to participate in the study.

Table 2

Participant Demographic Information

	Part 1: Questionnaire-Study (N = 582)	Part 2: IDLEL Study (N = 130)
Gender		
Female	487	112
Male	95	18
Academic Years		
Year 1	176	89
Year 2	142	41
Year 3	264	
Academic Disciplines		
English	427	130
Management	50	
Economics	34	
Information Technology	31	

Law	40	
University Type		
Key university	293	89
Non-Key university	289	41
University Geographical Context		
Middle China	521	130
Southeastern coastal region of China	61	

It is worth noting that, the gender distribution in the participant sample of this study was uneven, as shown in Table 2. According to existing research (Yu, 2015) and the researcher's personal experience, the number of male students majoring in English at Chinese universities can be significantly lower than that of female students. In the present study, English major undergraduates constituted the majority of participants in both the questionnaire study and the IDLEL study. As a result, a gender imbalance emerged in the participant sample. This issue will be acknowledged as a limitation of the study and will be discussed in detail in *Section 6.5*.

3.3 Data Collection: Research Instruments and Procedures

As stated above, this research employed a mixed-method research design, meaning that both qualitative and quantitative data needed to be collected. Therefore, this section will discuss the collection of quantitative and qualitative data in detail. English listening

questionnaires, E-logs, and listening comprehension tests were used as research instruments to collect data for the present research, and each instrument is outlined below.

3.3.1 English Listening Questionnaire

The English listening questionnaire developed and used in the current research consisted of three sub-sections (see Appendix 2 for the complete questionnaire). The first sub-section was the *Self-Regulated L2 Listening Questionnaire (SRLLQ)*, which firstly aimed to validate the self-regulated L2 listening model and secondly to investigate participants' level of self-regulation in English listening. The second sub-section was the *L2 Listening Anxiety Scale (LLAS)*, designed to explore learners' English listening anxiety, and the third sub-section was the *Listening Self-efficacy Questionnaire (LSEQ)*, which aimed to examine participants' English listening self-efficacy.

3.3.1.1 Self-Regulated L2 Listening Questionnaire (SRLLQ)

The *Self-Regulated L2 Listening Questionnaire (SRLLQ)* was specifically designed for the present study to validate the hypothesised structure of the newly proposed SRL model while measuring participants' L2 listening self-regulation. Following a similar process to the development procedures of other questionnaires (Teng & Zhang, 2016; Zhou et al., 2024), the development of the SRLLQ began with item generation, followed by expert review and piloting. The item generation process typically starts with a review of relevant literature, which serves two purposes: 1) providing a theoretical basis for the design of specific items and the overall structure of the questionnaire, and 2) drawing insights from similar questionnaires within the same research domain. Since one of the core objectives of the development of SRLLQ was to validate the structure of the self-regulated L2 listening model,

this model thus served as the theoretical framework for the questionnaire design. Specifically, the five phases of the proposed self-regulated L2 listening model were used as five corresponding constructs in SRLLQ, with SRL listening strategies relevant to each phase forming the items within each construct. Specifically, seven items (Item 1-7) were generated for *Task Representation*, six (Item 8-13) for *Goal Setting & Strategy Planning*, 14 (Item 14-15; Item 17-28) for *Performance*, eight (Item 39-46) for *Attribution & Adjustment*, and 11 (Item 16; Item 29-38) for *Monitoring & Control*. Additionally, to validate the dual-level structure of the model, which integrates cognitive and motivational/affective self-regulation, each construct included items related to both listening self-regulation and motivational/affective self-regulation.

The questionnaire items were primarily created by the researcher based on SRL theories (Efklides, 2011; Oxford, 2017; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000), as well as on theoretical and empirical literature on listening strategies (e.g., Goh, 2002). Additionally, inspiration from existing representative measurement instruments in the listening strategy research field, such as Vandergrift et al. (2006)'s *Metacognitive Awareness Listening Questionnaire (MALQ)*, also contributed to the development of SRLLQ items. As the first developed metacognitive awareness questionnaire for listening, the MALQ integrates core concepts of SRL and metacognition, assessing students' metacognitive awareness and self-regulated listening strategy use from five dimensions: *Directed Attention, Problem Solving, Planning-Evaluation, Mental Translation, and Person Knowledge*. In the current research, some items from the MALQ, such as "I translate in my head as I listen", were adapted into the SRLLQ as "During listening, I translate words, phrases, or sentences into

Chinese in my mind to aid comprehension” to better suit the research context and needs of this study.

In terms of its format, the SRLLQ was designed with reference to existing research instruments. Both widely recognised and extensively used classic tools, such as the MALQ, and emerging instruments reflecting contemporary developments, such as Zhou et al. (2024)’s *Mobile-Assisted Self-Regulated Listening Strategy Questionnaire (MSRLS-Q)*, which was designed for L2 listening research in mobile-assisted language learning (MALL) contexts, tend to adopt a six-point Likert scale, allowing participants to rate each statement based on their level of agreement (1 = strongly agree, 6 = strongly disagree). This design can be considered beneficial in preventing students from selecting the neutral option as a strategy to avoid giving a clear response, thereby reducing ambiguity in their answers (Rahimi & Abedi, 2014; Zhou et al., 2024). Consequently, the SRLLQ also employed a six-point Likert scale.

Subsequently, the SRLLQ, consisting of 46 items, was assessed by two experts in the field of L2 acquisition to ensure its “redundancy, content validity, clarity, and readability” (Vandergrift et al., 2006, p. 440). Since the self-regulated L2 listening model applies to various learning contexts, the SRLLQ, which was developed based on this model, can also be suitable for diverse listening contexts.

3.3.1.2 L2 Listening Anxiety Scale (LLAS)

The *L2 Listening Anxiety Scale (LLAS)* was developed and employed to measure participants’ L2 listening anxiety in the current study. This scale was adapted from Elkhafaifi’s (2005) Foreign Language Listening Anxiety Scale (FLLAS). While the FLLAS was originally designed to measure foreign language listening anxiety among university-level

Arabic learners, it has been widely adapted in various studies as a tool for measuring foreign language listening anxiety among university EFL learners in countries such as China (Liu & Xu, 2021), Saudi Arabia (Alshahrani & Almanea, 2023), Iran (Valizadeh & Alavinia, 2013). Not only did Elkhafaifi (2005)'s original FLLAS demonstrate high internal consistency (Cronbach's $\alpha = .96$), but its adapted versions have also shown good internal consistency, with Cronbach's α ranging from .74 to .89. Therefore, it was selected for use in this study in an adapted form.

Unlike other studies that adapted the FLLAS simply by replacing the target language "Arabic" with "English" (Alshahrani & Almanea, 2023; Liu & Xu, 2021; Valizadeh & Alavinia, 2013), the adaptations in the present study included other three aspects. Firstly, the original FLLAS uses a five-point Likert scale. As previously mentioned, a six-point Likert scale appears to be more effective in preventing learners from giving ambiguous responses to item statements (Rahimi & Abedi, 2014). Consequently, the adapted questionnaire employed a six-point Likert scale, with higher scores indicating higher levels of anxiety.

Secondly, the FLLAS was adapted from the *Foreign Language Reading Anxiety Scale (FLRAS)* by Saito et al. (1999). The adaptation was limited to replacing the word "reading" in the FLRAS with "listening" in the FLLAS and substituting the target languages "French, Russian, Japanese" with "Arabic", while the number and content of the items were barely modified (Elkhafaifi, 2005). Although both listening and reading are receptive language skills, there are still significant differences between the two. Therefore, items suitable for measuring reading anxiety may not be entirely applicable to listening anxiety. Upon evaluation by the researcher and two L2 experts, it was determined that 10 (i.e., Item 9 and Items 12 to 20) out

of the 20 items in the FLLAS were inappropriate for understanding learners' listening anxiety and thus were removed. Moreover, the FLRAS assesses learners' reading anxiety from three aspects: learners' self-reports of anxiety over various aspects of reading (i.e., reading processing difficulties or general reading difficulties), their perceptions of reading difficulty (i.e., specific reading difficulties), and their perceptions of the relative difficulty of reading compared to other language skills (Saito et al., 1999). Accordingly, the FLLAS evaluates learners' listening anxiety from the three aspects above. In the FLLAS, items 8 and 10, which were used to understand learners' perceptions of specific listening difficulty, and item 4, which was used to understand learners' difficulties in listening processing, were considered insufficient by the researcher and L2 experts to serve their intended purposes. Therefore, these three items were modified for greater clarity and precision, after reaching a consensus.

Thirdly, as discussed in *Section 2.2.3.2*, anxiety can be situation-specific (Horwitz et al., 1986). Therefore, when assessing learners' listening anxiety, situational factors should be considered. After discussions between the researcher and the two experts, items involving seven specific English listening situations were added to the questionnaire, including two in-class situations (i.e., English courses and tests) and five out-of-class situations (i.e., English radio, movies and TV, songs, social media platforms, and presentations). Additionally, the questionnaire included negatively worded items (e.g., "I do not feel nervous when listening to passages from my English textbook") to prevent respondents from consistently scoring on one side of the scale (Vandergrift et al., 2006).

Finally, an LLAS with 17 items using a six-point Likert scale was developed. The questionnaire aimed to measure learners' listening anxiety from three aspects: anxiety over

listening processing difficulties (Item 1-4), anxiety over specific listening difficulties (Item5-10), and contextual listening anxiety (Item 11-17). These aspects constituted the three constructs of the questionnaire (i.e., listening processing anxiety, listening difficulty anxiety, and contextual listening anxiety), with each construct comprising four, six, and seven items respectively.

3.3.1.3 Listening Self-Efficacy Questionnaire (LSEQ)

Self-efficacy in practice is typically measured by learners' confidence in achieving specific outcomes (Efklides, 2011). Self-efficacy questionnaires, therefore, often use a 0 to 100 scale, following Bandura's (2006) lead. For instance, Graham and Macaro's (2008) questionnaire, on which the *Listening Self-Efficacy Questionnaire (LSEQ)* used in this study was based, used a 0-100 scale to assess learners' confidence in different aspects of listening, including understanding the main idea, understanding details, inferring the meaning of unknown or difficult words, and identifying opinions expressed in the text. Considering the task-specific nature of self-efficacy (Pintrich, 2000), the LSEQ retained the same items and scale format as Graham and Macaro's (2008) questionnaire but was designed around eight specific types of English listening tasks relevant to both classroom (i.e., listening to textbooks and listening tests) and out-of-class contexts (i.e., English video, TV/movies, videos on social media platforms, presentation, audiobook, songs). Participants were asked to rate their confidence in relation to these eight types of listening tasks on a scale from 0 to 100 from two perspectives: understanding the gist of the listening input (i.e., comprehending the main idea and identifying the viewpoints expressed in the passage) and understanding detailed information (i.e., grasping details and inferring the meanings of unknown or difficult words). Hence the questionnaire assessed two main constructs: listening self-efficacy for gist

comprehension and listening self-efficacy for comprehension of detailed information, each comprising 16 items.

To prevent participants' L2 proficiency from hindering their understanding of the questionnaire items, the three questionnaires were first translated into Chinese by the first researcher who is a native Chinese speaker. Then, another expert, also a native Chinese speaker, translated the Chinese version back into English. The back translation method can help ensure the consistency of the bilingual versions of the questionnaires (Teng & Zhang, 2016).

The password-protected electronic *English Listening Questionnaire*, which consisted of three sub-questionnaires (i.e., SRLLQ, LLAS, and LSEQ), was created by the researcher on the professional online questionnaire platform *WenJuanXing* (问卷星) and distributed via a QR code generated by the platform to the English listening teachers of the participants' classes. Immediately after both the English listening pretest and post-test were completed, the teachers shared the QR code and questionnaire access password in the class WeChat groups. Participants then accessed the questionnaire to complete it by scanning the QR code with their mobile devices and entering the access password. According to *WenJuanXing* platform, the questionnaire took approximately 15 minutes to complete.

3.3.2 IDLEL E-Logs

E-logs (electronic logs) were employed to collect both quantitative data and qualitative data regarding participants' IDLEL engagement (Appendix 3). A log is a tool for language learners to record, review and reflect on their language learning. By keeping logs, a language learner can "report on affective factors, language learning strategies, and his own

perceptions--facets of the language learning experience which are normally hidden or largely inaccessible to an external observer" (Bailey & Ochsner, 1983, p. 189). It has been argued that, for language learners, keeping learning logs can facilitate their awareness of and reflection on the process and behaviour of language learning, thus helping them recognise potential deficiencies and make changes (Bailey & Ochsner, 1983; Engin, 2011; Fry, 1988). For researchers and teachers, logs can help them gain comprehensive and detailed insights into learners' behaviours, strategies, perceptions, and other learner variables such as feelings, thus providing rich evidence for pedagogical improvement and research findings (Bailey & Ochsner, 1983; Fry, 1988). In addition, logs can serve as an essential source of data triangulation, used in conjunction with other data sources to answer research questions comprehensively and increase the quality of the study (Bailey & Ochsner, 1983).

When it comes to potential drawbacks of logs, their time-consuming nature, how far they raise learners' awareness of what they are doing and thus influence their behaviour, the nature of the retrospective data, and the nature of the self-report data are noteworthy (Bailey & Ochsner, 1983; Fredricks & McColskey, 2012; Fry, 1988; Garcia & Pintrich, 1996). Firstly, regarding the time issue, Fry (1988) believed that keeping logs could take a considerable amount of time and that learners would easily become fatigued and give up. To compensate for this shortcoming, E-logs in this research were designed in a semi-closed format, containing single choice, multiple choice, and open-ended short answer questions, to reduce the amount of time participants spent on this task, thus reducing their fatigue.

Secondly, it can be beneficial for learners to raise their consciousness of specific learning behaviours and learning strategies by keeping logs. However, for the researcher, to

some extent, this may affect the results of the study. As Fry (1988) concluded, “the act of recording aspects of learning behaviour will raise consciousness of that behaviour and may change it” (p. 161). This means that keeping logs may turn out to be an invisible intervention or prompt in the research setting, causing participants to become more aware of a particular behaviour and thus influencing the outcome of the study. For example, participants may raise their consciousness of the process and benefits of self-regulated learning through log-keeping and apply it to their IDLEL activities. However, prior to the log-keeping task, participants may not act in a self-regulated manner in IDLEL activities. Hence, the logs may lead students to self-regulate their learning. This drawback seems inevitable but can be observed, acknowledged and commented on in the study’s findings. If the log-keeping duration is long enough, in the absence of other interventions, it is possible for the researcher to detect participants’ changes in particular behaviours through the content of the logs, and thus can speculate whether the log is an invisible intervention or prompt for participants to perform particular behaviours. Taking this issue into account, this study positioned the length of the log-keeping task at one month and the researcher did not make any interventions with the participants during this process. Such a design allowed the researcher to gain insight into the process of participants’ engagement in IDLEL activities and their self-regulated learning behaviours on the one hand, and on the other hand, to understand whether the E-log could serve as a prompt to some extent.

Thirdly, retrospective data refers to data “collected after the event” (Fry, 1988, p. 160) and logs can be regarded as an example. Constrained by memory capacity, the later participants report retrospective data after an event, the more likely they are to forget their

complete and accurate perceptions of the event, and therefore the more likely unconscious editing of initial perceptions will occur (Fry, 1988). In order to reduce the potential negative impact of the nature of the retrospective data on research findings, this study required participants to complete weekly E-logs so that they were less likely to forget their initial perceptions and strategies regarding their engagement in the IDLEL activities. In addition, the E-logs allowed participants to edit their work in a real-time manner and save the work for future reediting and submission. Therefore, the researcher suggested that participants record each IDLEL activity immediately after they had undertaken it and submit a complete weekly E-log at the beginning of the following week. Moreover, in the middle of each week, the researcher posted a reminder in the *WeChat* group to remind participants to keep an immediate record of their engagement in IDLEL activities.

Fourthly, as a self-report instrument, E-logs effectively and practically collected self-report data on participants' perceptions of IDLEL activities and self-regulation behaviours that could not be directly observed by the researcher. Even though the validity of self-reported data may often be questioned, such as their ability to reflect participants' true thoughts or behaviours and the social desirability of participants' answers, the researcher compensated for this shortcoming in the current research by allowing participants to answer anonymously, encouraging honest responses, and promising that answers would not be linked to participants' academic results (Fredricks & McColskey, 2012; Garcia & Pintrich, 1996).

As stated above, both qualitative data and quantitative data were collected using E-logs. Regarding quantitative data collection, participants recorded the frequency, diversity, and duration of their engagement in IDLEL activities as well as the reasons for their engagement

and achievements in the various IDLEL activities by answering single and multiple-choice questions. A section of the E-log entitled “How did I engage in IDLEL activities?” collected qualitative data. In this section, participants briefly described how they made use of the IDLEL activities they had engaged in to enhance their English listening. Three open-ended questions were included: “Any mental preparation or action I performed before I started being involved in the activity”; “What did I do during the activity to make it beneficial for improving my listening comprehension?”; and “What did I do after the activity?” These questions were designed to gain insights into learners’ SRL strategy use in relation to IDLEL activities.

In the second stage of the IDLEL study (i.e., Week 2 to Week 5), the weekly E-logs were presented to participants in the form of an online questionnaire via the *WenJuanXing* platform. Specifically, starting from Week 2, following the same procedure as accessing the electronic questionnaires, participants accessed and completed the password-protected weekly E-logs by scanning the QR codes shared by their English teachers in the class WeChat groups every Monday morning via their mobile devices. The E-logs featured an auto-save function, allowing participants to edit and modify their entries at any time before final submission each week. This functionality helped ensure that participants could promptly and accurately record their self-regulated listening behaviours and psychological states after each IDLEL activity, thereby enhancing the reliability of the log data. During the four-week IDLEL study, in addition to completing the tasks assigned by their teachers, participants chose to engage in any number of IDLEL activities and described their engagement in IDLEL activities in their weekly E-logs. They were encouraged to use English in their E-logs but could use Chinese for

expressions they were unsure about or unable to articulate. Participants were then required to submit their E-logs from the previous week at the beginning of the following week (e.g., the E-log for Week 2 was submitted on Monday of Week 3). According to *WenJuanXing* platform, each E-log took approximately 10-15 minutes to complete.

3.3.3 Listening Comprehension Tests

Three listening comprehension tests (see Appendix 4) were conducted at the beginning and end of the IDLEL study (Week 1 and Week 6), and three months after the study ended (Week 19) to assess participants' English listening proficiency.

The listening tests were specifically designed for this study. To examine the short-term effects of IDLEL engagement on listening, the first and second tests were of comparable difficulty (i.e., similar speech rate, linguistic complexity, and topic familiarity, see discussion below), but different listening materials and test items were used to avoid potential short-term memory effects that could influence the results if the same materials were repeated.

The first and the third tests, however, used identical listening materials and test items, with only the order of the passages and the answer options in multiple-choice questions altered. Such a design was based on the following considerations. Firstly, using the same listening materials and test items can maximise score comparability between the two tests, eliminating the influence of test material variation and thus providing a more precise reflection of the long-term effects of IDLEL engagement on L2 listening. The researcher and two experts in the field of L2 acquisition reached a consensus that a three-month interval would be sufficient for participants' memory of the listening materials to fade, reducing the likelihood that they would rely on recollection from the first test when taking the third test.

Additionally, after the first test, the researcher did not disclose the correct answers or the transcripts of the listening materials to the participants. Moreover, in the third test, the order of answer options in multiple-choice questions was adjusted. These measures could help minimise the chances of participants answering the third test based on memory alone, ensuring that the test results accurately reflected changes in their listening proficiency under IDLEL and non-IDLEL research conditions.

3.3.3.1 Listening Materials

The four sets of listening materials used in the listening comprehension tests comprised both authentic listening sources, such as TED Talks, BBC radio programmes and English films, and non-authentic or semi-authentic materials, such as passages from IELTS listening practice tests. This diverse selection of materials was intended to provide a more comprehensive assessment of participants' listening proficiency across both formal (e.g., academic and test-based listening) and informal (e.g., real-life conversational and media-based listening) contexts, compared to relying solely on standardised test materials such as those used in IELTS, TOEFL, or CET.

To ensure that the two listening comprehension tests were of comparable difficulty, the researcher controlled the speech rate, linguistic complexity, and text content of the listening materials used in both tests (Brunfaut & Revesz, 2015), as these factors have been shown to be significant determinants of listening task difficulty (Brindley & Slatyer, 2002; Mohamadi, 2013).

The effect of speech rate on listening comprehension has been widely recognised and it is therefore an important variable in evaluating the difficulty of listening tasks (Griffiths,

1990; Hayati, 2010; Zhao, 1997). In most studies, the speech rate is measured by calculating the number of words per minute (WPM) without excluding silent pauses (Brindley & Slatyer, 2002; Griffiths, 1990; Hayati, 2010; Zhao, 1997). Experimental findings suggest that fast speech rate may hamper listeners' comprehension while slower speech rate can facilitate their comprehension (Griffiths, 1990; Hayati, 2010; Zhao, 1997). This may be because the slow speech rate allows listeners more time to perceive and process the input information, thus facilitating their comprehension of the listening text. Fast speech rate, however, may cause the listener to miss a lot of information and result in perception problems, which will negatively influence their comprehension (Nowrouzi et al., 2015; Tran & Duong, 2020).

Table 3 shows the speech rate (WPM) of each listening passage in the two listening tests.

Table 3

The Speech Rate (WPM) of Each Listening Passage in the Two Listening Tests

Listening Test 1						
	Passage 1	Passage 2	Passage 3	Passage 4	Total	Average
Duration						
(min)	1.73	1.21	1.79	1.82	6.55	≈1.64
Number of words						
	257	198	367	156	978	≈244
Speech rate						
(wpm)	146	163	205	85	N/A	≈149
Listening Test 2						
	Passage 5	Passage 6	Passage 7	Passage 8	Total	Average

Duration	1.38	1.10	1.96	1.97	6.41	≈ 1.60
(min)						
Number of words	266	170	310	164	910	≈ 227
Speech rate	192	154	158	83	N/A	≈ 141
(wpm)						

The most straight forward way to decide whether passages are of comparable difficulty in terms of speech rate is to use commonly employed bands that represent slow, medium, and fast, and so on (Griffiths, 1990, 1992; Zhao, 1997). These bands have a range of 50-70 wpm. As can be seen from Table 3, therefore, the four passages in Listening Test One and the four passages in Listening Test Two had a comparable average speech rate (149 wpm and 141 wpm respectively). Therefore, two listening comprehension tests can be considered to have comparable levels of difficulty in terms of the speech rate of listening passages.

Unfamiliar topics and vocabulary in listening passages may also hamper learners' comprehension (Bloomfield et al., 2010; Mohamadi, 2013). Unfamiliar topics may prevent listeners from using background knowledge about a topic stored in long-term memory to aid comprehension, while unfamiliar or infrequent vocabulary may directly impede listeners' word recognition during listening, thus hindering their listening comprehension. Table 4 displays the topics and potential unfamiliar vocabulary of listening passages.

Table 4*The Topics and Potential Unfamiliar Vocabulary of Listening Passages*

Listening Test 1				
	Passage 1	Passage 2	Passage 3	Passage 4
Topic	daily life: dream & sleep	lifestyle: recycling	language learning	life attitude
Potential		upcycled		
unfamiliar	segmented	ornaments	polyglots	None
vocabulary		Eskilstuna		
		Stockholm		
Listening Test 2				
	Passage 5	Passage 6	Passage 7	Passage 8
Topic	daily life: diet & obesity	lifestyle: food waste	language learning	life attitude
Potential		leftover	apprehensive	
unfamiliar	None	niche		None
vocabulary		breweries		

According to Table 4, the topics of the listening passages in the two listening tests were generally consistent and closely related to participants' lives and studies. Regarding vocabulary, the first-year participants from University A were preparing for the TEM-4 at the time of this study, while the second-year participants from University B had just completed

the TEM-4. Therefore, it can be assumed that most participants' vocabulary levels met the requirements of the TEM-4.

A comparison of the vocabulary in the eight listening passages from the two listening comprehension tests with the TEM-4 vocabulary syllabus revealed that no more than ten words (six for Listening Test One and four for Listening Test Two) were outside of the syllabus. In addition, these words were considered to be unlikely to pose a significant obstacle to participants' completion of the listening comprehension tasks. The meaning of "upcycled" was explained in the passages, and the meaning of "segmented", "leftover", "niche", and "breweries" could be inferred from the passage. Other words, such as "polyglots", "Eskilstuna", and "Stockholm", did not have direct relevance to the comprehension tasks. In summary, the two listening tests were of comparable difficulty in terms of the topics and vocabulary of the listening passages.

3.3.3.2 Listening Tasks

Two types of tasks were included in each listening test: free-recall tasks and multiple-choice questions. Each listening comprehension test contained four free-recall tasks and eight multiple-choice questions.

3.3.3.2.1 Free-Recall Task

In the free-recall task, participants heard four audio passages. Every passage was only played once, and immediately after each passage participants were required to write down everything that they could understand in either English (L2) or Chinese (L1). There are three reasons for including the free-recall task in the listening tests. First, they have been employed as effective tools to explore participants' listening comprehension in many studies (Graham et

al., 2008; Graham & Macaro, 2008; Leeser, 2004; Simasangyaporn, 2016; Vogely, 1995).

Free-recall tasks can not only reflect participants' understanding of the main ideas of the listening passages but can also indicate their grasp of the details of the listening input, thus providing a comprehensive picture of listening comprehension (Graham et al., 2008; Riley & Lee, 1996). Second, compared to other assessment measures, such as multiple-choice questions, free-recall is able to measure listeners' comprehension of the listening passages in a straightforward manner as it does not require listeners to guess based on the question prompts, but only to write down all their understanding of the listening input (Berkemeyer, 1989). Third, free-recall can minimise the confounding effect of other language skills, such as reading and speaking, on the listening assessment. Although participants were required to write down their understanding, the influence of writing skills on listeners' listening comprehension was minimised by allowing listeners to use their first language to answer, making it possible for listening skill to be tested independently.

Concerns about free-recall tasks primarily focus on discussions of the memory confounding effect (Sakai, 2005). Due to the limited capacity of human memory, longer listening passages may pose a challenge to memory, making it difficult for participants to write complete or correct answers. This is because they may forget the information they have already understood, thereby affecting the test results.

To address this issue, in both listening tests, all listening materials except for Passage Four were segmented into two audio clips based on time and semantic structure, ensuring that the materials remained coherent. A designated pause (i.e., one minute) was provided between the two clips and after the second clip to allow participants to write down the information

they had understood from both parts. This design was believed to help to reduce participants' memory burden, allowing them to focus more on comprehending the content they were listening to. Such an approach is also commonly used in standardised listening tests. For example, in IELTS listening tests, listening materials are sometimes played in segments, with multiple-choice questions and other tasks immediately following each segment to assess candidates' comprehension. Additionally, in real-life listening scenarios, such as watching an English movie on a mobile device, it is common to pause briefly to recall the information just understood; similarly, in academic lectures or speeches, professors and speakers often use pauses to help listeners recall and process previously delivered information.

The audio for Passage 4 in both tests was not segmented because, compared to the other listening materials, its speech rate was slower, reducing the necessity for additional pauses. Moreover, these two passages were dialogues, which are more coherent and difficult to segment without disrupting their natural flow.

3.3.3.2.2 Multiple-Choice Tasks

The second listening comprehension task was multiple-choice questions. Learners heard the four passages played in the free-recall task again and answered questions after each one. Every passage was only played once. Participants had thirty seconds to go through the questions before listening and two minutes to answer questions of each passage.

Multiple-choice is a type of widely used listening test format “with a stem and three or more options from which learners are required to select one” (In’nami & Koizumi, 2009, p. 220). The multiple-choice format has many strengths including “high marker reliability, ease of scoring, and objective scoring” (Hemmati & Ghaderi, 2014, p. 639). In addition, it has

been shown that multiple-choice questions are easier than open-ended questions (In’nami & Koizumi, 2009). Hence, given that the first task in this study was the free-recall task with an open-ended question format, selecting multiple choice questions as the second task was believed to balance the difficulty level of the tests to some extent. Moreover, the scoring criteria for free-recall tasks may involve a certain degree of subjectivity, as different participants may express the same information in various ways, making the evaluation process somewhat dependent on subjective judgment. In contrast, multiple-choice questions can provide clear, standardised answers, reducing subjectivity in scoring and making the test more objective. This can ensure that participants’ scores can be directly compared, thereby enhancing reliability. Furthermore, the free-recall task in this study mainly focused on reflecting participants’ understanding of the main ideas rather than systematically covering all key details. Multiple-choice questions, however, can more precisely assess participants’ comprehension of critical details, such as important numbers, times, locations, specific cause-and-effect relationships, or the speaker’s attitude and implicit information. As a result, multiple-choice questions can help compensate for potential omissions in free-recall tasks, allowing the researcher to gain a more comprehensive understanding of participants’ overall listening comprehension proficiency.

However, when using multiple-choice questions as a measure of listening proficiency, problems like “guessing, limitation in testing different components of a language, difficulty in preparing the successful items, harmful backwash” (Hemmati & Ghaderi, 2014, p. 639) as well as the confounding effect of reading skills should be taken into account. Considering that different formats of multiple-choice questions may result in different levels of difficulty, the

full question preview format was used in this study, whereby (participants are allowed to see both the stem and answer options before listening. That format can facilitate participants' listening comprehension (Hemmati & Ghaderi, 2014).

The current research adopted a three options per question format, which, compared to five-option and four-option items, is considered to offer lower levels of difficulty as well as appropriate item discrimination and reliability (Dehnad et al., 2014; Rodriguez, 2005). Among the three options of each question, only one of them was the correct answer and all the options were as similar in length as possible, reducing the likelihood of participants guessing the answer by the length of the option.

The listening tests took place during normal face-to-face English listening lessons. Although the two universities participating in the IDLEL study had resumed in-person classes, the COVID-19-related isolation measures remained in place. That meant that if paper-based answer sheets were used, the researcher would not be able to immediately collect them for data analysis. Therefore, the listening test answer sheets were distributed to participants as password-protected online questionnaires via *WenJuanXing*. Following the same procedure as accessing the electronic questionnaires and E-logs, participants scanned the QR code shared by the teachers in the class WeChat group using their electronic devices to obtain the password-protected electronic listening test answer sheet before the listening test. Once all participants had successfully accessed the answer sheets, either in electronic or paper format, the teacher began playing the listening audio, continuing until the test ended. Participants then submitted their electronic answer sheets. Each listening test lasted approximately 35 minutes.

3.4 Pilot Study

A pilot study refers to “a small-scale version of a planned study conducted with a small group of participants similar to those to be recruited later in the larger scale study” (Doody & Doody, 2015, p. 1074). In the present study, a pilot study was therefore conducted to test and refine the data collection and analysis techniques, identify potential methodological issues, and allow adjustments before the main study began (Doody & Doody, 2015).

3.4.1 Procedure of Pilot Study

The pilot study took place in the first semester of 2022 at University A, lasting for a total of four weeks. Selecting participants for the pilot study who are as similar as possible to those in the main study can be beneficial for ensuring the effectiveness of research instruments, assessing the appropriateness of tasks, maintaining consistency in experimental procedures, and identifying potential issues to optimise the research process. The planned participants for the main study in this research were first- and second-year English majors at Chinese universities, therefore, 33 Chinese EFL first-year undergraduates (13 males and 20 females) majoring in English were recruited to participate in the pilot study. In Week One, the first listening test (Appendix 4) was completed by 31 participants and the questionnaire (Appendix 2) by 26 participants. In Week Two, participants were told to choose to complete as many IDLEL activities as they wished in addition to the work set by their teacher. Two participants (one male and one female) volunteered to keep a record of their IDLEL engagement in E-logs (Appendix 3), which they submitted to the researcher via *WenJuanXing* by the Monday of Week Three. The second listening test (Appendix 4) was due to take place in Week Three, but did not, due to unexpected changes in teaching arrangements. Instead, it took place in Week Four and was completed by 31 participants. Teaching schedule constraints meant that the

second listening questionnaire and the third listening test could not be included in the pilot study as planned. The pilot study timeline is shown in Table 5.

Table 5

Pilot Study Timeline

	Phase 1	Phase 2	Phase 3
Week	1 (24/5/2022)	2	3
Action	First English listening comprehension test; English listening questionnaire (self-regulation, anxiety, self-efficacy)	E-logs	Delayed Second English listening comprehension test (14/5/2022)

3.4.2 Data Analysis and Results for the Pilot Study

The primary aim of this pilot study was to simulate the main procedures of the main study, assess the effectiveness of research instruments, and identify potential issues in the research process. Factors such as the small sample size, limited E-log data, and the absence of data from the second listening questionnaire and the third listening test meant that the data analysis in the pilot study could not be carried out in a way that would answer the research questions. Instead, it primarily focused on evaluating the scoring criteria of the listening tests and checking the reliability of the questionnaire. Beyond this, the researcher conducted basic descriptive analyses of the questionnaire and listening test results and performed a paired-samples t-test to compare the two listening test results.

3.4.2.1 English Listening Questionnaire

The descriptive analysis of the SRLLQ (Appendix 2) revealed that participants ($N=26$) used various SRL strategies to self-regulate their English listening ($M=4.6$, $SD=.25$). The most used SRL strategy was “*When I receive a listening task, I try to figure out what the demands of the task are*” ($M=5.35$, $SD=.75$), while the least used was “*I check whether my understanding of the passage is consistent with my linguistic knowledge*” ($M=4.12$, $SD=1.12$). In addition, the descriptive analysis of the LLAS revealed that participants ($N=26$) exhibited a certain degree of English listening anxiety ($M=3.98$, $SD=.42$). However, at the same time, they also demonstrated a relatively high level of listening self-efficacy ($M=63.67$, $SD=2.92$). This suggests that although some students experienced anxiety during listening tasks, they still perceived themselves as capable of handling them. Particularly, participants had the weakest sense of self-efficacy in the typical VOA/ BBC English radio programme setting ($M=57.79$, $SD=5.96$), while they had the strongest sense of self-efficacy when they encountered listening passages in the textbook ($M=67.12$, $SD=5.61$). Moreover, in each listening context, participants had the greatest confidence in extracting the main ideas of listening passages and were least sure that they could infer the meaning of unknown or incomprehensible words.

3.4.2.2 English Listening Tests

Scoring is the prerequisite for analysing qualitative data and quantitative data collected through listening comprehension tests. In terms of the free-recall task used in the listening test (Appendix 4), the most frequent used approach for scoring is to calculate the number of

correct idea units in the listener's answers (Berkemeyer, 1989; Graham et al., 2008; Leeser, 2004; Riley & Lee, 1996).

As stated in *Section 3.3.3.2.2*, the free-recall task in the present study focused on assessing participants' overall comprehension of the listening material rather than word-for-word memorisation. Therefore, inspired by Macaro and Erler's (2008) scoring criteria for reading comprehension tasks, the scoring of the free-recall task in this study was based on whether participants could accurately capture the main ideas and express them in a clear and coherent manner. Five idea units (propositions) were pre-identified in each listening passage with each idea unit representing a key piece of information or a main idea from the passage. Participants' responses were then scored as follows: one point was awarded if they accurately expressed a complete idea unit; 0.5 points were given if the response contained partial information related to the idea unit but lacked completeness or key details; and 0 points were assigned if the response was incorrect or did not mention the idea unit. Participants' total score for the free-recall task of listening comprehension tests, therefore, was the sum of the number of correct unit ideas they recalled. The maximum score for each listening passage was five points, and since each listening test included four passages, the total score for the free-recall task in each listening test was 20 points. All recall protocols were independently rated by two raters as outlined below. The two raters' scores were then compared to assess whether the ratings were consistent. Any scoring discrepancies were resolved through discussion.

To identify the idea units contained in the listening materials, firstly, the researcher of the current study and another PhD student majoring in L2 education separately listed the

potential idea units (main ideas in the current study) in each listening passage. Then two versions of the idea units were compared and, after consultation between the two raters, a revised list of idea units for scoring purposes was produced (see Appendix 5). To further ensure the reasonableness of the idea units, the revised idea unit list was shared with two experts in the field of L2 acquisition for a second round of review. After the researcher and the two experts reached a consensus on the new revisions, the final list of idea units was established and used for scoring the free-recall tasks in the listening tests.

Regarding the multiple-choice question task, each question was marked with one point. Each listening test contained eight multiple-choice questions, so the total score for the multiple-choice task in each listening test was eight points and the total score for each listening test was 28 points.

The descriptive analysis of the two listening comprehension test scores revealed that the mean score of the first listening tests was 10.81 ($SD=4.44$), and the mean score of the second listening test was 12.48 ($SD=3.42$). Specifically, the mean score for the free-recall tasks in the first listening test was 5.16 ($SD=3.40$) while the mean score for the multiple-choice question tasks was 5.65 ($SD=1.76$). For the second listening test, the mean free-recall tasks score was 6.77 ($SD=2.91$) while the mean multiple-choice question task score in was 5.71 (8 in total, $SD=1.10$). In other words, there was a small improvement between the two tests across both tasks.

3.4.2.3 Analysis of English Listening Tests

The pilot study collected E-log data from only two participants over one week meaning that a systematic data analysis could not be conducted. However, the researcher gained

valuable insights from examining the limited data, which could be beneficial for the E-log data analysis in the main study. Specifically, the E-logs of the two participants indicated that their descriptions of engaging in IDDEL activities were related to the SRL strategies embedded within the five phases of the self-regulated L2 listening model. For example, participants mentioned that their purposes for engaging in different listening activities varied, leading to differences in the listening strategies they adopted. One participant mentioned that when listening to test-related materials on her phone, she tended to focus on understanding every word rather than just grasping the main idea. Driven by this goal, if she encountered difficulties during listening, she would either listen repeatedly or use Chinese subtitles to aid her comprehension until she could understand every word. However, if her goal was to relax and learn about the target culture by watching an English film, she would not attempt to understand every word but would instead focus on grasping the main ideas. If she came across interesting words, she might imitate their pronunciation but would not necessarily write them down deliberately.

Thus, the E-log data from the pilot study suggested that the qualitative E-log data in the main study could be well-suited for thematic analysis. For instance, the data could be coded and categorised based on SRL strategies and SRL phases.

3.4.3 Reliability and Validity

Rigour refers to “the extent to which researchers work to enhance the quality of their studies” (Heale & Twycross, 2015, p. 66), and enhancing reliability and validity can be an effective way to ensure research rigour. In quantitative research, reliability concerns the consistency of measurement, meaning “the extent to which a research instrument consistently

produces the same results when used in the same situation on repeated occasions" (Heale & Twycross, 2015, p. 66). In other words, it relates to "the repeatability of results or observations" (Golafshani, 2003, p. 598). A commonly used indicator for measuring the reliability of a research instrument is Cronbach's α coefficient, which assesses the internal consistency of a measurement tool, determining whether the items in a questionnaire or test measure the same construct (Vaske et al., 2017). A commonly used rule of thumb for evaluating internal consistency suggests that $\alpha \geq .9$ = Excellent, $\alpha \geq .8$ = Good, $\alpha \geq .7$ = Acceptable (George & Mallery, 2003). In the present pilot study, the internal consistency coefficient of the SRLLQ ($N=26$), LLAS ($N=26$), and LSEQ ($N=26$) were .98, .89, and .99 respectively. Therefore, the questionnaires devised for this study had high reliability.

Inter-rater reliability assesses the degree of consistency between two or more raters when independently scoring the same data (Heale & Twycross, 2015). To evaluate whether the ratings are stable and consistent, research typically involves multiple raters independently scoring open-ended questions or subjective items, followed by a comparison of their ratings to determine inter-rater reliability. Pearson's correlation coefficient is a commonly used measure for assessing inter-rater reliability (Graham et al., 2008; Macaro & Erler, 2008). In this pilot study, the correlation coefficients for the free-recall task in the two listening tests were .92 and .94, with all correlations being statistically significant ($p < .01$). Therefore, the free-recall task in both listening tests can be considered to have had high inter-rater reliability.

As for validity in quantitative research, it refers to the extent to which an instrument measures what it is intended to measure (Heale & Twycross, 2015). Two types of validity that are frequently discussed in quantitative research are content validity and construct validity.

Content validity refers to the extent to which an instrument measures all aspects of a variable. It is often evaluated by “using a panel of experts to evaluate instrument elements and rate them based on their relevance and representativeness to the content domain” (Almanasreh et al., 2019, p. 216). A similar concept to content validity is face validity, whereby “experts are asked their opinion about whether an instrument measures the concept intended” (Heale & Twycross, 2015, p. 66). Face validity can be regarded as a sub-type of content validity (Heale & Twycross, 2015), however, compared to content validity, face validity seems to be more informal. The instruments employed in this study were examined and discussed by several experts in the language research domain, thus the content validity and face validity can be assured to a large extent.

As for construct validity, it relates to the extent to which an instrument assesses the intended variable and can be measured through the confirmatory factor analysis (CFA) or exploratory factor analysis (EFA). The sample size of the questionnaire ($N=26$) collected in the pilot study did not meet the requirements for factor analysis, which typically requires the number of participants to be at least five times the number of questionnaire items (Sun et al., 2021). Therefore, the pilot study was insufficient to test construct validity; however, CFA was conducted on the questionnaire in the main study.

Reliability and validity are controversial concepts in qualitative research. While many scholars believe that reliability and validity are essential factors that qualitative researchers should consider when conducting studies (Johnson, 1997; Patton, 1990), others argue that reliability and validity, while widely used in quantitative research, cannot be used as a measure of research quality due to the different epistemological and ontological assumptions

and the different purposes of measuring the research quality of two types of research (Golafshani, 2003; Johnson, 1997). In addition, some scholars claim that it is necessary to assess the quality of qualitative research, but that reliability and validity are criteria for evaluating the quality of quantitative research, and qualitative research should have quality assessment criteria that fit its own paradigm (Healy & Perry, 2000). As a result, various terms corresponding to the concepts of reliability and validity but specifically used to evaluate the quality of qualitative research, have been proposed, such as “Credibility, Neutrality or Confirmability, Consistency or Dependability and Applicability or Transferability” (Golafshani, 2003, p. 601). Although the concepts of reliability and validity may be expressed in other terms in qualitative research, it can be seen that in qualitative research, reliability relates to “the degree to which other researchers performing similar observations in the field and analysis...would generate similar interpretations and results” (Franklin & Ballan, 2001, p. 356) while validity concerns credibility, which involves the “truthfulness” of the results of the research (Franklin & Ballan, 2001).

To improve the reliability and validity of qualitative research, a triangulation strategy can be helpful. By using multiple data sources (data triangulation), research methods (methods triangulation), theories (theory triangulation), or investigators (investigator triangulation) for data collection and analysis, qualitative researchers can obtain more “valid and reliable multiple and diverse realities” (Golafshani, 2003, p. 604), which is consistent with the philosophical assumptions of qualitative research (Johnson, 1997). The present study employed both quantitative and qualitative research methods, used multiple instruments to collect data, combined different theories as the theoretical framework, and invited multiple

researchers to analyse qualitative data, maximising the rigour of the study. In addition, to enhance the validity of the self-report data collected from E-logs, the researcher encouraged honest responses and promised that participants' answers would not be linked to their academic results (Fredricks & McColskey, 2012; Garcia & Pintrich, 1996).

3.4.4 Lessons & Implications

One of the major purposes of the pilot study was to simulate the research process and identify any potential gaps so that they could be addressed before the main study. The researcher identified several flaws in the pilot study, particularly in terms of data collection methods and instrument design. These limitations provided valuable insights for refining the research and instrument design in the main study.

Firstly, the sample size of the questionnaire survey in the pilot study was insufficient to allow its validity to be assessed. Seven participants did not submit the questionnaire, even though the researcher had asked the teacher to remind them to complete it on time. This issue may have arisen for two reasons. The first reason was that the questionnaire was not administered in the same classroom immediately after the listening test. Participants may have had other learning activities after class and therefore did not have time to complete the questionnaire. The second reason was that, since the questionnaire could be completed out-of-class, students did not attach as much importance to it as they did to the listening test.

To address this issue, participants in the main study were asked to complete the questionnaire in the same classroom immediately after the listening test. On the one hand, this approach helped the researcher collect adequate questionnaire data for further analysis. On the other hand, it enabled participants to describe their self-regulatory behaviours and feelings

about English listening more accurately, based on their immediate experience of having just taken the listening test.

Secondly, the electronic answer sheets for the second listening test did not include the “back to previous page” function. Some participants accidentally skipped to the next page during the test but were unable to return to the original page to continue, which affected the results to some extent. The researcher addressed this issue in the main study by adding the “back to previous page” function in the electronic answer sheet.

Thirdly, some participants reported that the time allocated for reading and answering questions in the multiple-choice task was too long. Therefore, in the main study, the previewing time for questions and options in the multiple-choice task was reduced from one minute to 30 seconds, and the answering time was shortened from two minutes to one minute.

Fourthly, as the E-log data collected in the pilot study were limited, a systematic data analysis was not conducted. However, the examination of the data inspired the researcher to code and categorise participants’ E-log data in the main study based on the self-regulated L2 listening model and relevant SRL strategies for thematic analysis.

3.5 Data Analysis

In the main study, both quantitative and qualitative data were collected to address the research questions. The quantitative data were gathered through the multiple-choice questions in the listening tests, the English listening questionnaires, and the multiple-choice questions in the E-logs. The qualitative data were primarily collected through the free-recall tasks in the listening tests and the open-ended questions in the E-logs. The analysis methods for research questions are detailed in this section.

3.5.1 RQ1: Can the hypothesised five-phase dual-level construct of the self-regulated L2 listening model be validated?

3.5.1.1 Confirmatory Factor Analysis (CFA)

To validate the structure of the self-regulated L2 listening model, the *Self-Regulated L2 Listening Questionnaire (SRLLQ)* (Appendix 2) was developed based on the model as its theoretical framework. Specifically, the five phases of the model corresponded to the five constructs in the SRLLQ, with the listening self-regulation and motivational/affective regulation strategies frequently used by participants in each phase serving as the measurement items for the respective constructs. If the structure of the SRLLQ could be confirmed, the five-phase structure of the self-regulated L2 listening model would thus be validated. To achieve this, confirmatory factor analysis (CFA) was conducted on the SRLLQ data collected from 582 participants through the questionnaire survey.

CFA is an analytical method used to “test a theoretical model and a tighter specification of multiple hierarchies by utilizing the factor, correlation, covariance patterns, and residual or error values within a data matrix” (Teng & Zhang, 2016, p.11). It is commonly employed when researchers have expectations regarding the number of factors and the structure of the model and wish to validate those expectations.

As mentioned in *Section 3.2.1*, the number of participants required for CFA should be no less than five times the number of questionnaire items (Sun et al., 2021). That is, the CFA sample size for this study should be no less than 230, given that the SRLLQ consisted of 46 items. Therefore, the 582 SRLLQ responses indicated that the sample size in this study was sufficient for conducting CFA. Before conducting CFA, the researcher first screened the

questionnaires to ensure valid completion and excluded any invalid responses to prevent them from affecting the analysis results. A total of 59 responses were identified as invalid due to unusually short response times or identical answers for all questions and were removed. The remaining 523 valid SRLLQ responses were retained for further analysis.

The screened data were first subjected to the Mardia Multivariate Normality Test (Mardia, 1970) in R (R Core Team, 2024), using the “MVN” package (Korkmaz, 2014), to assess normality, which determined the estimator used in CFA. The results showed that the *p*-values for the skewness and kurtosis statistics of SRLLQ were both below .05, indicating that the data did not follow a normal distribution (Oppong & Agbedra, 2016). Consequently, the robust maximum likelihood estimator, which is suitable for non-normally distributed data, was employed to perform CFA in R using the “semTools” package (Jorgensen et al., 2022), on the SRLLQ data (*N* = 523) to validate the underlying constructs of the questionnaire.

As a multivariate statistical method, CFA can be used to examine the fit of a measurement model, determining whether the observed data align with the theoretical model. If the model fit is acceptable, it may indicate that the structure of the theoretical model can be supported by empirical data. In the field of L2 research, commonly used indices for evaluating model fit include : the ratio of χ^2 to its degrees of freedom (χ^2/df), the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), the comparative fit index (CFI), and the Tucker-Lewis index (TLI) (In’nami, 2006; Sun et al., 2021; Xu & Huang, 2018). Specifically, a χ^2/df value less than 3.0, RMSEA and SRMR values less than .08, and CFI and TLI values greater than or equal to .90 all indicate an acceptable model fit (Hu & Bentler, 1999; Sun et al., 2021; Tucker & Lewis, 1973).

After conducting CFA, some items that failed to effectively measure their intended constructs were removed from the original SRLLQ, resulting in a revised and final version of the SRLLQ (see *Section 4.1.1*). Subsequently, the convergent validity and internal consistency reliability of the five constructs were assessed. Specifically, convergent validity was evaluated by calculating the Average Variance Extracted (AVE) using the “summary_model” function in the “semTools” package. AVE measures the shared variance among the observed variables within the same latent construct, indicating the extent to which these observed variables explain the variance of the underlying construct (Hair & Alamer, 2022). Conventionally, an AVE value of .50 or higher is considered evidence of adequate convergent validity (Hair et al., 2017). Additionally, internal consistency reliability, which measures the reliability of the items within each construct, was assessed using Cronbach’s α and composite Reliability (CR) via the “compRelSEM()” function in the “semTools” package. Similar to the case of Cronbach’s α (George & Mallery, 2003), a widely accepted rule of thumb suggests that CR values above .70 are considered acceptable (Hair et al., 2017). As CR tends to be a more lenient measure of reliability than Cronbach’s α , it is recommended to use CR as the upper bound and Cronbach’s α as the lower bound (Hair et al., 2019), providing a more comprehensive evaluation of internal consistency reliability.

This validated SRLLQ was then used as a tool for assessing participants’ self-regulated L2 listening in subsequent analyses of the current study and can also be applied in future research.

3.5.2 RQ2: Can the two hypothesised mechanisms through which L2 listening is jointly predicted by listening self-efficacy, listening anxiety, and self-regulation be validated?

Before validating the hypothesised joint mechanisms of listening, the self-regulation, self-efficacy, and listening anxiety of IDLEL-study participants ($N = 130$) needed to be measured. As previously mentioned, the validated SRLLQ served as the instrument for assessing participants' listening self-regulation. Similarly, to measure participants' listening self-efficacy and listening anxiety, the *L2 Listening Anxiety Scale (LLAS)* (Appendix 2) and the *Listening Self-efficacy Questionnaire (LSEQ)* (Appendix 2) also required validation. Similar to the validation process of the SRLLQ, since the researcher had prior expectations regarding the number of factors and the structural composition of LLAS and LSEQ (as outlined in *Section 3.3.1.2* and *Section 3.3.1.3*) and aimed to validate these expectations, CFA was employed.

Following the same procedure as the CFA conducted for the SRLLQ data, the 582 LLAS and LSEQ data entries were first screened and cleaned. Subsequently, the Mardia Multivariate Normality Test was performed in R to assess their normality. The results indicated that the p -values for the skewness and kurtosis statistics of the screened LLAS and LSEQ data were both below .05, suggesting that the data did not follow a normal distribution (Oppong & Agbedra, 2016). Consequently, the robust maximum likelihood estimator was employed to conduct CFA in R on the remaining LLAS data ($N = 427$) and LSEQ data ($N = 435$), and model fit was evaluated based on fit indices (e.g., χ^2/df , TLI, CFI, RMSEA, and SRMR). Ultimately, the validated versions of LLAS and LSEQ, with items that failed to

effectively measure their respective constructs removed, were used to assess the listening anxiety and listening self-efficacy of participants of the IDDEL study ($N = 130$).

It is important to note that the data used to validate the joint predictive mechanisms, including self-regulation, self-efficacy, listening anxiety, and listening test results, were all derived from the pre-test listening test and pre-test questionnaires administered before the four-week IDDEL study began. This design was chosen because post-test data might be influenced by participants' learning experiences during the four-week IDDEL study. If post-test data were used, the researcher might not be able to distinguish whether the relationships between variables stemmed from the predictive mechanisms themselves or were affected by the IDDEL study. Therefore, such a design helped to minimise the impact of the IDDEL study on the data, enhancing the stability and generalisability of the predictive mechanisms. Finally, the scoring method for the listening tests in the main study was consistent with that used in the pilot study. For a detailed explanation, refer to *Section 3.4.2.2*.

3.5.2.1 Structural Equation Modelling (SEM)

Structural Equation Modelling (SEM) was conducted in R using the “semTools” package to validate the hypothesised joint predictive mechanism of SRL, self-efficacy, and listening anxiety on listening comprehension. Before conducting SEM, the normality of the new data ($N = 130$) was assessed in R using the Mardia Multivariate Normality Test. The results indicated that the p -values for the skewness and kurtosis statistics of the new SRL, self-efficacy, and listening anxiety data were all below .05, suggesting that the data did not follow a normal distribution (Oppong & Agbedra, 2016). Therefore, robust maximum likelihood estimation was employed to conduct CFA to evaluate the fit between the new data

and the previously validated models of SRL, self-efficacy, and listening anxiety. After the correlation analysis, SEM was performed.

There are different approaches to executing SEM, including Covariance-Based Structural Equation Modelling (CB-SEM) and Partial Least Squares Structural Equation Modelling (PLS-SEM) (Hair & Alamer, 2022). The latter was chosen for the following reasons. Firstly, unlike CB-SEM, which is primarily used for testing well-established theories that have been empirically validated, PLS-SEM does not require the theory being tested to have strong empirical support (Lowry & Gaskin, 2014). Therefore, PLS-SEM is suitable for exploratory research testing developmental theories, which was the case for this study.

Secondly, unlike CB-SEM, PLS-SEM does not have strict requirements for the normality of data distribution and sample size (Dash & Paul, 2021; Hair & Alamer, 2022). Considering the non-normality of the questionnaire data in the current study, PLS-SEM would be more applicable. Additionally, it has been suggested that the minimum sample size for SEM using approaches like CB-SEM is 200 (Dash & Paul, 2021). However, PLS-SEM can perform well with a smaller sample size of 69 to 100 (Hair & Alamer, 2022). Hence, given the sample size of 130 in this study, PLS-SEM seemed most appropriate. Thirdly, PLS-SEM is better suited than CB-SEM for handling formative constructs, which indicate that the latent variable is formed or defined by multiple observed variables (Hair & Alamer, 2022). In this study, SRL can be considered a formative construct defined by its five sub-phases, namely *Task Representation, Goal Setting & Strategy Planning, Performance, Attribution & Adjustment*, and *Monitoring & Control* phase. Self-efficacy and listening anxiety are further formative constructs, respectively defined by the components they

encompass. In summary, from the perspectives of research objectives, sample size, data distribution characteristics, and the relationship between latent and observed variables, PLS-SEM is more applicable to this study.

Moreover, PLS-SEM does not assess the overall “fit” of a proposed model by comparing observed and proposed covariance matrices as CB-SEM does (Hair et al., 2022). Instead, PLS-SEM is “developed as a predictive approach similar to multiple regression analysis”, helping researchers to understand “the relationships, directionality, and predictive capability between variables” (Hair & Alamer, 2022, p. 9). Therefore, model fit indices commonly used in CB-SEM studies (e.g., χ^2 , df, p, χ^2/df , CFI, RMSEA) are typically not applicable to PLS-SEM (Sharma et al., 2019). Rather, to understand the pathways and mechanisms of the influence among variables, the evaluation of the PLS-SEM model mainly focuses on assessing the significance and size of the path coefficients (Hair & Alamer, 2022). Furthermore, PLS-SEM path analysis does not generate *p*-values by default (Hair & Alamer, 2022). Therefore, to calculate *p*-values and confidence intervals, a bootstrap method with 5,000 iterations and the bias-corrected confidence interval method were used.

3.5.2.2 Mediation Analysis

The hypothesised joint predictive mechanisms involve multiple mediation pathways, such as the mediating roles of SRL and listening anxiety between self-efficacy and listening, as well as the mediation of self-efficacy between SRL and listening. The mediation analysis was conducted in a step-by-step manner. First, the significance of the indirect effect (a * b) was assessed using bootstrapping, where “a” represents the direct effect of the independent variable on the mediator, and “b” represents the direct effect of the mediator on the dependent

variable. Mediation is considered valid only when the indirect effect is statistically significant (Zhao et al., 2010). Next, the significance of the direct effect of the independent variable on the dependent variable was examined. If the indirect effect is significant but the direct effect is not, it indicates indirect-only mediation (Zhao et al., 2010). If both the indirect and direct effects are significant and in the same direction, it suggests complementary mediation, whereas if both are significant but in opposite directions, it indicates competitive mediation (Zhao et al., 2010). The mediation analysis was also conducted in R using the “semTools” package.

3.5.3 RQ3: What are the characteristics of Chinese undergraduates' IDLEL Engagement, in terms of its quantity (i.e., frequency and duration) and quality (i.e., diversity and strategy-use)?

To explore the characteristics of participants' engagement in IDLEL activities, namely frequency, diversity, duration, and strategy use, the E-log data were subjected to three types of analyses: descriptive analysis, thematic analysis, and cluster analysis. The frequency, diversity, and duration of IDLEL activity engagement were measured quantitatively in the E-logs (Appendix 3) and could thus be directly subjected to descriptive analysis. The SRL strategies used by participants in IDLEL activities were reflected through their descriptions of their mental, psychological, and physical “actions” before, during, and after each IDLEL activity, which were qualitative data in nature. As for qualitative data, however, before conducting descriptive analysis, they needed to be subjected to thematic analysis to identify

key themes that described the research phenomenon, which in the present study is self-regulated listening strategies (Kuckartz, 2019).

It should be noted that not all of the 130 students participating in the study completed both valid and usable E-logs, listening tests and questionnaires. Specifically, ninety-one (out of 130) participants completed valid E-Logs, two listening tests (pre- and post-tests) and two questionnaires (pre- and post-questionnaires). Their E-log data were analysed. Of these 91, 37 were second-year students from University B and 54 were first-year students from University A. Additionally, of the 91 participants, only 60 (32 from University B, 28 from University A), completed the delayed post-test conducted three months (Week 19) after the listening post-test. This was because 26 of them from University A were assigned to classes that were not part of the study in the new semester, and five of them from University B were absent from the third listening test.

It should also be noted that the number of E-logs completed by participants during the four-week study varied. Hence, when addressing research questions, such as exploring the relationship between participants' IDLEL activity engagement and their L2 listening, self-regulation, self-efficacy, and listening anxiety, or establishing profiles of participants' L2 individual differences, participants' weekly IDLEL activity engagement data were used instead of their overall engagement throughout the four-week study. This was done in order to help eliminate bias caused by the differing number of E-logs, making the data more comparable.

3.5.3.1 Thematic Analysis

The first step of thematic analysis is to become familiar with the data (Braun & Clarke, 2006). In this study, many participants used their native language, namely, Chinese, to complete their E-logs (Appendix 3). The method of maintaining the original language during the analysis process was used and only the most critical findings were translated into English for result presentation (Al-Shboul et al., 2013). The method was employed to minimize the time required for data processing and the misinterpretation of participants' original intentions due to language differences.

The second step of thematic analysis is developing categories, which are descriptive groupings of data that share common aspects, such as different methods for implementing different strategies (Kuckartz, 2019). Developing categories can be a crucial stage in thematic analysis, as categories provide a direct description of the data, while themes are typically derived from a group of related categories (Kuckartz, 2019). This study employed a hybrid method that combines both theory-driven and data-driven approaches for the development of the category framework (Fung & Lo, 2023; Kuckartz, 2019; Xu & Zammit, 2020). On one hand, the self-regulated L2 listening model and questionnaire proposed in this study provided a theoretical basis and direction for constructing the initial category framework focused on self-regulated L2 listening strategies. On the other hand, participants' records in their E-logs involved strategies beyond the theory-based category framework. Therefore, it was necessary to enrich the theory-driven category framework with categories emerging from the data.

Specifically, the researcher first employed the L2 self-regulated listening model and SRLLQ questionnaire (Appendix 2) as a theoretical basis to design the initial category

framework, also known as the codebook. To verify whether the initial coding framework was sufficiently clear and robust to allow for a consistent interpretation across researchers, intercoder reliability (ICR), a “numerical measure of the degree of agreement between coders on how to code the same data” (O’Connor & Joffe, 2020, p. 2), was calculated, using the most common method of dividing the number of agreements by the sum of the number of agreements and disagreements (Miles & Huberman, 1994). The ICR is usually calculated for only a subset of the data, and while there are no universal standards about what proportion of the data in a dataset contributes to a plausible estimate of the ICR, 10%-25% of the units of data are generally considered acceptable depending on the size of the dataset, and samples should be selected as far as possible either by randomisation or by using criteria that have some justification (O’Connor & Joffe, 2020). In the present study, therefore, the researcher and another second language Ph.D. student with expertise in L2 education research and who speaks Chinese as a first language used the initial codebook to conduct the first round of coding separately on the E-logs of ten participants randomly selected from each of the two universities. Comparing the coding of the E-logs by the researcher and the other expert, the ICR result was 95%, meeting the widely accepted standard of 80% (Miles & Huberman, 1994). After the discussion of areas where coding had differed, the researcher revised the initial codebook and used the adapted codebook (Appendix 6), which contained 46 task-level self-regulation strategies and 21 motivation/affect-level self-regulation strategies, to conduct the second round of coding on the remaining E-log data.

3.5.3.2 Cluster Analysis

In addition to thematic analysis and descriptive analysis, the E-log data were also subjected to cluster analysis, which is a “multivariate exploratory procedure used to group cases” (Staples & Biber, 2015, p.243). The current study employed cluster analysis for three reasons. Firstly, learners’ IDLE activity engagement experiences could be characterised as “idiosyncratic and individualised” (Lee & Xie, 2023). Therefore, participants with different L2 listening proficiency, self-regulation, listening self-efficacy, and listening anxiety may have shown differences in the diversity, frequency, duration, and strategy use for their IDLEL activity engagement. Although from an individual perspective, the number of differences in participants’ L2 listening, self-efficacy, and listening anxiety might be limitless, on a macro level, typical patterns representing similarities across individual diversity often exist within a given population (Peng et al., 2022). Identifying these typical patterns and exploring their differences in IDLEL activity engagement can be more feasible and insightful for L2 learners and educators compared to exploring participants’ differences on an individual level. Hence, it was deemed useful to conduct cluster analysis in the present research to identify typical patterns by clustering participants into groups based on their similar characteristics, such as L2 listening scores, self-regulation, self-efficacy, and listening anxiety.

Secondly, in researching learners’ IDLE experiences, it is common to adopt a variable-centered approach, which analyses individual variables through correlation and multiple regression analyses (Lee, 2022). However, the variable-centered approach has its limitations. On the one hand, it may fail to explain the diverse pathways or characteristics of learners who achieve similar learning outcomes (Lee & Xie, 2023). On the other hand, their learning outcomes may be influenced by the interplay and interconnection of multiple factors

rather than by individual variables alone (Peng et al., 2022). Therefore, adopting a person-centered approach that dynamically interrelates different variables and views their interactions as influencing and shaping individual learning patterns seems more advantageous in explaining the relationships between L2 learners' IDDEL experiences and other variables such as language proficiency, self-regulation, anxiety, and self-efficacy. Cluster analysis is a common method used in person-centered research. For instance, Lee and Xie (2023) employed cluster analysis to profile the affective characteristics of 764 Korean secondary EFL learners and 501 Korean university-level EFL learners' IDLE. Using the frequency (e.g., once a week, once a day) of participants' engagement in different types of IDLE activities (i.e., Gaming, Entertainment, English Learning, and Socializing) as cluster variables, five distinct IDLE profiles were identified. Learners with different IDLE profiles showed significant differences in their L2 motivation, enjoyment, anxiety, and grit levels ($p < .001$). Among them, the two groups of EFL learners who made full use of the out-of-school digital environment gained more affective benefits compared to participants from other groups. Similarly, by employing the time spent daily on informal mobile language learning (IMLL) activities related to English reading, listening, speaking, writing, as well as language features such as grammar and vocabulary, as cluster variables, Peng and colleagues (2022) identified six distinct learner types among 238 Chinese university EFL learners. In addition to learners who hardly spent time studying English outside the classroom and those who invested much time in IMLL activities, there were learner types who preferred engaging in comprehension-based IMLL activities (e.g., listening- and reading-related activities).

After examining the performance of these six learner types in terms of their self-reported language proficiency, learning motivation, and anxiety, Peng and colleagues (2022) found that each learner type exhibited unique motivational and emotional characteristics. For instance, two types of participants both spent considerable time on receptive learning activities (albeit to different extents), but they showed significant differences in their motivation, anxiety, and self-reported language proficiency. Conversely, some learners, even though consistent in their linguistic, motivational, and emotional states, differed in their approaches to engaging in IMLL activities. For instance, some learners tended to use more traditional methods to improve their English, while another type of learners showed greater awareness of exploring various learning resources and methods. The authors thus concluded that similar learning patterns resulted from the interplay of different levels of motivation, anxiety, and language proficiency.

The aforementioned two studies that employed cluster analysis adopted a person-centred approach to research IDLE, identifying learner types with similar IDLE experiences from a bottom-up perspective and explaining each pattern in relation to learners' affective, motivational, and linguistic characteristics. Considering the purpose of the present study, participants' pre-test L2 listening scores, self-regulation, self-efficacy, and listening anxiety were used as cluster variables to group them, and the clusters were further interpreted in relation to the characteristics of participants' IDLEL activity engagement (i.e., diversity, frequency, duration, and strategy use diversity).

Thirdly, cluster analysis can provide a basis for personalised instruction and learning for L2 educators and learners (Lee & Xie, 2023; Peng et al., 2022). Specifically, it may help

educators understand different types of learners, allowing them to design personalised teaching content and methods for various learner groups rather than adopting a traditional “one-size-fits-all” approach. Learners can use the results of the cluster analysis to identify which type of learner they are, thus better understanding their learning characteristics and needs, and adjusting their learning strategies accordingly. For this study, the findings might help L2 educators understand the types of learners with different levels of L2 listening proficiency, self-efficacy, and listening anxiety. Thus, they could encourage learners who perform poorly in L2 listening, self-regulation, self-efficacy, and listening anxiety to emulate those who excel in these areas when engaging in IDLEL activities. Similarly, learners could gain insights from the study’s findings on how to engage in IDLEL activities effectively to benefit their L2 listening, self-efficacy, and listening anxiety.

There are two main types of cluster analysis: hierarchical cluster analysis (HCA) and K-means. Neither agglomerative hierarchical clustering, which gradually merges each observation from its own cluster into a larger cluster, nor divisive hierarchical clustering, which gradually splits all observations from one large cluster into smaller subclusters, requires researchers to determine the optimal number of clusters in advance. The number of clusters needs to be decided by the researcher after the data are analysed using methods such as the elbow method, average silhouette method, and gap statistic method (Crowther et al., 2021). K-means, by contrast, requires the researcher to determine the number of clusters in advance, and then the observations are combined into the specified number of clusters to maximize inter-group differences while minimizing intra-group variation (Crowther et al., 2021).

The present study employed agglomerative hierarchical clustering with Ward's method as the linkage method and Euclidean distance as the inter-object distance measure to cluster participants based on their pre-test L2 listening scores, self-regulation, self-efficacy, and listening anxiety. The agglomerative hierarchical clustering, along with Ward's method and Euclidean distance, are commonly used in L2 research (Crowther et al., 2021; Staples & Biber, 2015). Therefore, on the one hand, the effectiveness of these methods has been widely validated. On the other hand, the extensive existing research using these methods can provide insights and references for the cluster analysis in this study. After HCA, to confirm the validity of the grouping, parametric and non-parametric tests were conducted depending on the number of groups and the distribution of the data, to determine whether clusters significantly differed on criterion variables, which were participants' IDLEL engagement (i.e., frequency, diversity, duration, and strategy-use diversity) in the case of the current study. Recommendations for the minimum sample size for cluster analysis vary, ranging from 120 to 500 (Breckenridge, 2000; Sarstedt & Mooi, 2011). Some recommendations suggest that the sample size should be determined based on the number of predictor variables. For instance, if k indicates the number of predictor variables, the minimum sample size should be $2k$ while the ideal sample size is $5*2k$ (Dolnicar, 2002). According to this principle, the sample size used for the cluster analysis in this study was 91, which met the requirement of the preferred sample size of $5 * 2k$ ($k=4$). Both the descriptive analysis and HCA in the current study were performed in R, while the thematic analysis was conducted using Excel.

3.5.4 RQ4: How do learners' IDLEL engagement (i.e., quantity and quality) and SRL predict their English listening comprehension, listening anxiety and listening self-efficacy respectively?

3.5.4.1 Linear Mixed-Effects Models (LMMs)

To identify potential predictors affecting the participants' listening proficiency, listening self-efficacy, and listening anxiety, linear mixed-effects models (LMMs), which can serve as powerful statistical analysis tools for second-language researchers (Cunnings, 2012), were constructed. LMMs were selected because they could not only specify the fixed effects that reflect the influence of independent factors (i.e., main effects) and their interactions on the dependent factor but also take the item random effects and the participant random effects into account (Baayen et al., 2008; Cunnings, 2012; Matuschek et al., 2017).

LMM analysis was conducted in R using the “lmerTest” package (Kuznetsova et al., 2017). When constructing the LMMs for listening, self-efficacy, and listening anxiety, the researcher followed the “make it maximal” principle proposed by Barr et al. (2013). This principle ensures that the model is built based on theory-driven critical hypotheses and experimental design, incorporating the most complex random-effects structure, which includes random intercepts and random slopes for every fixed effect of theoretical interest (Barr et al., 2013). In other words, the maximal model, constructed in a theory-driven manner, includes all theoretically relevant random factors that may contribute to variability in the dependent variable (Zhang, 2022). The purpose of this approach is to minimize model specification bias, ensuring that all potential influences on the dependent variable are adequately accounted for.

After constructing the theory-driven maximal model, the model was run in R to check for convergence (i.e., whether it could be computed). If the model did not converge, it was gradually simplified by removing interactions between random slopes and/or eliminating the least contributing interaction terms and random effects until a final converged model was obtained (Zhang, 2022). If the model successfully converged, the `tab_model` function from the `sjPlot` package (Lüdecke, 2021) was used to calculate marginal R^2 and conditional R^2 to assess model fit. Marginal R^2 measures the contribution of fixed effects, while conditional R^2 accounts for the combined contribution of both fixed and random effects (Plonsky & Ghanbar, 2018). According to Plonsky and Ghanbar (2018), $R^2 < 0.2$ is considered a small effect, while $R^2 > 0.5$ is considered a large effect.

However, even if the maximal model successfully converges, it is not necessarily the optimal choice. Maximal models may bear risks such as overfitting and multicollinearity, leading to a reduction in power, that is, the probability of correctly rejecting false null hypotheses (i.e., detecting actual effects), even though they can “provide the best fit for a given data set” (Matuschek et al., 2017, p. 308). When the model loses power, it may not be able to effectively detect important relationships or effects, even if they exist. Therefore, parsimonious mixed models, which include only the variance components supported by the data, can serve as an alternative approach, and they have been shown to effectively address the issue of loss of power (Matuschek et al., 2017).

Parsimonious mixed models in the current study were achieved through model selection, which aimed to find “the best approximating model” (Buscemi & Plaia, 2020, p. 530) that provided the best explanation of the data within a class of candidate models (Cunnings, 2012).

Model selection was conducted by applying the backward-selection heuristic approach, namely reducing the model complexity by removing insignificant fixed effects one by one (Matuschek et al., 2017). After each removal, the `anova()` function in R was used to compare the new model with the previous model and the maximal model, obtaining key statistical indicators, including p-values and Akaike Information Criterion (AIC) values. AIC reflects the amount of unexplained variation in the model, so a lower AIC value indicates that the model explains a higher proportion of variance (Cunnings, 2012). Additionally, the key criterion for determining whether to continue removing fixed effects was to observe whether the remaining predictors remained significant. The process was stopped only when all remaining predictors were significant, ensuring that the final model was both convergent and interpretable while avoiding overfitting.

The LMMs for self-efficacy, listening anxiety, and listening comprehension were constructed and optimised in the present study to address RQ4. It should be noted that among the 91 participants, only 60 took part in the third listening test, namely the delayed post-test. Therefore, the LMM analysis for listening was conducted under two conditions: with and without the delayed listening post-test. In other words, two separate LMMs were constructed to examine the short-term and long-term effects of IDLEL engagement on participants' listening development respectively.

3.5.5 RQ5: To what extent does learners' level of SRL moderate the relationship between the IDLEL engagement and their L2 listening comprehension, listening anxiety, and listening self-efficacy?

RQ5 aimed to examine whether the effects of IDLEL engagement on listening comprehension, self-efficacy, and listening anxiety varied depending on learners' self-regulation abilities. IDLEL provides a highly autonomous and flexible language learning context, but not all learners may benefit equally from such a context, as individual differences in self-regulation may influence their learning outcomes when engaging in IDLEL.

For instance, self-regulated learners are more likely to actively select suitable listening materials and activities, set listening goals, and employ strategies to enhance comprehension, thereby improving their listening skills more effectively. In contrast, learners with low levels of self-regulation may passively engage with English input without effectively utilising IDLEL resources, leading to lower learning gains. Similarly, self-regulated learners may be better able to set goals, monitor their progress, and adjust strategies when facing challenges, which enables them to have more positive learning experiences in IDLEL, leading to higher listening self-efficacy and better control over listening anxiety. On the other hand, less self-regulated learners may only engage in passive listening practice without managing their learning process, resulting in limited improvements in listening self-efficacy and ineffective anxiety regulation despite IDLEL engagement.

Thus, it is essential to examine whether the impact of IDLEL engagement on listening comprehension, self-efficacy, and listening anxiety depended on learners' SRL levels through moderation effect analysis. If SRL moderated the relationship between IDLEL engagement and listening comprehension, this would suggest that enhancing learners' SRL abilities may improve the effectiveness of IDLEL, highlighting the importance of self-regulation in listening development (Vandergrift & Goh, 2012), particularly in informal L2 listening

contexts. Moreover, if SRL strengthened the positive effect of IDDEL on self-efficacy, this may imply that improving learners' self-regulation may be a crucial pathway for fostering their listening efficacy in IDDEL engagement. Likewise, if SRL moderated the relationship between IDDEL engagement and listening anxiety, this may suggest that developing learners' self-regulation may help them manage their listening anxiety in IDDEL engagement.

3.5.5.1 Moderation Analysis

To answer RQ5, the moderation effects of SRL on the relationship between IDDEL engagement and listening proficiency, listening efficacy and listening anxiety were examined. Moderation effect refers to the influence of a moderator variable (M) on the strength or direction of the relationship between an independent variable (IV) and a dependent variable (DV) (Baron & Kenny, 1986). In other words, the moderator determines whether the effect of the independent variable on the dependent variable changes across different levels of the moderator.

Moderation effects can be tested using LMM analysis. In this study, moderation effects were analysed using the "lmerTest" package in R. Specifically, in LMMs a moderation effect can be represented by the interaction term between the independent variable (i.e., IDDEL engagement) and the moderator (i.e., SRL), denoted as $IDL\text{E}L \times SRL$. If the interaction term is statistically significant ($p < .05$) in the listening model, self-efficacy model, and/or listening anxiety model, this indicates that SRL moderates the relationship between IDDEL engagement and listening, self-efficacy, and/or listening anxiety.

Generally, a moderation effect is meaningful when there is a correlation between the independent variable and the dependent variable (Baron & Kenny, 1986). In the current study,

IDLEL engagement was measured using four indicators: frequency, diversity, and duration of activity engagement, and diversity of strategy use in activities. Therefore, four separate LMMs should ideally have been constructed to examine whether the interaction terms between each IDLEL engagement indicator and SRL were significant, thereby determining SRL's moderation effect. However, the actual number of LMMs was determined based on the findings from RQ4, which explored the predictive effects of IDLEL engagement on listening, self-efficacy, and listening anxiety (see *Section 4.4*). Not all IDLEL engagement indicators were found to significantly predict listening, self-efficacy, or listening anxiety, hence, only the indicators that showed significant predictive power in RQ4 were included in the moderation analysis.

3.6 Ethical Considerations

This study obtained Ethical Approval from the University of Reading before the commencement of the pilot study (Appendix 7), and ethical considerations were prioritised throughout the entire data collection process. First, regarding informed consent (Appendix 7), participants had been informed of the purpose, procedures, risks, and benefits of the research before they decided to take part in the research or not. Second, participants were given the right to withdraw from the research at any point and for any reason. Third, as for the confidentiality issue, apart from basic information about the participants, such as gender, other private data were not collected. Fourth, anonymity was addressed as participants were not be required to use their real names throughout the research. Instead, they used the last four digits of their student ID numbers as pseudonyms when participating in the study (in the target schools, the last four digits of the student ID number can be used to distinguish between

students of the same year in the same faculty). Such a method on the one hand protected the students' privacy as it was impossible for others to identify any of the participants by the four numbers. On the other hand, it allowed the researcher to track participants across different data sources. However, the last four digits of participants' student ID numbers are not presented in this research report. Instead, participants are referred to by an English pseudonym assigned by the researcher in the data analysis and reporting where necessary.

Fifth, the time needed to complete E-logs (Appendix 3) was minimised by using a structured and semi-closed format. It was hoped that the time students needed to devote to the study would be offset by the benefits they gained from completing additional listening practice, alongside potential benefits from reflecting on their listening and self-regulation. Sixth, for both the pilot study and the main study, the results of the listening tests, questionnaires, and E-logs did not affect participants' grades in their formal courses and were not shared with their teachers. Participants were informed of this principle both in the PowerPoint introducing the research and in the Information Sheet.

Additionally, all the electronic research instruments (i.e., questionnaires, Appendix 2; answer sheets for listening tests, Appendix 4; E-logs, Appendix 3) were created on *WenJuanXing*. The listening test data, questionnaire data, and E-log data were also collected through this platform and securely stored in the researcher's password-protected University OneDrive account. *WenJuanXing* is a professional online questionnaire platform that enables the creation of questionnaires as well as the collection and analysis of data. Regarding the security and privacy of the data, *WenJuanXing* has established a comprehensive information security management system and holds corresponding certifications. In addition,

WenJuanXing has passed the Chinese national network security level protection assessment (Level 3). It therefore has the capacity to protect data privacy and security. Related documents can be reviewed here: <https://www.wjx.cn/wjx/license.aspx?type=1>

Moreover, all the electronic research instruments were password protected so that only users who knew the password, namely the researcher, teachers and participants of the research, were able to access to and/or complete them. Additionally, the questionnaire (Appendix 2), answer sheet (Appendix 4), and E-log (Appendix 3) could not be re-edited after submission, ensuring that the responses submitted by participants could not be altered by others. The researcher set the total number of submitted questionnaires, answer sheets, and E-logs to match the number of participants, so that the QR codes automatically expired once all participants' questionnaires had been collected.

Furthermore, the listening tests in the pilot study and the listening tests and questionnaire survey in the main study were conducted during class time. To avoid disrupting the class time of non-participating students, it was agreed, after consultation with the teacher, that any student who did not consent to participate in the study would still complete the listening tests as part of their regular in-class listening practice using paper answer sheets. The teacher did not collect these answer sheets but provided them with transcriptions of the test audio recordings to help them assess their listening proficiency. Additionally, non-participants who wished to gain insight into their self-regulation and perceptions of English listening could complete a paper version of the questionnaire but were not required to submit it. If they were not interested in the questionnaire, the teacher assigned them alternative tasks, such as previewing or reviewing materials, to complete instead.

Taken together, this chapter discussed the methodology of the study in terms of aspects such as research design, participants, data analysis, and considerations of reliability and validity, laying the groundwork for the presentation of research findings that followed.

Chapter 4. RESULTS

4.1 Introduction

This chapter presents the results of the CFA conducted on the SRLLQ (*Self-Regulated L2 Listening Questionnaire*) data as a response to RQ1, which examines whether the structure of the self-regulated L2 listening model is validated. Additionally, the CFA results for the LLAS (*L2 Listening Anxiety Scale*) and LSEQ (*Listening Self-efficacy Questionnaire*) data are also presented. As mentioned in *Section 3.5.2*, the validated SRLLQ, LLAS, and LSEQ served as instruments to measure participants' self-regulation, listening anxiety, and listening self-efficacy. Therefore, presenting their CFA results can help prepare for answering RQ2, which investigates whether the hypothesised joint predictive mechanisms of self-regulation, self-efficacy, and listening anxiety on listening comprehension can be validated. The SEM results testing the two hypothesised joint mechanisms are also reported in this chapter in response to RQ2.

Moreover, the descriptive analysis, thematic analysis, and cluster analysis results of the E-log data are presented in this chapter to address RQ3, which focuses on the characteristics of participants' IDLEL engagement. RQ4 examines the relationships between IDLEL engagement, self-regulation and L2 listening comprehension, self-efficacy, and listening anxiety. To answer these research questions, the LMM analysis results are reported, demonstrating the predictive effects of IDLEL engagement and self-regulation on listening comprehension, self-efficacy, and listening anxiety. Furthermore, the moderating effects of SRL on the relationships between IDLEL engagement and listening comprehension, self-efficacy, and listening anxiety are also presented in this chapter in response to RQ5.

4.2 RQ1: Validation of the Self-Regulated L2 Listening Model

The validation process for the SRLLQ was iterative. Seven rounds of CFA were conducted to evaluate and refine the questionnaire items. Firstly, the CFA conducted on 532 SRLLQ data initially indicated that the five-factor structure consisting of 46 items did not achieve an acceptable model fit ($\chi^2/df = 2.92$; TLI = .737; CFI = .751; RMSEA = .075; SRMR = .077). Following the recommendations of Bostancıoğlu & Handley (2018) and Sun et al. (2021), 21 items with low standardised regression weights and/or high error covariance and standardised residual covariances were removed. After seven rounds of modifications, the revised model demonstrated a strong fit to the data ($\chi^2/df = 2.14$; TLI = .904; CFI = .915; RMSEA = .057; SRMR = .057). Furthermore, all factor loadings met the established cut-off value of .50 (Teng & Zhang, 2016), confirming the model's acceptable effect size (see Figure 5). Additionally, the parameter estimates for the remaining 25 items were statistically significant ($p < .001$). As a result, the five-factor structure of the 25-item SRLLQ (Appendix 2) was validated, supporting the theoretical foundation of the five-phase SRL model on which the questionnaire was based. Additionally, the items within each construct of the model addressed both task-level (i.e., cognitive) regulation and motivational/affective regulation (see Table 5), supporting the validity of the dual-level structure of the self-regulated L2 listening model.

When assessing the convergent validity and internal consistency reliability of the constructs, according to Table 6, although the Average Variance Extracted (AVE) for the overall five-factor model and for individual constructs, except for Construct One and Five, fell below the conventional threshold of .50 (Hair & Alamer, 2022), their composite

reliability (CR) values were all well above the .70 benchmark (Hair et al., 2017). This indicated that the observed variables still demonstrated a high degree of shared variance in explaining their respective latent constructs. Therefore, the convergent validity of these constructs could still be considered acceptable (Fornell & Larcker, 1981). Additionally, the CR and Cronbach's α values for all constructs ranged from .76 to .89, indicating a high level of internal consistency reliability. Moreover, the 25-item SRLLQ demonstrated good reliability, as the Cronbach's α value for the overall SRL model reached .89.

In summary, the five-phase, dual-level structure of the self-regulated L2 listening model was validated, thereby addressing RQ1. The subsequent section will present the results for RQ2, that is, the verification outcomes of two hypothesised joint predictive mechanisms of self-regulation, self-efficacy, and listening anxiety on listening.

Figure 5

CFA Results for the Five-Construct Self-Regulated L2 Listening Model

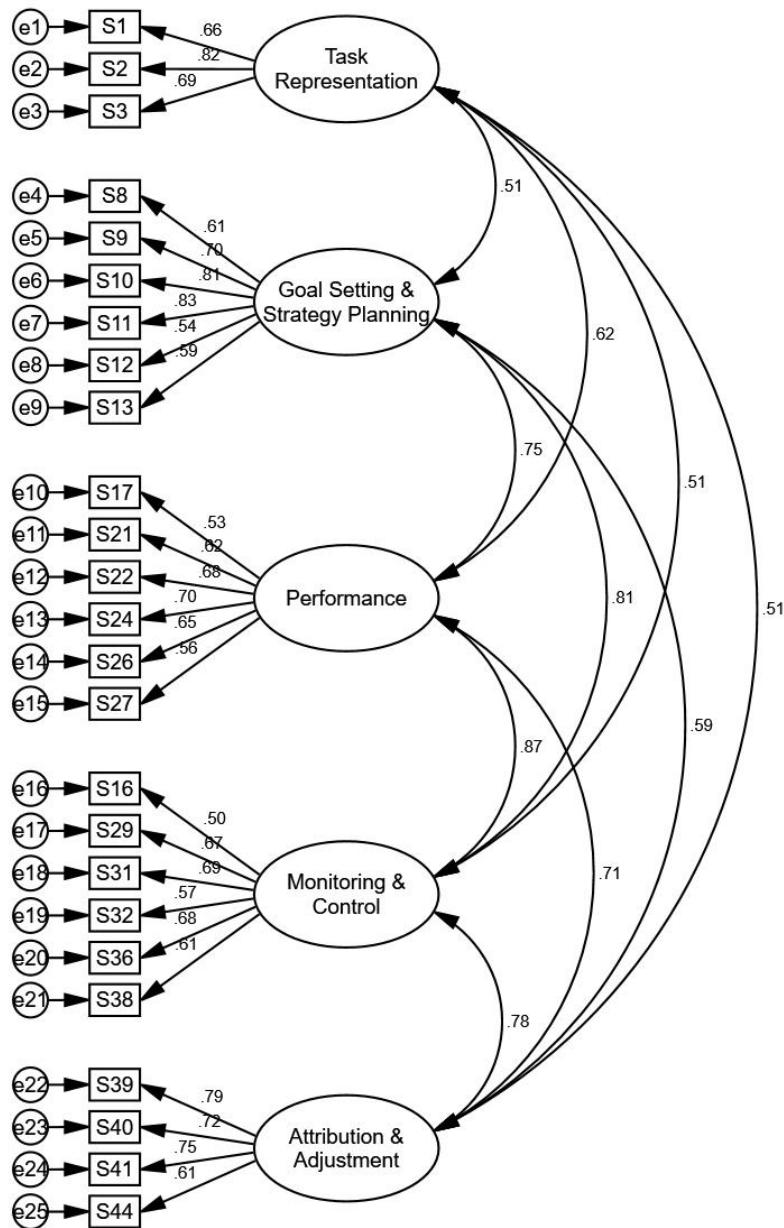


Table 6*The Items, Convergent Validity and Internal Consistency Reliability of SRL Constructs*

SRL Constructs	Task-Related Items	Motivation/Affect-Related Items	AVE	Cronbach's α	CR
Construct 1: Task Representation	2	1, 3	.53	.76	.77
Construct 2: Goal Setting & Strategy Planning	8, 9, 10	11, 12, 13	.47	.83	.85
Construct 3: Performance	17, 21, 22, 24	26, 27	.46	.79	.80
Construct 4: Monitoring & Control	16, 29, 31, 38	32, 36	.46	.79	.79
Construct 5: Attribution & Adjustment	39, 41, 44	40	.53	.81	.82
Five Constructs Together: SRL Model	15 items	10 items	.46	.89	.80

Note. AVE = Average Variance Extracted; CR = Composite Reliability

4.3 RQ2: Validation of the Hypothesised Joint Predictive Mechanisms

4.3.1 CFA for Questionnaire Validation: LSEQ and LLAS

As mentioned in *Sections 3.6.2* and *4.1*, the SRLLQ, LLAS, and LSEQ scales all required validation before being used to measure participants' self-regulation, listening anxiety, and listening self-efficacy. Since the researcher had prior expectations regarding the number of factors and the structural composition of both LLAS and LSEQ (as detailed in *Sections 3.3.1.2* and *3.3.1.3*) and sought to validate these expectations, CFA was employed for validating these two scales, similar to the validation process of the SRLLQ.

The initial CFA result of the two-factor construct generated from the 435 LSEQ data failed to demonstrate an acceptable goodness-of-fit ($\chi^2/df = 10.62$; TLI = .666; CFI = .688; RMSEA = .165; SRMR = .076). After six rounds of CFA, twenty-four items were thus removed to improve the model fit following the same item screening criteria as the SRLLQ, and the adapted model fit of the two-construct self-efficacy model was good ($\chi^2/df = 3.67$; TLI = .951; CFI = .964; RMSEA = .078; SRMR = .030). Similarly, after removing five items through three rounds of CFA, the adapted three-factor listening anxiety model ($N = 427$) demonstrated a good model fit ($\chi^2/df = 2.33$; TLI = .956; CFI = .966; RMSEA = .063; SRMR = .051). All factor loadings of the two models ranged from .68 to .91, exceeding the cut-off value of .50 (Teng & Zhang, 2016), demonstrating a strong effect size (Figure 6 and Figure 7). Additionally, the parameter estimates for all items within the two models were all significant ($p < .001$). Finally, both the two constructs of the self-efficacy model and the three constructs of the listening anxiety model were validated, and the 9-item LSEQ and the 12-item LLAS were hence validated, with good (Cronbach's $\alpha = .80$) and acceptable (Cronbach's $\alpha = .73$) reliability, respectively.

The items included, as well as the convergent validity and internal consistency reliability of the two-construct self-efficacy model and the three-construct listening anxiety model, are presented in Table 7. The indicators used to measure the convergent validity (i.e., AVE) and the internal consistency reliability (i.e., Cronbach's α and CR) of the self-efficacy model and its included constructs all exceeded their respective acceptable thresholds (i.e., .50, .70) (DeVellis, 2012; Shrestha, 2021). The same was true for the three-construct listening anxiety model. Taken together, the overall two-construct self-efficacy model and the three-construct

listening anxiety model, as well as all the constructs they contain, exhibited adequate convergent validity and internal consistency reliability.

Figure 6
CFA Results for the Listening Self-Efficacy Model

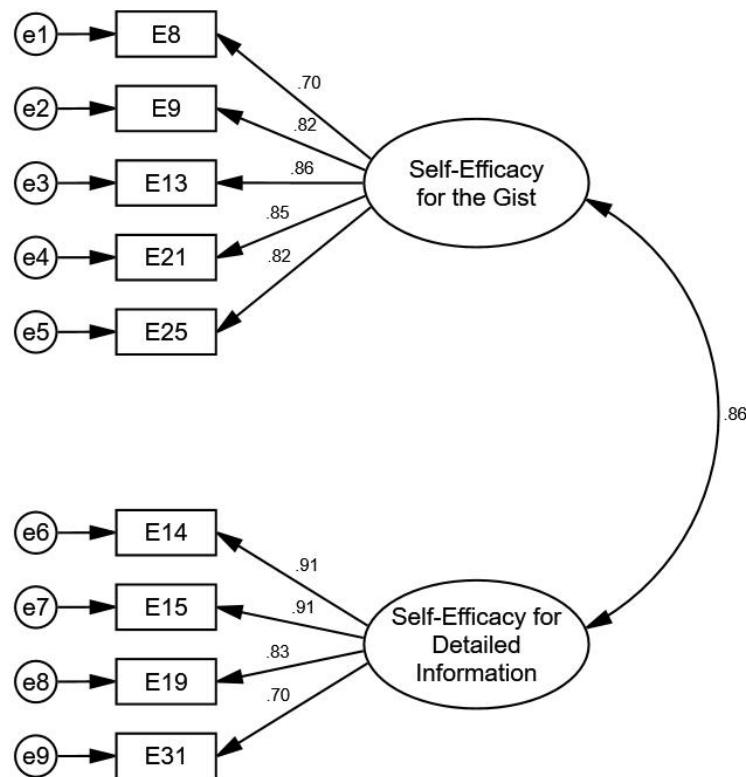


Figure 7
CFA Results for the Listening Anxiety Model

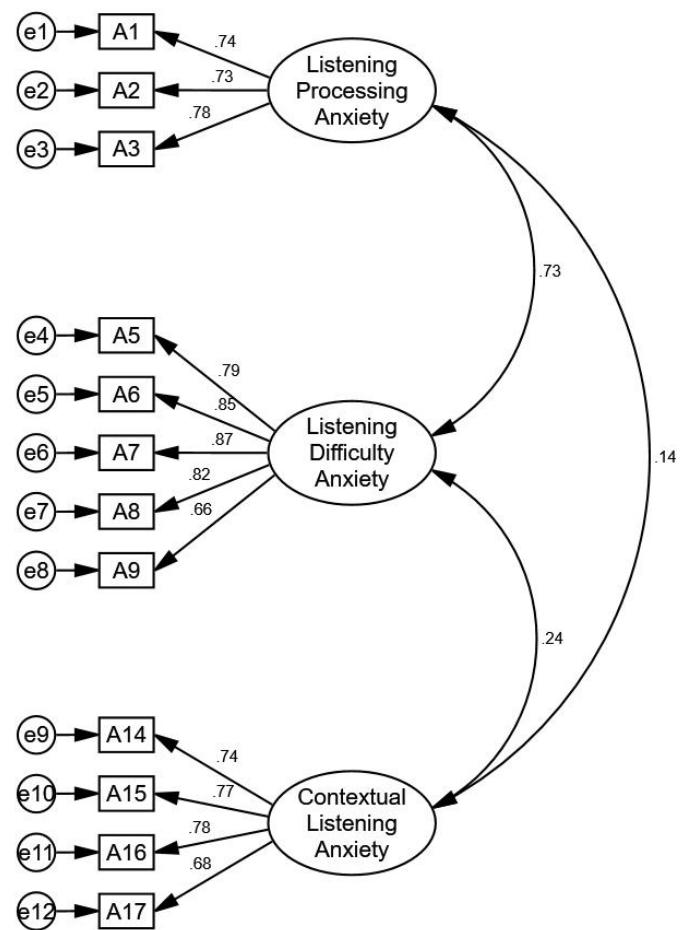


Table 7

The Items, Convergent Validity and Internal Consistency Reliability of Self-Efficacy and Listening Anxiety Constructs

Self-Efficacy Constructs	Items	AVE	Cronbach's α	CR
Construct 1: Self-Efficacy for Gist	8, 9, 13, 21, 25	.66	.90	.90
Construct 2: Self-Efficacy for Detailed Information	14, 15, 19, 31	.70	.90	.90
Two Constructs Together: Self-Efficacy Model	9 items	.67	.80	.79
<hr/>				
Listening Anxiety Constructs	Items	AVE	Cronbach's α	CR
Construct 1: Listening Processing Anxiety	1, 2, 3	.57	.80	.80
Construct 2: Listening Difficulty Anxiety	5, 6, 7, 8, 9	.64	.90	.90
Construct 3: Contextual Listening Anxiety	14, 15, 16, 17	.56	.82	.83
Three Constructs Together: Listening Anxiety Model	12 items	.56	.73	.79

Note. AVE = Average Variance Extracted; CR = Composite Reliability

4.3.2 SEM for the Validation of the Hypothesised Joint Predictive Mechanisms

4.3.2.1 CFA for SEM Purpose

As mentioned in the *Section 3.5.2.1*, CFA was conducted prior to SEM to assess the fit between the new data ($N = 130$) and the previously validated models of SRL, self-efficacy, and listening anxiety. The results indicated a strong model fit for the SRL model ($\chi^2/df = 1.36$; TLI = .901; CFI = .913; RMSEA = .063; SRMR = .075), the self-efficacy model ($\chi^2/df = 1.12$;

TLI = .990; CFI = .993; RMSEA = .041; SRMR = .047), and the listening anxiety model ($\chi^2/df = 1.47$; TLI = .946; CFI = .959; RMSEA = .070; SRMR = .065). These findings also suggest that SRLLQ, LSEQ, and LLAS performed well with the new data ($N = 130$), demonstrating the robustness of the measurement model structure and providing a solid foundation for the implementation of SEM.

4.3.2.2 Basic Statistics

Apart from the normality issue discussed in the *Section 3.5.2.1*, no violations were detected concerning linearity, homogeneity of variance, multicollinearity, or the presence of outliers in the dataset. The descriptive statistics and Spearman's correlation coefficients for the variables indicated that all variables were significantly related to one another (Table 8). Specifically, listening anxiety showed significant negative correlations with the other three variables, namely self-efficacy, self-regulation, and listening test scores, while all remaining variables demonstrated significant positive correlations with each other.

Table 8

Descriptive Statistics and Spearman's Correlation Coefficients (N= 130)

Variables	Mean (SD)	Listening Anxiety	Listening Self-Efficacy	Listening Self-Regulation
Listening Anxiety	3.74 (.85)			
Self-Efficacy	54.25 (15.53)	-.34***		
Self-Regulation	4.36 (.73)	-.21*	.31***	
L2 Listening Test Scores	11.53 (4.76)	-.29***	.30***	.22*

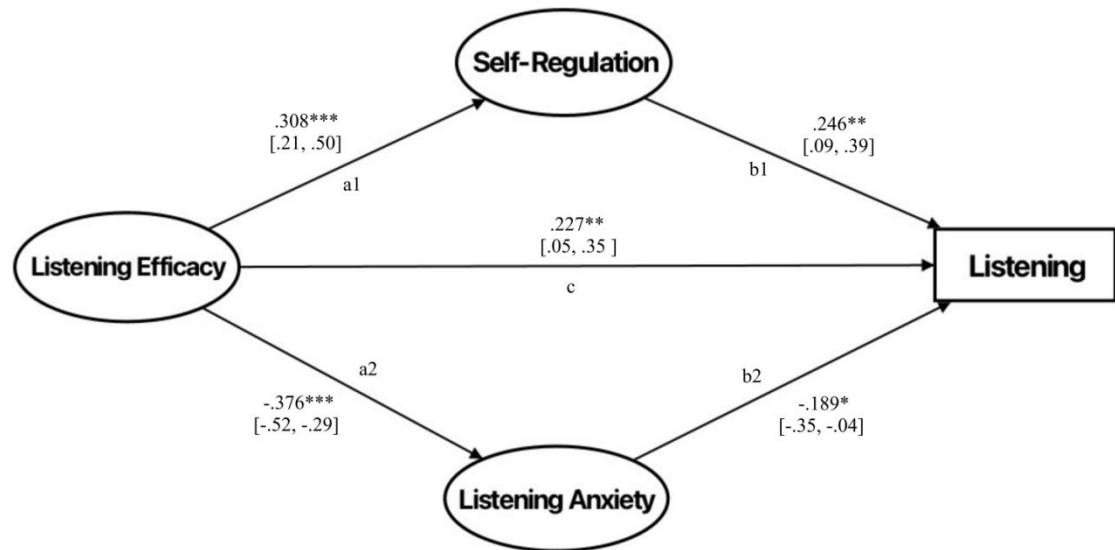
Note. SD = standard deviation; * $p \leq .05$; *** $p \leq .001$.

4.3.2.3 Indirect Effect of Self-Efficacy on Listening Through SRL and Listening Anxiety

The first hypothesised model examined the mediating role of SRL and listening anxiety in the relationship between self-efficacy and listening. Following the method for calculating mediation effects outlined in the *Section 3.5.2.2*, the results confirmed the mediation effects of SRL and listening anxiety, identifying them as complementary mediation (Zhao et al., 2010). This classification was based on the fact that both the direct effect of self-efficacy on listening (Figure 8) and its indirect effects through SRL and listening anxiety (Table 9) were significant and in the same direction.

Figure 8

The Hypothesised Model Showing the Indirect Effect of Self-Efficacy on Listening through SRL and Listening Anxiety



Note. Values in brackets are confidence interval (CI) 95%

Table 9*Indirect Effects of the First Hypothesised Model*

Indirect Effect	β	95% CI		p	Value & Sign
		Lower	Upper		
a1 * b1: efficacy>SRL> listening	.076	.031	.153	.033*	
a2 * b2: efficacy>anxiety> listening	.071	.016	.154	.047*	
a1 * b1 * c					.017 (+)
a2 * b2 * c					.016 (+)

Note. β = Standardized Coefficients Beta; * $p \leq .05$

After confirming the significance of the hypothesised paths, the coefficient of determination (R^2), which represents in-sample predictive power by indicating “the variance in the outcome(s) explained by the predictor constructs” (Hair & Alamer, 2022, p.8), was further assessed. The results demonstrated that self-efficacy accounted for 14.1% of the variance in L2 listening anxiety and 9.5% of the variance in SRL, reflecting a modest level of explanatory power (11% to 30%) for listening anxiety and a weak level (0% to 10%) for SRL (Hair & Alamer, 2022). This suggests that self-efficacy exerted a stronger predictive effect on listening anxiety than on the other mediator, SRL. Regarding listening anxiety, self-efficacy exhibited a significant negative predictive effect but was not the primary predictor, while for SRL, self-efficacy was a significant but not powerful predictor. Additionally, listening anxiety, SRL, and self-efficacy jointly explained 23.1% of the variance in L2 listening, indicating a modest level of explanatory power (Hair & Alamer, 2022).

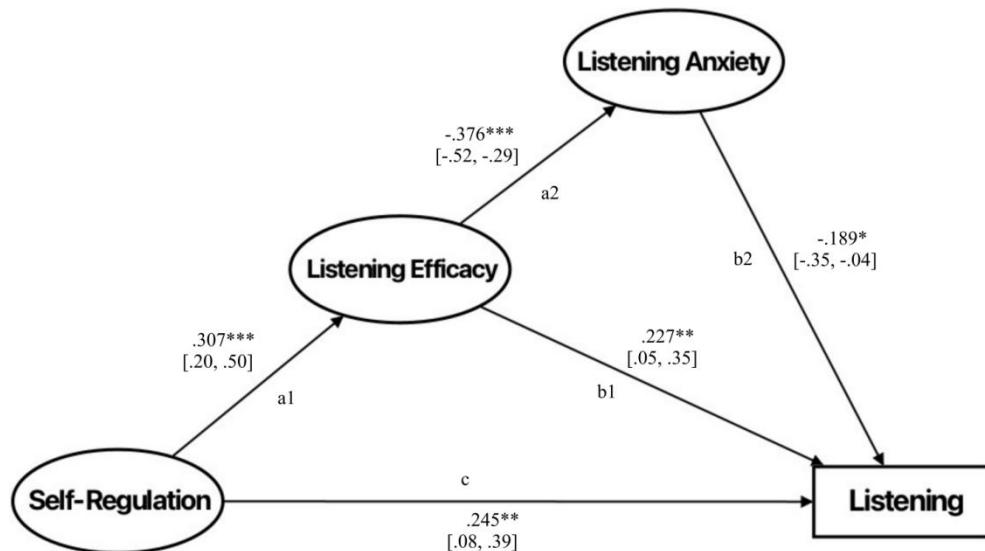
Furthermore, the path coefficient (β), which serves as a key indicator for evaluating the impact of predictor variables on outcome variables, is widely recognised as an essential measure of a model's predictive power (Lowry & Gaskin, 2014). As illustrated in Figure 6, SRL exhibited the strongest predictive power for L2 listening in the first hypothesised model, followed by self-efficacy and listening anxiety. In summary, the first hypothesised joint predictive mechanism for listening comprehension was successfully validated.

4.3.2.4 Indirect Effect of SRL on Listening Through Self-Efficacy and Listening Anxiety

The second hypothesised model proposed an indirect predictive effect of SRL on listening comprehension, mediated by self-efficacy and listening anxiety. Similar to the findings of the first hypothesised model, the mediation effects in this model were also identified as complementary mediation (Zhao et al., 2010). As illustrated in Figure 9 and Table 10, the direct effect of SRL on listening ($c, p = .009$), as well as its indirect effect via self-efficacy ($a1 * b1$), were both significant and in the same direction. Additionally, the direct effect of self-efficacy on listening ($b1, p = .007$) and its indirect effect via listening anxiety ($a2 * b2$) were also found to be significant and consistent in direction.

Figure 9

The Hypothesised Model Showing the Indirect Effect of SRL on Listening through Self-Efficacy and Listening Anxiety



Note. Values in brackets are confidence interval (CI) 95%

Table 10

Indirect Effects of the Second Hypothesised Model

Indirect Effect	β	95% CI		p	Value & Sign
		Lower	Upper		
a1 * b1: SRL>efficacy>listening	.070	.015	.133	.032*	
a2 * b2: efficacy>anxiety>listening	.071	.016	.154	.048*	
a1 * b1 * c					.017 (+)
a2 * b2 * b1					.016 (+)

*Note. β = Standardized Coefficients Beta. * $p \leq .05$*

Similar to the first hypothesised model, listening self-efficacy was found to have a moderate negative predictive effect on listening anxiety (14.1%), indicating that self-efficacy was a significant negative predictor of listening anxiety but not the primary predictor. L2 listening was moderately predicted (23%) by the remaining three variables, with SRL being the strongest positive predictor, followed by the positive predictor self-efficacy, and the negative predictor listening anxiety. Additionally, SRL was identified as a significant but not strong predictor of self-efficacy, as it demonstrated weak explanatory power (9.4%) (Hair & Alamer, 2022).

Taken together, the two hypothesised mechanisms were both validated, and the RQ2 was addressed. The following section will display the findings for RQ3, which explores the characteristics of participants' IDLEL engagement (quantity and quality)

4.4 RQ3: The Characteristics of Participants' IDLEL Engagement

4.4.1 Diversity, Frequency and Duration: Descriptive Analysis

The frequency, diversity, and duration of participants' ($N = 91$) engagement in IDLEL activities are presented in Table 11. Participants engaged in a total of ten types of IDLEL activities during the four-week IDELL study, involving English movies, English TV series, English songs, listening practice for English tests, English learning mobile applications, English talk-shows, English presentations, English radio programs, English videos on social platforms, and English audiobooks. Among 91 participants, the majority ($N = 64$, 70.3%) spent less than one hour per day on IDLEL activities, followed by those who spent one to two hours ($N = 23$, 25.3%) and two to three hours ($N = 3$, 3.3%) per day. The fewest participants,

with only one person (1.1%), spent an average of three to four hours per day engaging in IDLEL activities.

Additionally, as shown in Table 12, the activity students participated in most frequently was listening to English songs, and it was also the activity chosen by the most participants. The activity students participated in the least often was listening to English audiobooks, and it was also the activity chosen by the fewest participants.

Table 11*Frequency, Diversity, and Duration of Participants' IDLEL Engagement*

N= 91		Frequency	Diversity	Duration
Four-Week Engagement	M(SD)	10.07 (4.33)	4.63 (1.43)	23.10 (13.39)
Weekly Engagement	M(SD)	2.85 (1.04)	1.41 (.71)	6.62 (3.86)

*Note. M = Mean; SD = standard deviation***Table 12***Frequency of Engagement and Number of Participants Choosing the Activity*

N= 91	English movies	English TV series	English songs	Listening practice for English tests	English learning mobile applications	English talk-shows	English presentations (e.g., TED Talks)	English radio programs	English videos on social platforms	English audiobooks
Frequency of engagement	101	137	158	120	131	41	68	32	108	11
Number of participants choosing this activity	48	58	67	42	57	25	38	16	55	8

The E-logs also required participants to report their reasons for engaging in various IDLEL activities. In addition to the reasons listed in the E-logs such as entertainment, improving English skills, and recommendations from friends or teachers, some participants also mentioned that they participated in IDLEL activities for test preparation. Table 13 shows that the most important stated reason for engaging with movies, TV series, talk-shows, videos on social media platforms, and English songs was for entertainment, while the main reason for engaging in activities such as radio programmes, audiobooks, and test-related listening practices was the teacher's recommendations. Additionally, the primary reason participants used English learning mobile apps and watched English presentations was to improve their English skills. Specifically, participants reported that the English skill they most expected to improve through these two activities was listening, followed by speaking, as shown in Table 14, which shows the English skills participants expected to improve through engagement in IDLEL activities. In addition to listening, participants also aimed to enhance their English speaking, reading, writing, vocabulary, and grammar through their engagement in IDLEL activities.

Furthermore, participants also reported that they gained different benefits from engaging in various types of IDLEL activities. Table 15 shows that the primary benefit reported by participants from engaging in IDLEL activities such as watching English movies, TV shows, talk shows and presentations, listening to songs and audiobooks, using English learning apps, and viewing English videos on social platforms, was a reduction in English listening anxiety, followed by an expansion of their vocabulary. For those who did test-related listening practice and listened to English radio programs, the greatest perceived benefit was the expansion of

their vocabulary, followed by a reduction in English listening anxiety and an increase in confidence, respectively.

Table 13*Frequency of IDLEL Activity Engagement for Different Reasons*

N= 91	Entertainment	To improve specific English skills	Friends' recommendation	Teachers' recommendation	Test preparation
English movies	101	71	77	81	1
English TV series	121	74	84	91	
English talk shows	38	27	31	29	
English videos on social platforms	69	44	44	59	
English learning mobile applications	103	104	62	94	1
English presentations	65	83	56	69	1
English radio programs	29	31	21	35	1
English audiobooks	9	11	6	12	
English songs	158	102	98	125	2
Test-related listening practice	70	98	72	99	4

Table 14*English Skills Participants Expected to Improve Through IDEL Activity Engagement*

N= 91	Listening	Speaking	Reading	Writing	Vocabulary	Grammar
English movies	25	13		1	7	1
English TV series	30	18	5	4	5	2
English talk shows	3	2			4	
English videos on social platforms	27	15	7	4	6	1
English learning mobile applications	80	35	11	6	14	1
English presentations	20	18	2	3	6	2
English radio programs	16	2	1		3	1
English audiobooks	3		1	1	1	1
English songs	29	20	1	3	9	1
Test-related listening practice	71	17	13	13	23	7

Table 15*Anticipated benefits from IDEL Activity Engagement*

<i>N= 91</i>	Vocabulary expanded	Became familiar with natural & authentic accent	Became familiar with natural & authentic speech speed	Learned about target culture	Understood long and complex sentences	Confidence in English listening increased	Anxiety in English listening decreased
English movies	107	46	51	40	16	50	133
English TV series	104	59	40	49	34	71	116
English songs	81	68	58	45	34	54	82
Listening practice for English tests	40	19	17	20	33	28	34
English learning mobile applications	95	53	49	41	29	58	103
English talk-shows	80	47	39	44	33	71	122
English presentations	77	35	31	24	21	27	82
English radio programs	77	56	40	39	37	72	69
English videos on social platforms	107	70	62	50	29	85	109
English audiobooks	77	54	58	51	29	67	91

4.4.2 Strategy Use: Thematic Analysis and Descriptive Analysis

A thematic analysis of the participants' E-log recorded mental, psychological, and physical "actions" before, during, and after IDLEL activities identified a total of 54 self-regulation strategies. These included 39 task-level self-regulation strategies and 15 motivation and affect-level self-regulation strategies. The definitions of all strategies can be found in Appendix 6.

4.4.2.1 Listening Task-level Strategy Use

The 39 task-level self-regulated L2 listening strategies could further be grouped into eight categories: task perception, goal setting, listening strategy planning, strategy employment, listening monitoring, reflection, attribution, and adjustment (Table 16). The frequency of use and the number of users for each strategy are presented in Table 17.

Listening material perception (when participants receive a listening task, they try to figure out the nature of the listening material, e.g. its genre, theme, cultural context, difficulty, etc.) was the most frequently used strategy reported by participants to build an understanding of the listening tasks, and *predicting listening content* (before listening, participants predict what they are going to listen to) was the most commonly used goal-setting strategy by participants before listening. Additionally, *strategy planning* (before listening, participants plan which listening strategies to use or not to use) was another frequently used strategy reported by participants at the pre-listening stage. During the listening task, the most widely chosen strategy to facilitate participants' listening comprehension was *using the dictionary or electronic translator*, followed by *replaying the listening material* and *using subtitles or scripts*. Moreover, few participants claimed that they monitored their listening comprehension

and the difficulties they encountered during listening. As for reflection on listening, the least frequently used strategy was *reflecting on strategy use* (participants reflect on the effectiveness of the strategies used in listening), while the most commonly used strategies reported by participants were *accumulating* (participants record and/or memorise unfamiliar or “useful” words, sentence structures, expressions, cultural phenomena or even other skills) and *reviewing the new knowledge* they acquired during the task. Very few participants claimed that they made *attributions* for their listening outcomes. However, some participants reported that they *planned for their future study* (participants set goals/ plans after the task for future. E.g., “Improve the frequency of speaking English in my daily life”, “Apply to real-world communication”).

Table 16*Task-Level Self-Regulation Strategies*

Task Perception	Goal Setting	Listening Strategy Planning	Strategy Employment	Listening Monitoring	Reflection	Attribution	Adjustment
Task demand perception	Setting goals for task completion		Selective attention	Imagination	Comprehension monitoring	Task performance reflection	Internal-factor attribution
Listening material perception	Setting goals for comprehension levels		Focus on ongoing contents	Noting		Strategy-use reflection	Help seeking
	Setting goals for selected attention	Listening strategy planning	Contextual-clue assistance	Repeat		Knowledge accumulation	
Previous knowledge recall	Setting goals for skill enhancement		Common-sense assistance	Subtitles/ Scripts	Difficulty-based monitoring		External-factor attribution
	Content prediction		Linguistic-knowledge assistance	Dictionary/ Translation websites/ apps		Task-gained knowledge review	Sharing

Topic-knowledge assistance	Pause/ Speed Adjustment
Cultural background assistance	Imitation
Understanding-based assistance	
Mental Translation	Question-clue assistance
Visualisation	Skip

Table 17*Frequency of Strategy Use and the Number of Participants Using Strategies*

N= 91	Task demand perception	Listening material perception	Previous knowledge recall	Setting goals for task completion	Setting goals for comprehension levels	Setting goals for selected attention	Setting goals for skill enhancement
Frequency of use	6	20	8	2	23	45	24
Number of participants using this strategy	4	15	4	2	10	22	15
N= 91	Content prediction	Listening strategy planning	Selective attention	Focus on ongoing contents	Contextual-clue assistance	Common-sense assistance	Linguistic-knowledge assistance
Frequency of use	278	158	64	1	78	4	36
Number of participants using this strategy	75	49	34	1	35	1	18

N= 91	Topic-knowledge assistance	Cultural background assistance	Understanding-based assistance	Mental Translation	Visualisation	Imagination	Noting
Frequency of use	1	1	15	2	7	2	53
Number of participants using this strategy	1	1	11	2	1	2	23
N= 91	Repeat	Subtitles/ Scripts	Dictionary/ Translation websites/ apps	Pause/ Speed Adjustment	Imitation	Question-clue assistance	Skip
Frequency of use	123	83	140	13	41	13	2
Number of participants using this strategy	43	33	49	8	23	9	2

N=91	Comprehension monitoring	Obstacle monitoring	Task performance reflection	Strategy-use reflection	Knowledge accumulation	Task-gained knowledge review	Internal-factor attribution
Frequency of use	4	2	46	2	477	356	2
Number of participants using this strategy	3	2	16	2	85	81	1
<hr/>							
N=91	External-factor attribution	Plan for future study	Help seeking	Sharing			
Frequency of use	1	83	2	5			
Number of participants using this strategy	1	30	1	5			

4.4.2.2 Motivation and Affect-level Strategies

Fifteen motivation and affect-level self-regulated L2 listening strategies could further be grouped into four categories: motivation and affect awareness, goal clarification, motivation and affect-level strategy employment, motivation and affect adjustment (Table 18). The frequency of use and the number of users for each strategy are presented in Table 19.

Before listening, participants claimed that they noted and dealt with their listening anxiety and self-efficacy. Aside from some unspecified strategies used to alleviate anxiety and boost self-efficacy (e.g., “adjusting mindset”, “relaxing”), the most frequently used and most commonly chosen strategy reported by participants before listening was *self-encouragement or self-comfort* (e.g., “I will tell myself that it doesn’t matter if I don’t know”; “I would tell myself that I can do a good job, and keep a positive attitude to this activity”). The second most frequently used strategy was *mental preparation for potential difficulties* (e.g., “Imagining what to do if I encounter something I don’t understand during listening”). Additionally, they *clarified their goals* for IDLEL activity engagement before listening (e.g., “I listen to English songs only for relaxation”; “I listen to English listening materials in English learning apps to learn knowledge”).

Participants reported some unspecified strategies used during listening to alleviate anxiety or enhance self-efficacy (e.g., “stay calm”; “relax my mood in listening”), but no participants listed specific strategies. Very few participants claimed that after the listening activities adjustments to motivation and affect occurred (e.g., “My confidence was enhanced, and my anxiety was relieved after activity”).

Table 18*Motivation and Affect-Level Self-Regulation Strategies*

Motivation & Affect Awareness	Goal clarification	Motivation and affect-level strategy employment	Motivation and affect adjustment
Self-efficacy awareness		Unspecified anxiety-relieving and efficacy-boosting strategies before listening	
Anxiety awareness		Anxiety-relieving physical strategies	
		Gather mind and focus	
Task-value awareness		Imaging good results	
	Goal clarification	Recall previous experiences	Motivation and affect adjustment
		Self-push/encourage/ comfort	
		Mental preparation for potential difficulties	
Potential difficulty awareness		Selecting/ Creating an environment conducive to concentration	
		Anxiety-relieving and efficacy-boosting strategies that were not explicitly stated during listening	

Table 19*Frequency of Strategy Use and the Number of Participants Using Strategies*

N= 91	Self-efficacy awareness	Anxiety awareness	Task-value awareness	Potential difficulty awareness	Goal orientation	Unspecified anxiety-relieving or efficacy-boosting strategies before listening	Anxiety-relieving physical strategies	Imaging good results
Frequency of use								
	14	137	22	8	31	145	17	1
Number of participants using this strategy								
	6	50	10	7	22	43	12	1
N= 91	Recall previous experiences	Gather mind and focus	Self-push/encourage/comfort	Mental preparation for potential difficulties	Selecting/Creating an environment conducive to concentration	Unspecified anxiety-relieving or efficacy-boosting strategies during listening	Motivation and affect adjustment	
	Frequency of use							
	1	22	42	26	8	24	5	

Number of participants using this strategy	1	14	20	9	6	13	4
---	---	----	----	---	---	----	---

As mentioned in *Section 2.4.4.2*, the diversity of learning strategies can, to some extent, reflect learners' adaptability to different task demands and learning environments, thereby helping researchers assess their self-regulation in IDLEL activities. Furthermore, examining the diversity of strategy use can assist researchers in determining whether learners have achieved a balance in self-regulation across cognitive, motivational, and affective dimensions. Therefore, the diversity of strategy use was included as an indicator for measuring participants' IDLEL engagement. Descriptive statistics showed that participants used an average of 10 to 11 self-regulated L2 listening strategies ($SD = 3.67$) over the four-week study period, with an average of 3-4 strategies used per week ($SD = 1.17$). Additionally, participants employed a maximum of nine different strategies and a minimum of one strategy during the four-week IDLEL study.

4.4.3 Cluster Analysis

To examine whether participants ($N=91$) with varying L2 listening proficiency, self-regulation, self-efficacy, and listening anxiety differed in their IDLEL engagement, their pre-test (T1) scores for L2 listening, self-regulation, listening self-efficacy, and listening anxiety were used to decide cluster membership, while the IDLEL variables, including frequency, diversity, duration, and strategy-use diversity, were used as criterion variables, respectively.

Since the scales of the variables used for cluster analyses in this research were inconsistent, it was necessary to transform them into standardised variables before conducting cluster analyses to prevent one or more of them from having a greater influence in distance calculations and thus affecting the clustering results (Crowther et al., 2021). This study used Z

scores, which is a commonly used standardisation method in cluster analysis, to achieve this purpose (Staples & Biber, 2015).

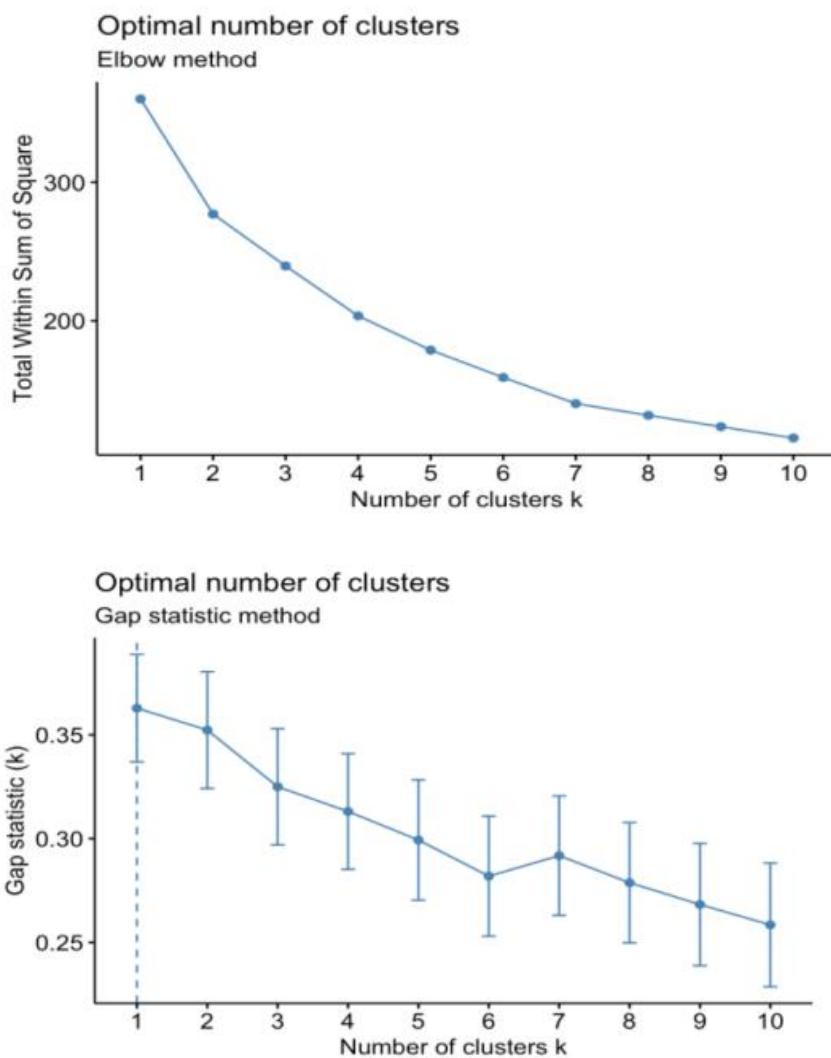
When determining the optimal number of clusters, the elbow method, average silhouette method, and gap statistic method were used (Crowther et al., 2021). The elbow method estimates the optimal number of clusters based on the total variance within the clusters (Crowther et al., 2021). Specifically, a lower variance indicates a more uniform cluster. The point where the variance starts to sharply decrease, known as the “elbow”, is usually considered the optimal number of clusters, which in this case is three. The average silhouette method finds the optimal number of clusters by calculating and comparing the average silhouette coefficients, which measure the distance between each data point and its own cluster as well as the nearest neighbouring cluster, for different numbers of clusters (Dinh & Huynh, 2019). A higher average silhouette coefficient indicates that data points are more tightly grouped within their own cluster and more distinctly separated from other clusters (Crowther et al., 2021). Therefore, the number of clusters that maximises this index can be considered the optimal number of clusters. The gap statistic method determines the optimal cluster solution by comparing the within-cluster dispersion of the actual dataset with that of random reference datasets for different numbers of clusters, selecting the number of clusters with the largest gap statistic (Tibshirani et al., 2001).

As shown in Figure 10, all three methods for estimating the optimal clustering scheme recommended two as the optimal number of clusters. Although overlaps between clusters seem inevitable in L2 research due to variable correlations (Crowther et al., 2021), the high degree of overlap between Cluster One and Cluster Two, as shown in the scatterplot of the

two-cluster solution (Figure 11), may indicate imperfect differentiation. This suggests that the two-cluster solution may not be optimal. Therefore, increasing the number of clusters could be considered to improve differentiation between clusters, thereby better reflecting the characteristics and structures of the original data (Kingrani et al., 2018; Milligan et al., 1983).

As a result, the three-cluster solution, which was closest to the recommended two-cluster solution, was ultimately chosen.

Figure 10
Indices Estimating the Optimal Number of Clusters



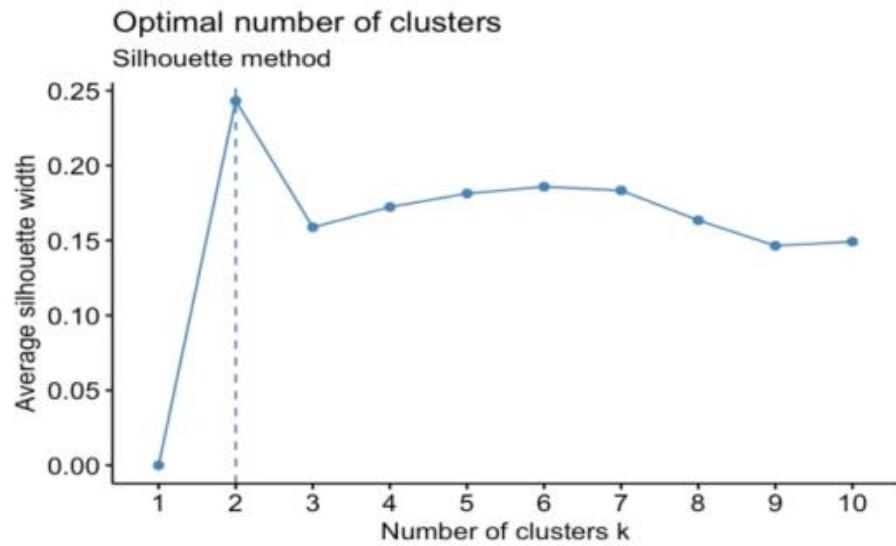
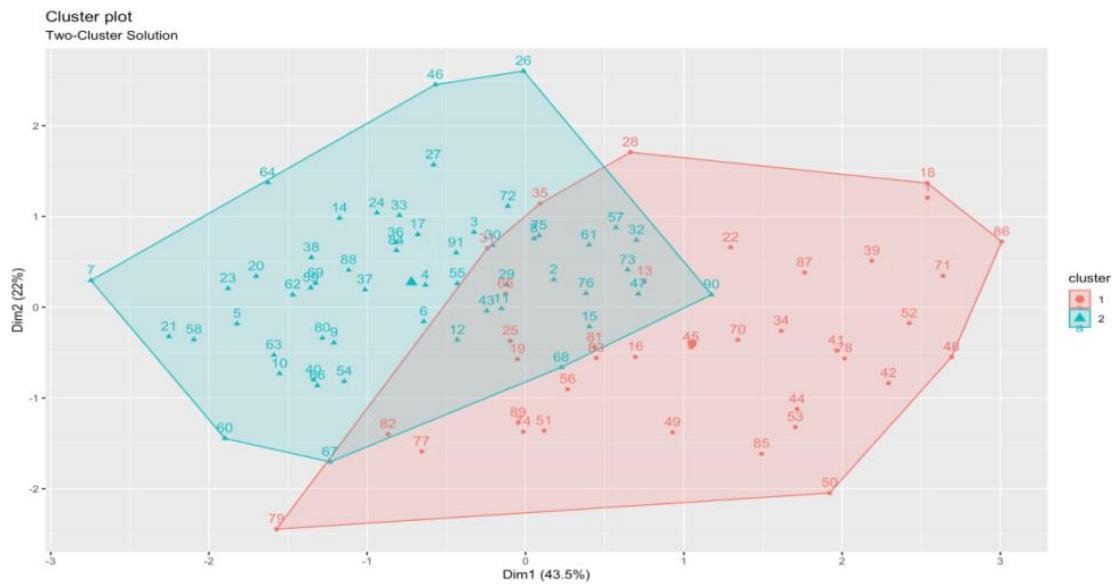


Figure 11
Scatterplot for the Two-Cluster Solution



As stated in *Section 3.5.3.2*, to confirm the validity of the grouping, determining whether clusters significantly differed on criterion variables, which were participants' IDLEL engagement (i.e., frequency, diversity, duration, and strategy-use diversity), parametric or non-parametric tests needed to be conducted depending on the number of groups and the

distribution of the data. The results of the Shapiro-Wilk univariate normality test indicate that all IDLEL variables (i.e., frequency, duration, diversity, and diversity of strategy use) and the listening test scores are non-normally distributed, while the self-regulation, self-efficacy, and listening anxiety are normally distributed. Based on the normality test results, ANOVA, which is suitable to determine whether significant differences exist in means across three or more groups, was conducted for the parametric cluster variables, namely self-regulation, self-efficacy, and listening anxiety. The results revealed that the three groups differed significantly on all of these variables (Table 20). Regarding all non-parametric IDLEL variables and the listening test scores, the Kruskal-Wallis test, which serves as a non-parametric method for determining whether there are statistically significant differences in medians among three or more independent groups, only showed significant differences in the three clusters for duration (Table 20). A Dunn-Bonferroni test and Tukey HSD test were conducted after the Kruskal-Wallis test and ANOVA, respectively, as non-parametric and parametric post-hoc tests to identify which specific groups differed from each other.

The three clusters can be described as follows:

Cluster 1: Positive Listeners With Extensive engagement. This cluster had the highest scores for self-regulation, self-efficacy as well as the duration of IDLEL engagement, while their listening anxiety was the lowest. This suggests that these participants were self-regulated and confident in L2 listening and were willing to spend their extracurricular time on IDLEL activities to learn English. Thus, they can be considered as positive L2 listeners with extensive IDLEL engagement.

Cluster 2: Negative Listeners with Limited engagement. This cluster was characterised

by the second-highest but below-average scores for self-regulation and the duration of IDDEL activity engagement, the lowest scores for self-efficacy, as well as the highest scores for listening anxiety. This suggests that these participants lacked confidence and self-regulation in L2 listening, felt anxious about listening, and were unwilling to spend much of their extracurricular time on IDDEL activities. Therefore, they can be regarded as negative L2 listeners with limited IDDEL engagement.

Cluster 3: Moderately Confident and Anxious Listeners with Low Self-Regulation and Limited engagement. These students had the second-highest, near-average scores for self-efficacy and listening anxiety, as well as the lowest scores for self-regulation and duration of IDDEL engagement. This indicates that, among all participants, those in Cluster Three had a moderate level of listening anxiety and confidence in listening, but they lacked self-regulation and were unwilling to take time to engage in IDDEL activities to improve their English listening in their spare time.

Taken together, this section presented the findings for RQ3, which examines the characteristics of participants' IDDEL engagement. The subsequent section will display the results for RQ4, focusing on the predictive effects of SRL and IDDEL engagement on L2 listening, listening self-efficacy, and listening anxiety.

Table 20*Cluster Profiles*

	Cluster 1 (n=25)		Cluster 2 (n=54)		Cluster 3 (n=12)		Whole Sample (n=91)		Kruskal-Wallis Test (P)	Dunn-Bonferroni Post-Hoc Test	ANOVA	Tukey HSD Post-Hoc Test
	M	SD	M	SD	M	SD	M	SD				
T1 Self-Efficacy	66.30	11.30	48.80	14.20	55.80	13.00	54.54	15.15			<.001***	1, 3>2
T1 Self-regulation	4.72	.61	4.27	.72	3.79	.51	4.33	.72			<.001***	1, 2>3
T1 Listening anxiety	2.78	.65	4.24	.59	3.91	.50	3.80	.86			<.001***	2, 3>1
Duration	8.65	5.07	5.98	3.17	5.24	1.40	6.62	3.86	.02*		1>2, 3	

Note. M = Mean; SD = standard deviation. * $p \leq 0.05$; *** $p \leq .001$. “>” in the Dunn-Bonferroni Post-Hoc Test and TukeyHSD Post-Hoc Test columns indicates the ranking of the means of the cluster

4.5 RQ4: The Prediction of IDLEL Engagement and SRL on L2 Listening, Listening

Anxiety and Self-Efficacy

Descriptive statistics for all measurements, including three listening tests (pre-, post-, and delayed post-test), two SRL, listening anxiety, and listening self-efficacy questionnaires (pre- and post- test), weekly IDLEL engagement (duration, diversity, frequency of activity engagement, and diversity of strategy-use) are displayed in Table 21.

Table 21

Descriptive Statistics for All Measurements

	Mean	SD	Minimum	Maximum
Listening				
pre-test (N=91)	10.80	4.54	4	22
post-test (N=91)	14.53	3.62	5.5	23.5
Delayed post-test (N=60)	11.38	5.45	2	26.5
Self-Regulation (N=91)				
Pre-test	4.33	0.72	2.64	6
post-test	4.71	0.61	3.08	6
Listening Anxiety (N=91)				
Pre-test	3.80	0.86	1.33	5.58
post-test	3.64	0.82	1.5	5.25
Self-Efficacy (N=91)				
Pre-test	54.54	15.15	18.89	86.67
post-test	62.63	14.71	18.89	90
Weekly IDLEL Engagement (N=91)				
Activity Diversity	1.41	0.71	0.67	5
Activity Duration	6.62	3.86	1.5	27
Activity Frequency	2.85	1.04	1	7
Strategy-Use Diversity	3.11	1.17	1.25	8

4.5.1 The Prediction of IDLEL Engagement and SRL on L2 Listening

As explained in *Section 3.5.4.1*, this study constructed two separate LMMs for listening analysis, due to participant attrition ($N=30$) in the third listening test, namely the delayed post-test. Using 91 participants' pre- and post- listening test data, the first listening model aimed to examine the short-term impacts of participants' IDLEL engagement on their listening development; the second listening model aimed to investigate the long-term effects of participants' IDLEL engagement on their listening development, using the 60 participants' pre-, post-, and delayed post- listening test data.

4.5.1.1 The Prediction of IDLEL Engagement and SRL on L2 Listening (Pretest vs. Post-test)

As detailed in *Section 3.5.4.1*, the present study implemented a backward-selection heuristic approach (Matuschek et al., 2017) to refine the theory-driven maximal model models (Barr et al., 2013), thereby achieving the parsimonious mixed models, which could provide the best explanation of the data within a class of candidate models (Cunnings, 2012). Therefore, the first listening LMM incorporated five fixed effects: (a) pretest self-efficacy (T1 self-efficacy), (b) post-test listening anxiety (T2 listening anxiety), (c) post-test self-regulation (T2 self-regulation), (d) weekly IDLEL engagement duration (Average duration), and (e) time (pretest vs. post-test). Additionally, the model included four two-way interactions: (a) T1 self-efficacy \times Time interaction, (b) T2 listening anxiety \times Time interaction, (c) T2 self-regulation \times Time interaction, and (d) Average duration \times Time interaction. Regarding random effect, the model included random intercept for participants.

Table 22*Results for the First Listening Model*

<i>Predictors</i>	<i>Listening</i>		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	10.80	9.98 – 11.62	<0.001
T1 self-efficacy	1.20	0.31 – 2.09	0.008
Time [Post-test]	3.73	3.05 – 4.41	<0.001
T2 listening anxiety	-0.94	-1.79 – -0.08	0.032
T2 self-regulation	-0.42	-1.34 – 0.49	0.361
Average duration	-0.06	-0.96 – 0.84	0.897
Time [Post-test] × T1 self-efficacy	-1.30	-2.05 – -0.56	0.001
Time [Post-test] × T2 listening anxiety	0.81	0.10 – 1.53	0.026
Time [Post-test] × T2 self-regulation	1.18	0.42 – 1.94	0.003
Time [Post-test] × Average duration	1.00	0.25 – 1.75	0.009
Random Effects			
σ^2	5.43		
$\tau_{00\text{ ID}}$	10.17		
ICC	0.65		
N_{ID}	91		
Observations	182		
Marginal R ² / Conditional R ²	0.266 / 0.745		

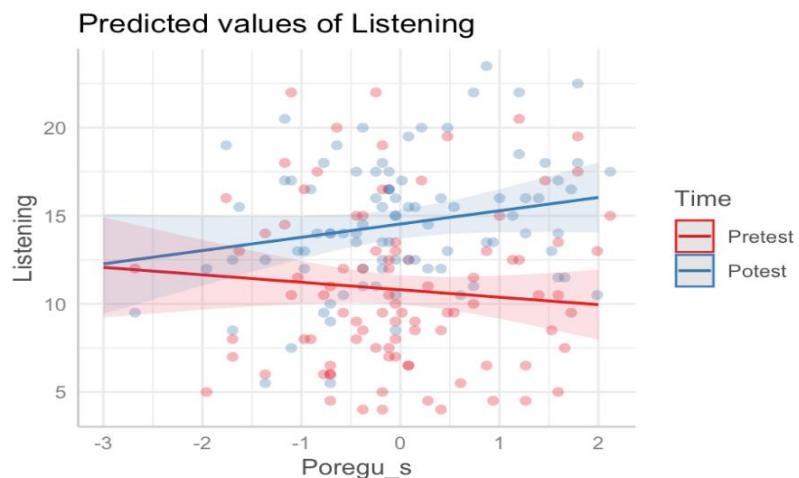
As demonstrated in Table 22, participants ($N= 91$) who took part in the four-week IDLEL study demonstrated an overall improvement in their listening test scores from the pretest (T1) to the post-test (T2). However, the extent of this progress differed based on T2 self-regulation, and the amount of time they spent on weekly IDLEL activities, as the two-way interactions Time × T2 self-regulation and Time × Average duration were both

significant in the final LMM listening model. In other words, T2 self-regulation and the weekly duration of IDDEL engagement were significant positive predictors of participants' listening improvement from the pretest to the post-test.

Specifically, as shown in Figure 12 and Figure 13, participants who dedicated more time to IDDEL activities weekly or exhibited higher T2 listening self-regulation showed greater progress in listening from T1 to T2; conversely, those who spent less time on IDDEL activities or had lower T2 listening self-regulation experienced smaller improvements in their listening proficiency from T1 to T2.

Figure 12

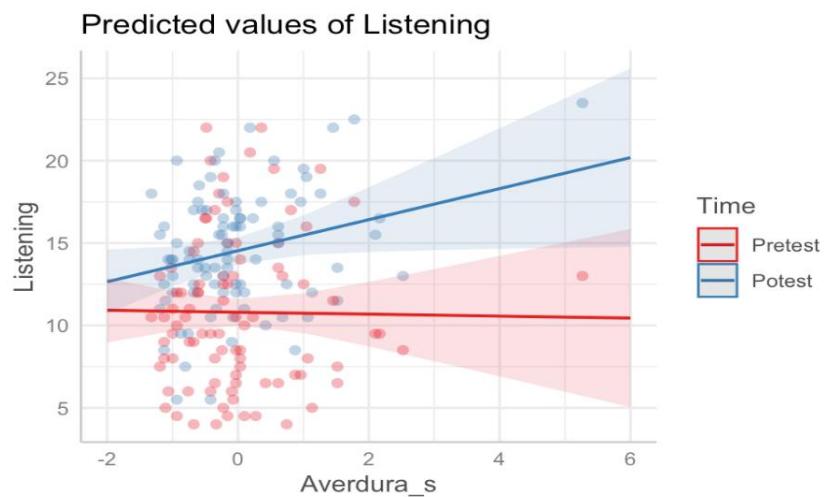
Effect Plots for the Time \times T2 Self-regulation Interaction



Note. Poregu_s = T2 Self-regulation

Figure 13

Effect Plots for the Time × Average duration Interaction

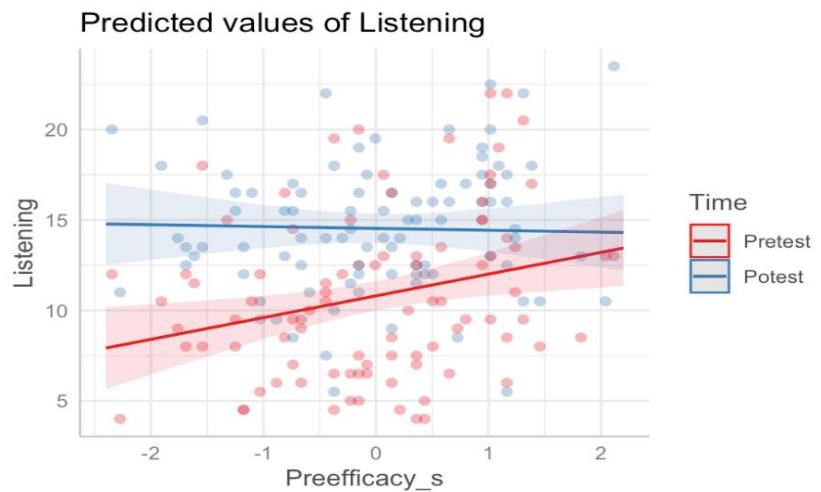


Note. $Averdura_s$ = Average Duration

In addition to T2 Self-regulation and IDLEL duration, the two-way interactions Time × T1 Self-efficacy and Time × T2 Listening Anxiety also showed significance in the LMM listening model. This indicates that T1 self-efficacy and T2 listening anxiety were also significant predictors of participants' listening improvement from the pre-test to the post-test. Specifically, according to Figure 14 and Figure 15, participants who were more confident in their English listening before the IDLEL study (i.e., higher T1 self-efficacy) showed smaller improvements in listening scores from T1 to T2, while those who were less confident in their English listening before the IDLEL study (lower T1 self-efficacy) showed greater improvement in listening from T1 to T2. Rather differently, participants with lower T2 listening anxiety exhibited smaller gains in listening performance from T1 to T2, whereas participants with higher T2 listening anxiety showed greater progress in listening from T1 to T2.

Figure 14

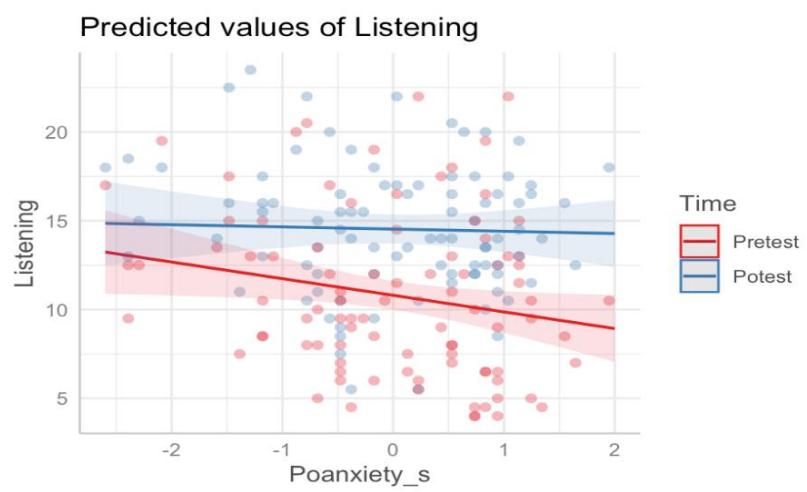
Effect Plots for the Time \times T1 Self-Efficacy Interaction



Note. Preefficacy = T1 Self-Efficacy

Figure 15

Effect Plots for the Time \times T2 Listening Anxiety Interaction



Note. Poanxiety = T2 Listening Anxiety

4.5.1.2 The Prediction of IDLEL Engagement and SRL on L2 Listening (Pretest vs.

Post-test vs. Delayed Post-test)

To investigate whether IDLEL engagement had a long-term impact on learners' listening development, a delayed post-test (i.e., T3 listening test) was conducted three months after the IDLEL study. As Table 23 shows, the second listening LMM included two fixed effects: (a) pretest self-efficacy (T1 self-efficacy) and (b) time (pretest vs. post-test vs. delayed post-test). Additionally, one two-way interaction was included in the model, namely T1 self-efficacy \times Time interaction. Similar to the first listening model, the model specified participant random intercept as the random effect.

Table 23

Results for the Second Listening Model

<i>Predictors</i>	Listening		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	10.30	9.14 – 11.46	<0.001
T1 self-efficacy	1.49	0.32 – 2.65	0.013
Time [Delayed]	1.08	0.17 – 1.98	0.021
Time [Post-test]	3.55	2.64 – 4.46	<0.001
T1 self-efficacy \times Time [Delayed]	-0.82	-1.73 – 0.09	0.077
T1 self-efficacy \times Time [Post-test]	-1.36	-2.28 – -0.45	0.004
Random Effects			
σ^2	6.37		
$\tau_{00\text{ ID}}$	14.52		
ICC	0.70		
N_{ID}	60		
Observations	180		

Marginal R ² / Conditional R ²	0.130 / 0.735
--	---------------

The LMM results indicated that participants' ($N = 60$) listening scores improved from T1 to T3; however, the improvement was smaller compared to the progress observed between T1 and T2. Additionally, non-significant interactions between SRL and Time (T1 vs. T2 vs. T3) as well as between IDLEL engagement and Time in the LLM indicated that neither SRL nor IDLEL engagement was a significant predictor of participants' listening progress from T1 to T2 or from T1 to T3.

4.5.2 The Prediction of IDLEL Engagement and SRL on Self-Efficacy

Apart from the listening models, the present study also constructed a self-efficacy LMM to examine the predictive effects of SRL and IDLEL engagement on participants' listening self-efficacy. According to Table 24, the final self-efficacy LMM contained eight fixed effects: (a) pre-test listening anxiety (T1 listening anxiety), (b) post-test listening anxiety (T2 listening anxiety), (c) pre-test listening test scores (T1 listening), (d) post-test listening test scores (T2 listening), (e) post-test self-regulation (T2 self-regulation), (f) weekly IDLEL engagement duration (Average duration), (g) weekly IDLEL engagement frequency (Average frequency) and (h) time (pretest vs. post-test). Moreover, the self-efficacy model included four two-way interactions: (a) T1 listening anxiety \times Time interaction, (b) T2 listening anxiety \times Time interaction, (c) T1 listening \times Time interaction, and (d) T2 listening \times Time interaction. Similar to the random effect contained in the listening LMMs, the self-efficacy LMM specified participant random intercept.

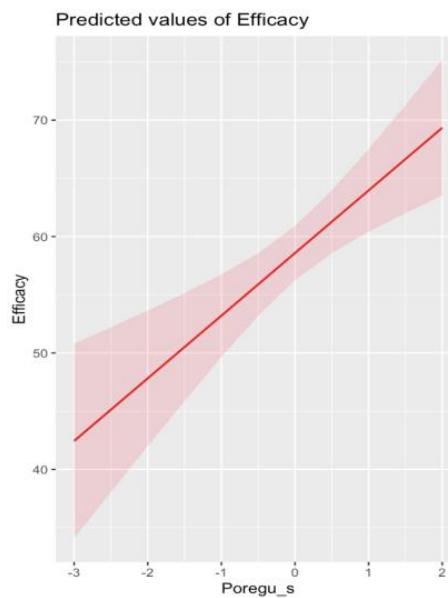
Table 24*Results for the Self-Efficacy Model*

Predictors	Efficacy		
	Estimates	CI	p
(Intercept)	54.54	51.95 – 57.13	<0.001
T2 self-regulation	5.38	2.71 – 8.05	<0.001
T1 listening anxiety	-4.26	-7.86 – -0.67	0.020
Time [Post-test]	8.08	5.82 – 10.35	<0.001
T2 listening anxiety	3.13	-0.49 – 6.76	0.090
T1 listening	6.96	3.48 – 10.45	<0.001
T2 listening	-5.41	-9.02 – -1.80	0.004
Average frequency	-3.31	-6.50 – -0.13	0.042
Average duration	5.53	2.16 – 8.90	0.001
Time [Post-test] × T1 listening anxiety	3.77	0.65 – 6.90	0.018
Time [Post-test] × T2 listening anxiety	-4.57	-7.68 – -1.46	0.004
Time [Post-test] × T1 listening	-3.95	-6.87 – -1.02	0.008
Time [Post-test] × T2 listening	5.05	2.16 – 7.95	0.001
Random Effects			
σ^2	59.99		
$\tau_{00\text{ ID}}$	96.63		
ICC	0.62		
N_{ID}	91		
Observations	182		
Marginal R ² / Conditional R ²	0.380 / 0.762		

The results of the LMM analysis indicated that participants' ($N = 91$) listening self-efficacy showed an overall upward trend from T1 to T2, suggesting that they became

more confident in their English listening after participating in the IDLEL study. Additionally, as Figure 16, 17 and 18 show, the study found that participants with higher T2 listening self-regulation also tended to exhibit higher listening self-efficacy at both T1 and T2; those who spent more time engaging in IDLEL activities generally demonstrated higher listening self-efficacy at T1 and T2, whereas participants who engaged in IDLEL activities more frequently tended to have lower listening self-efficacy. These findings confirm the positive predictive role of SRL and IDLEL duration, as well as the negative predicting effect of IDLEL frequency on participants' overall performance across the two listening tests. However, neither SRL nor IDLEL was found to predict participants' improvement in self-efficacy from T1 to T2, as the two-way interactions of SRL and IDLEL with Time were not significant in the LMM analysis.

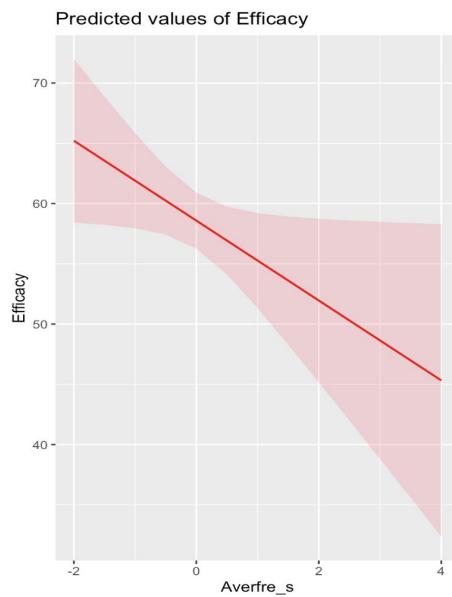
Figure 16
Effect Plots for T2 Self-regulation



Note. Poregu = T2 Self-regulation

Figure 17

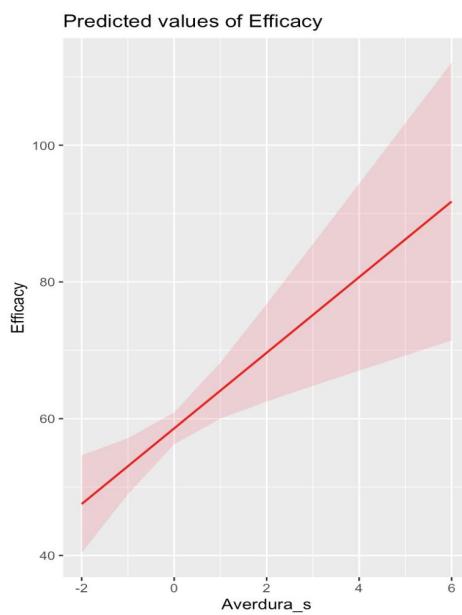
Effect Plots for IDLEL Frequency



Note. Averfre = IDLEL Frequency

Figure 18

Effect Plots for IDLEL Duration



Note. Averdura = IDLEL Duration

The factors found to predict participants' self-efficacy improvement from T1 to T2 were listening anxiety (T1 and T2) and listening test scores (T1 and T2), as their two-way interactions with Time, namely Time \times T1 listening anxiety, Time \times T2 listening anxiety, Time \times T1 listening test scores, and Time \times T2 listening test scores, were significant (Table 24).

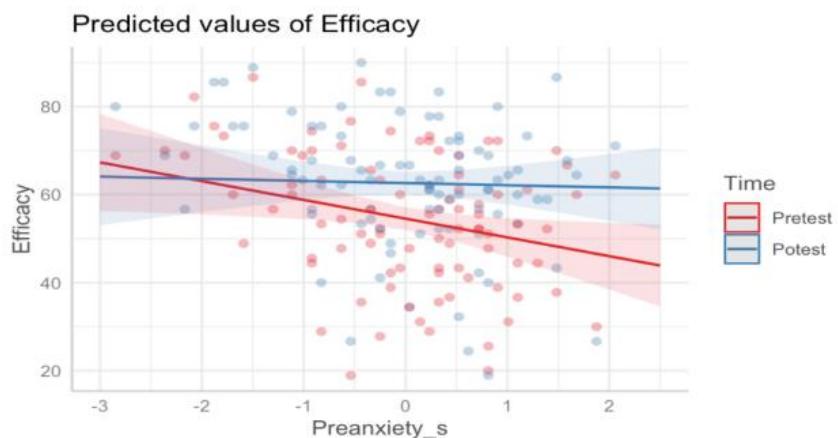
Specifically, as Figure 19 displays, participants who experienced higher listening anxiety before the IDLEL study (i.e., higher T1 listening anxiety) generally exhibited lower initial listening self-efficacy (T1). However, they demonstrated a more significant increase in listening confidence from T1 to T2. In contrast, participants with lower listening anxiety before the study (i.e., lower T1 listening anxiety) tended to have higher T1 self-efficacy but showed smaller gains in listening confidence over time. Additionally, according to Figure 20, participants with lower listening anxiety after the IDLEL study (i.e., lower T2 listening anxiety) often had lower initial listening self-efficacy (T1) but experienced a notable improvement in self-efficacy by the end of the study. By contrast, participants who exhibited higher listening anxiety after the IDLEL study (i.e., higher T2 listening anxiety) generally had higher listening self-efficacy before the study, but their self-efficacy showed only a slight improvement by the end of the study.

Regarding the impact of listening test scores on changes in self-efficacy, as shown in Figure 21, the results indicated that participants with lower T1 listening test scores typically demonstrated lower T1 listening self-efficacy but experienced a significant increase in self-efficacy from T1 to T2. Conversely, participants with higher T1 listening test scores generally had higher T1 self-efficacy but exhibited smaller improvements in self-efficacy

over time. Furthermore, as Figure 22 demonstrates, participants who obtained higher T2 listening test scores tended to have lower initial listening self-efficacy (T1) but showed greater gains in self-efficacy from T1 to T2, whereas participants with lower T2 listening test scores generally displayed higher T1 listening self-efficacy but experienced a smaller increase in self-efficacy over time. Another interesting finding was that participants with lower T2 listening test scores tended to report higher T1 and T2 listening self-efficacy, whereas those with higher T2 listening test scores typically exhibited lower T1 and T2 listening self-efficacy.

Figure 19

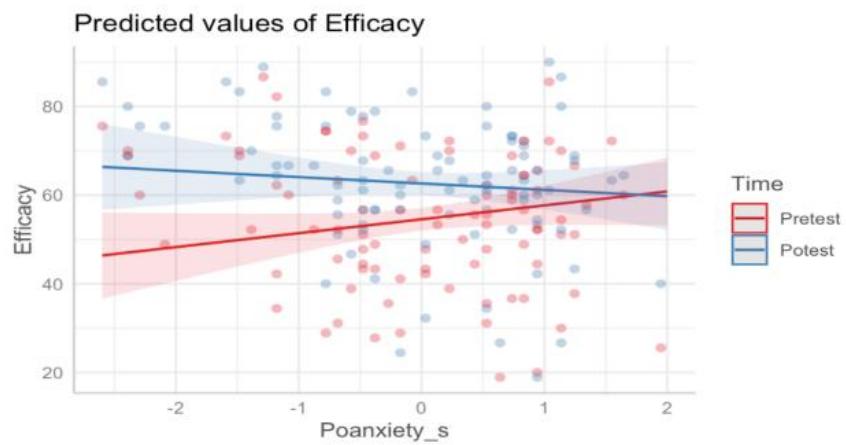
Effect Plots for the Time \times T1 Listening Anxiety Interaction



Note. Preanxiety = T1 Listening Anxiety

Figure 20

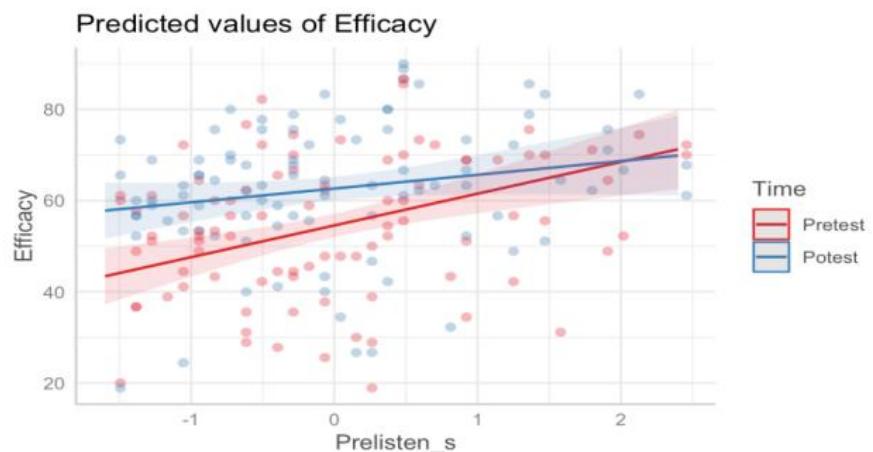
Effect Plots for the Time \times T2 Listening Anxiety Interaction



Note. Poanxiety = T2 Listening Anxiety

Figure 21

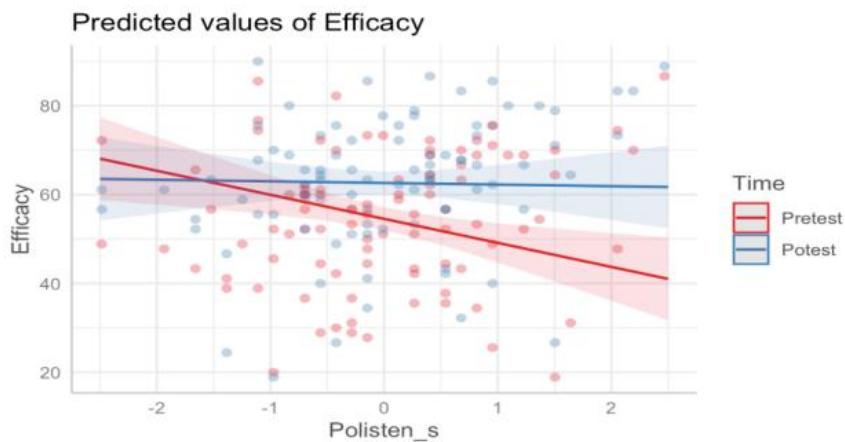
Effect Plots for the Time \times T1 Listening Test Scores Interaction



Note. Prelisten = T1 Listening Test Scores

Figure 22

Effect Plots for the Time × T2 Listening Test Scores Interaction



Note. Polisten = T2 Listening Test Scores

4.5.3 The Prediction of IDLEL Engagement and SRL on Listening Anxiety

The listening anxiety LMM was also constructed to explore the predictive effects of SRL and IDLEL engagement on participants' listening anxiety. According to Table 25, the final listening anxiety LMM included three fixed effects: (a) pre-test listening (T1 listening), (b) post-test self-regulation (T2 self-regulation) and (c) time (pretest vs. post-test). No significant interactions were contained in the model, and random intercept for participants was incorporated as the random effect.

Table 25

Results for the Listening Anxiety Model

Predictors	Anxiety		
	Estimates	CI	p
(Intercept)	3.80	3.63 – 3.96	<0.001
Time [Post-test]	-0.15	-0.30 – -0.01	0.034
T2 self-regulation	-0.21	-0.36 – -0.06	0.006
T1 listening	-0.19	-0.34 – -0.04	0.013

Random Effects

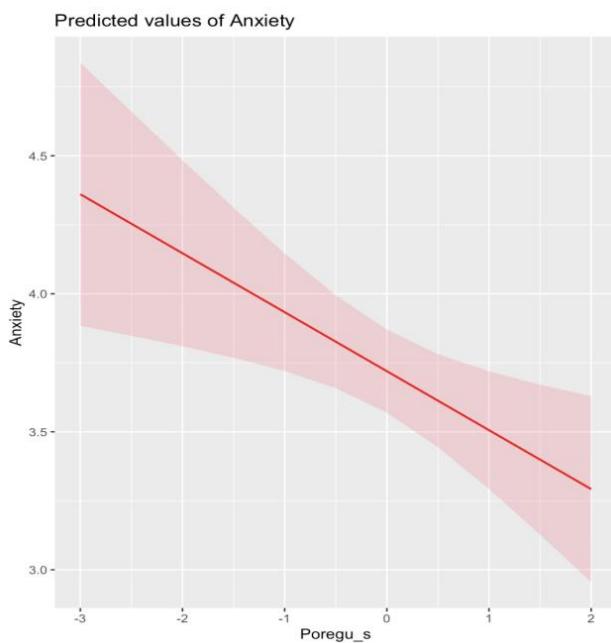
σ^2	0.24
$\tau_{00 \text{ ID}}$	0.41
ICC	0.64
N_{ID}	91
Observations	182
Marginal R^2 / Conditional R^2	0.124 / 0.681

Participants' ($N = 91$) listening anxiety decreased from T1 to T2. SRL and IDEL engagement, however, were not found to significantly predict the decrease in listening anxiety or the overall listening anxiety (T1 and T2). In other words, neither the two-way interaction effects of SRL and IDEL with Time (T1 vs. T2) were significant, nor were SRL and IDEL themselves significant predictors. As a result, they were removed during the model simplification process. Moreover, no factors were identified as significant predictors of changes in participants' listening anxiety from T1 to T2. However, T2 listening self-regulation (Figure 23) and T1 listening test scores (Figure 24) were found to have a significant negative predictive effect on listening anxiety. Specifically, participants with lower T2 self-regulation or lower T1 listening test scores exhibited higher listening anxiety at both T1 and T2, whereas those with higher T2 self-regulation or higher T1 listening test scores demonstrated lower listening anxiety at both time points.

This section presented the findings for RQ4, which investigates the predictive effects of SRL and IDEL engagement on L2 listening, listening self-efficacy, and listening anxiety. The following section will elucidate the results for RQ5, focusing on the moderating role of SRL in the relationships between IDEL engagement and listening, self-efficacy, and listening anxiety.

Figure 23

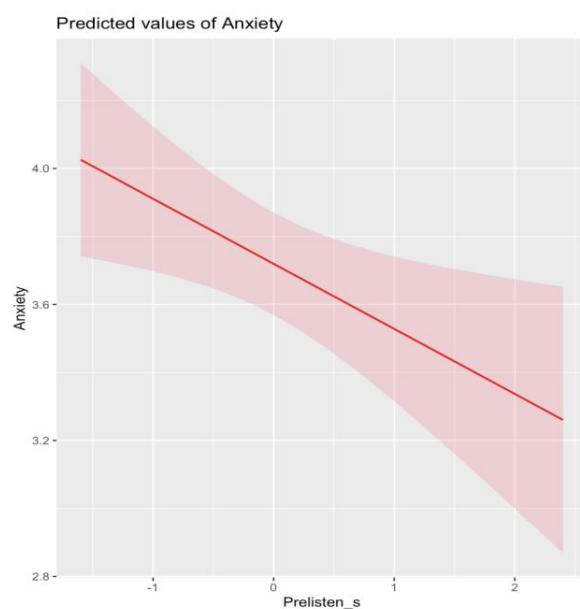
Effect Plots for the T2 listening Self-regulation



Note. Poregu = T2 Self-regulation

Figure 24

Effect Plots for the T1 Listening Test Scores



Note. Prelisten = T1 Listening Test Scores.

4.6 RQ5: The Moderating Effect of SRL on the IDLEL Engagement and L2 Listening,

Listening Anxiety, and Self-Efficacy

To investigate whether the effects of IDLEL engagement on listening comprehension, self-efficacy, and listening anxiety varied depending on learners' self-regulation abilities, moderation analysis was conducted. A moderating effect analysis is typically conducted under the premise that a relationship already exists between the independent variable and the dependent variable, as noted in the *Section 3.5.5.1* (Baron & Kenny, 1986). As indicated in the previous discussion of the LMM results, the duration of IDLEL engagement significantly predicted participants' listening improvement from T1 to T2, while both the duration and frequency of IDLEL engagement significantly predicted participants' overall listening self-efficacy (i.e., T1 and T2). However, IDLEL engagement was not found to be a significant predictor of listening anxiety. Therefore, this study examined whether SRL moderated the relationship between IDLEL engagement and listening, as well as the relationship between IDLEL engagement and self-efficacy.

Specifically, the first moderation LMM explored the moderating role of SRL on IDLEL engagement and listening. According to Table 26, four fixed effects were included in the model: (a) pre-test self-regulation (T1 self-regulation), (b) post-test self-regulation (T2 self-regulation), (c) weekly IDLEL engagement duration (Average duration) and (d) time (pretest vs. post-test). Additionally, the first moderation model contained two three-way interactions: (a) Average duration \times T1 self-regulation \times Time interaction and (b) Average duration \times T2 self-regulation \times Time interaction. Moreover, participant random intercept was incorporated in the model as the random effect. However, no three-way interactions were

found significant, indicating that SRL failed to moderate between participants' IDLEL engagement duration and listening.

Table 26
Results for the First Moderation Model

Predictors	Listening		
	Estimates	CI	p
(Intercept)	10.63	9.71 – 11.54	<0.001
Time [Post-test]	3.86	3.04 – 4.67	<0.001
T1 self-regulation	0.72	-0.30 – 1.74	0.165
Average duration	-0.11	-1.22 – 0.99	0.837
T2 self-regulation	-0.22	-1.25 – 0.82	0.681
Time [Post-test] × T1 self-regulation	-0.30	-1.21 – 0.61	0.519
Time [Post-test] × Average duration	0.88	-0.10 – 1.86	0.078
T1 self-regulation × Average duration	0.31	-0.96 – 1.57	0.635
Time [Post-test] × T2 self-regulation	0.81	-0.11 – 1.74	0.084
Average duration × T2 self-regulation	0.24	-0.90 – 1.38	0.678
Time [Post-test] × T1 self-regulation × Average duration	0.15	-0.98 – 1.29	0.787
Time [Post-test] × Average duration × T2 self-regulation	-0.51	-1.52 – 0.51	0.324
Random Effects			
σ^2	6.60		
$\tau_{00\text{ ID}}$	10.03		
ICC	0.60		
N ID	91		
Observations	182		
Marginal R ² / Conditional R ²	0.226 / 0.693		

To understand the moderating effect of SRL on IDLEL engagement and listening self-efficacy, the second moderation LMM was constructed. As Table 27 demonstrates, the second model contained five fixed effects: (a) pre-test self-regulation (T1 self-regulation), (b) post-test self-regulation (T2 self-regulation), (c) weekly IDLEL engagement duration (Average duration), (d) weekly IDLEL engagement frequency (Average frequency) and (e) time (pretest vs. post-test). Four three-way interactions were included: (a) Average duration \times T1 self-regulation \times Time interaction and (b) Average duration \times T2 self-regulation \times Time interaction, (c) Average frequency \times T1 self-regulation \times Time interaction and (d) Average frequency \times T2 self-regulation \times Time interaction. As for the random effect, participant random intercept was incorporated. Similar to the first moderation model, all three-way interactions were found insignificant, suggesting that SRL did not moderate between participants' IDLEL engagement and listening self-efficacy.

In summary, this chapter presented the analytical results of the five research questions, and the following chapter will provide an in-depth discussion of these results.

Table 27
Results for the First Moderation Model

Predictors	Efficacy		
	Estimates	CI	p
(Intercept)	55.71	52.64 – 58.79	<0.001
Time [Post-test]	8.19	5.46 – 10.93	<0.001
T1 self-regulation	0.48	-3.09 – 4.05	0.791
Average duration	7.70	2.77 – 12.63	0.002
T2 self-regulation	2.72	-0.64 – 6.08	0.112
Average frequency	-2.98	-6.88 – 0.92	0.133

Time [Post-test] \times T1 self-regulation	-1.67	-4.85 – 1.50	0.299
Time [Post-test] \times Average duration	-0.96	-5.34 – 3.42	0.666
T1 self-regulation \times Average duration	-5.41	-13.13 – 2.31	0.168
Time [Post-test] \times T2 self-regulation	4.22	1.23 – 7.21	0.006
Average duration \times T2 self-regulation	-0.33	-6.83 – 6.16	0.919
Time [Post-test] \times Average frequency	1.01	-2.46 – 4.47	0.568
T1 self-regulation \times Average frequency	5.45	-0.64 – 11.54	0.079
T2 self-regulation \times Average frequency	-0.34	-5.73 – 5.05	0.900
(Time [Post-test] \times T1 self-regulation s) \times Average duration	0.93	-5.93 – 7.79	0.790
(Time [Post-test] \times Average duration s) \times T2 self-regulation	-1.47	-7.24 – 4.30	0.616
(Time [Post-test] \times T1 self-regulation s) \times Average frequency	-1.34	-6.75 – 4.07	0.626
(Time [Post-test] \times T2 self-regulation s) \times Average frequency	1.19	-3.61 – 5.98	0.626

Random Effects

σ^2	68.59
$\tau_{00\text{ ID}}$	105.07
ICC	0.61
N ID	91
Observations	182
Marginal R ² / Conditional R ²	0.324 / 0.733

Chapter 5. DISCUSSION

5.1 Introduction

As articulated in *Sections 1.1.1* and *1.1.2*, the researcher has identified a critical dilemma faced by Chinese university EFL learners: the paramount importance of L2 listening contrasts sharply with insufficient instructional support in classroom contexts. Consequently, engaging in IDLEL activities may serve as a viable alternative for enhancing English listening proficiency. However, in informal learning contexts where external regulation from authority figures (e.g., teachers) can be absent, SRL assumes heightened significance. Moreover, given that Chinese university students often struggle with moderate listening anxiety and less-than-optimal listening self-efficacy, this study seeks not only to examine learners' IDLEL engagement patterns (RQ3) but also to investigate the influences of SRL and IDLEL engagement on their L2 listening, listening self-efficacy, and listening anxiety (RQ4), while exploring the moderating role of SRL in the relationships between IDLEL engagement and L2 listening, listening self-efficacy, and listening anxiety (RQ5). During the selection of theoretical frameworks and measurement tools for SRL, a notable gap emerged: the absence of a comprehensive L2 listening-specific SRL model to elucidate the mechanisms of self-regulated L2 listening development (RQ1). Meanwhile, the joint predictive effects of SRL, self-efficacy, and listening anxiety on L2 listening remain underexplored (RQ2).

To answer these research questions, CFA, SEM, mediation analysis, descriptive analysis, thematic analysis, cluster analysis, and LMMs were conducted. The preceding

chapter presented the analytical results for the five research questions, and this chapter will systematically discuss these findings following the sequence of the research questions.

5.2 Discussion for RQ1: Validation of the Construct of the Self-regulated L2 Listening Model

The new model aimed to depict the structure and mechanisms of self-regulated L2 listening at both the task-level (i.e., cognition-level) and the motivation/affect-level. Five phases were included in the model and each phase involved both cognitive regulation as well as motivational and affective regulation. CFA results provided substantial evidence for the five-construct structure of the *Self-Regulated L2 Listening Questionnaire (SRLLQ)*, which was developed using the newly proposed model as the theoretical framework, with items within each construct related to both cognitive and motivational/affective regulation. In other words, the five-phase dual-level construct of the self-regulated L2 listening model was empirically supported.

On the one hand, the key phases of previous SRL models (e.g., preparatory, performance, and appraisal phases, Puustinen & Pulkkinen, 2001) were also included in the self-regulated L2 listening model but presented in a more detailed and clear manner. For instance, the preparatory phase corresponds to the *Task Representation* and *Goal Setting & Strategy Planning* phases in the new model. Separating goal setting and strategy planning from task perception as the new model does can highlight the importance of the latter in influencing subsequent SRL phases (Ranalli, 2012). The task execution (performance) phase in other models corresponds to the *Performance* and *Monitoring & Control* phases in the new model. Although presented as an independent phase, the monitoring and control of

task processing and motivational/affective states pervade the other four phases of the self-regulated L2 listening model. Finally, the appraisal phase corresponds to the *Attribution & Adjustment* phase in the new model.

On the other hand, the new SRL model, designed specifically for L2 listening and applicable to different learning contexts, addresses the limitations of many classic SRL models that fail to reflect the uniqueness of language learning (e.g., the Cyclical Phases Model, Zimmerman, 2000). It also compensates for the shortcomings of models like the S²R model (Oxford, 2017), which is targeted at L2 learning but does not fully capture the characteristics of L2 listening. Moreover, the new SRL model details the mechanisms of motivational and affective self-regulation, following the same process as cognitive regulation. This addresses the limitations of some models (e.g., the COPES model, Winne & Hadwin, 1998) that lack regulation of motivation/affect, as well as other models (e.g., the MASRL model, Efklides, 2011) that fail to clearly demonstrate the process of motivational and affective self-regulation. By emphasising both the multi-dimensional and multi-stage features of SRL, the model not only advances current SRL theory but also broadens its application to the domain of L2 learning, particularly L2 listening.

Along with the validation of the self-regulated L2 listening model, the SRLLQ was also validated. Unlike most listening questionnaires incorporating SRL components, such as the MALQ (Vandergrift et al., 2006), the SRLLQ is grounded in SRL theory, namely, the self-regulated L2 listening model. Its items not only cover the five phases of SRL and the different stages of listening comprehension (before, during, and after listening) but also encompass both task-level and motivational/affective self-regulation strategies, reflecting

the multidimensional nature of self-regulated learning. However, while the SRLLQ is one of the few listening questionnaires with these advantages, it is not the only one.

For instance, Zhou et al. (2024) developed the *Mobile-assisted Self-Regulated Listening Strategy Questionnaire (MSRLS-Q)* based on Zimmerman's (2000) Cyclical Phase Model to explore students' self-regulated learning strategies in mobile-assisted language learning (MALL) contexts when practicing listening. The MSRLS-Q includes items designed to comprehensively cover the three phases of SRL: forethought, performance, and reflection, aiming to capture learners' self-regulation before, during, and after listening, and the SRLLQ exhibits similar characteristics in this regard. Additionally, the items in MSRLS-Q address learners' cognitive, motivational, and social strategy use, thereby accounting for the multidimensionality of SRL (Zhou et al., 2024).

However, differences also exist between the two questionnaires. Firstly, the MSRLS-Q was designed for a specific application context, namely MALL. Consequently, many of its items are closely related to learners' self-regulated listening activities and behaviours within MALL contexts. By contrast, as the self-regulated L2 listening model does not target a specific application setting, the SRLLQ, which was developed based on this model, is not restricted to any particular context. Secondly, while the MSRLS-Q emphasises the importance of maintaining motivation (e.g., interest) in self-regulated listening, the SRLLQ incorporates items addressing the regulation of learners' listening self-efficacy and listening anxiety, considering their potential impact on learners' listening development.

5.3 Discussion for RQ2: Validation of the Joint Predictive Mechanisms

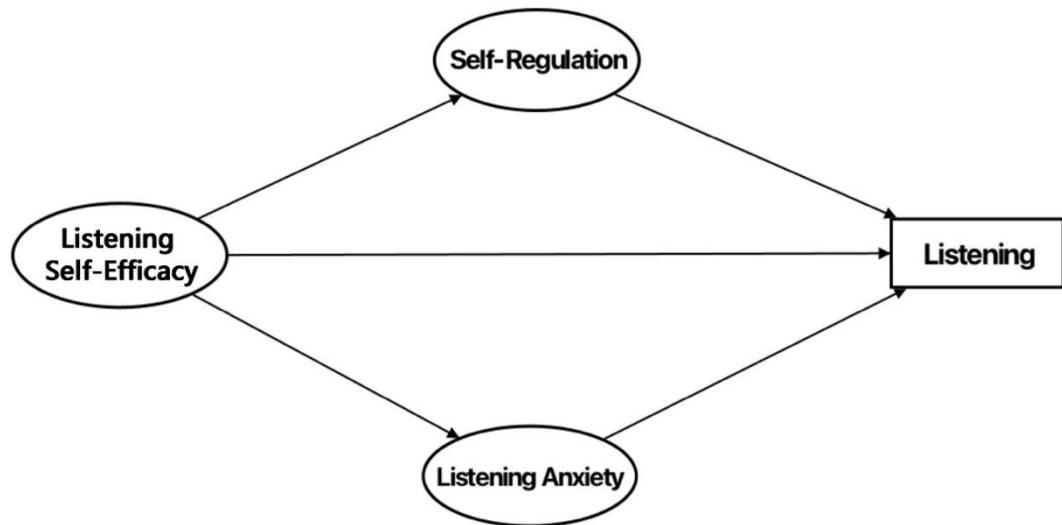
5.3.1 The Mediating Role of SRL and Listening Anxiety

Framed within Social Cognitive Theory (SCT, Bandura, 1990, 1991), two hypothesised models depicting the joint predictive mechanisms were proposed. These two models complement each other while revealing the distinct roles SRL may play within the joint prediction mechanisms (see *Section 2.3.4.4* for details).

Specifically, the first model (Figure 25), validated by SEM, focused on the direct predictive effect of self-efficacy on listening and its indirect predictive effect on listening mediated by SRL and listening anxiety.

Figure 25

Hypothesised Model One: the Mediating Role of Self-Regulation and Listening Anxiety



The findings first confirmed the direct predictive role of self-efficacy on listening, indicating that L2 learners with higher self-efficacy were likely to have better listening performance. Similar results can be found in previous studies. For instance, a study conducted by Du and Man (2023) involving Chinese university EFL learners showed that participants' listening self-efficacy had a significant positive direct predictive effect on L2 listening, measured by a nationwide standardised English language test using non-authentic,

scripted listening materials, that is, from the CET-4 test. Additionally, Chen's (2007) study with Taiwanese college-level EFL learners also validated the significant positive effect of listening self-efficacy on learners' L2 listening proficiency, which was measured by their listening course grades over a period of time, including listening quizzes, tests, and final exam scores. The listening measurement instrument used in this study differed from those used in the research by Du and Man (2023) and Chen (2007); it included not only the non-authentic or semi-authentic listening materials commonly used in standardised listening tests and formal language learning contexts but also authentic materials derived from real-life situations, which are commonly encountered in informal learning contexts. The results of this study thus further empirically support the direct predictive role of listening self-efficacy on listening, suggesting that the influence of self-efficacy on listening can be consistent across different types of listening, both formal and informal.

This study also confirmed the indirect predictive effect of self-efficacy on listening through SRL, revealing the joint predictive power of motivational and cognitive variables on L2 listening. According to SCT, self-efficacy can be an important factor influencing SRL, which can affect the functioning of SRL subfunctions such as goal setting and attribution (Bandura, 1991). Learners with higher self-efficacy are more likely to set higher goals, use more effective analytical strategies, and attribute their failure to a lack of effort rather than low ability (Bandura, 1991). Therefore, high self-efficacy L2 listeners are more likely to experience a greater sense of control over listening, more likely to actively engage in metacognitive listening comprehension, and more likely to successfully apply strategies to complete listening tasks, leading to better listening comprehension (Du & Man, 2023;

Rahimi & Abedi, 2014). This study's findings are consistent with Du and Man (2023) as well as Zhang and Xu (2024). Both of these studies found that listening self-efficacy indirectly predicted listening comprehension through metacognitive awareness, which serves as a key component of SRL. The findings of the present study thus provide empirical evidence for the joint predictive power of the interplay between self-efficacy and SRL on listening.

SEM also validated the indirect predictive effect of self-efficacy on listening through listening anxiety, revealing the mechanism by which affective factors act as the mediator, interacting with motivational factors to jointly predict listening. This result provides empirical support for SCT regarding the relationship between self-efficacy and anxiety, indicating that self-efficacy can be an essential factor in triggering anxiety in L2 listening. Specifically, anxiety, triggered by low self-efficacy, can cause learners to shift their attention from task processing to concerns about their perceived inadequacies, thereby negatively impacting their academic performance (Bandura, 1990). Conversely, high self-efficacy may have an inhibitory effect on the arousal of anxiety, helping learners focus on the current task and invest sustained effort, which may positively influence their academic performance. The indirect effect of self-efficacy on linguistic performance through anxiety has been previously observed for L2 speaking. For example, Passiatore and colleagues (2019) discovered the indirect effect of participants' foreign language self-efficacy on their in-class English speaking performance mediated by foreign language anxiety in their study among 132 Italian high-school EFL learners. However, in the field of L2 listening research, the indirect predictive effect of self-efficacy on L2 listening through

listening anxiety has not been explored, to our knowledge. The findings of the current study thus provide supporting evidence to fill this research gap.

Self-efficacy was found to be a weak predictor of SRL in the present study, explaining only 9.8% of its variance, that is, below 10% and therefore having low explanatory power (Hair & Alamer, 2022). One possible explanation is that, besides SRL, the current research also included another mediator, namely listening anxiety, in the analysis. The two mediators likely competed with each other, thereby affecting the explanatory power of self-efficacy on SRL. The fact that self-efficacy was found in this study to have a higher explanatory power for listening anxiety (14.4%) than for SRL can support this explanation. A second possible explanation is that apart from self-efficacy, other factors had a stronger explanatory power for SRL. For instance, as previously demonstrated, listening anxiety and listening comprehension have been found to be significantly negatively and positively correlated with SRL, respectively. Therefore, although the directional predictive effects of listening anxiety and listening comprehension on SRL were not verified in the hypothesised model, they may have also influenced participants' SRL.

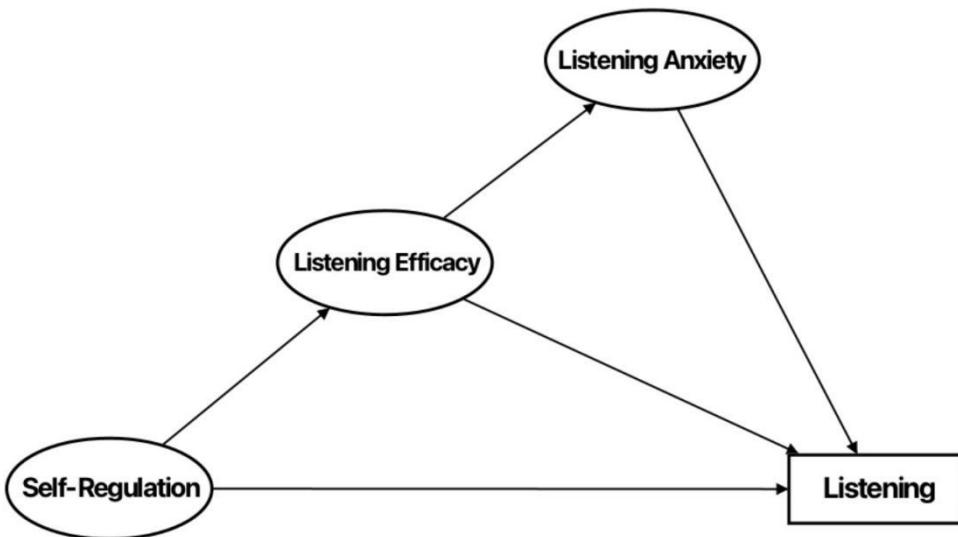
5.3.2 The Mediating Role of Self-Efficacy and Listening Anxiety

As Figure 26 shows, the second hypothesised model validated by SEM focused on the direct predictive effect of SRL on listening and its indirect predictive effect on listening through self-efficacy, revealing the mechanism by which motivational factors act as the mediator, interacting with cognitive factors to jointly predict listening. At the same time, similar to the first model, the direct predictive effect of self-efficacy on listening and its

indirect predictive effect on listening through listening anxiety were also predictive pathways of interest in the second hypothesised model. The interaction between self-efficacy and listening anxiety in predicting listening has been elaborated on earlier, so that discussion will not be repeated here.

Figure 26

Hypothesised Model Two: The Mediating Role of Self-Efficacy and Listening Anxiety



The finding that SRL directly predicted listening aligns with Yabukoshi (2024), who explored the impact on the listening proficiency of Japanese undergraduate EFL learners of their perceived SRL strategy use during out-of-class listening practice. The results indicated that L2 learners with higher self-regulation, as manifested in strategy use, are more likely to have better listening performance. Similarly, the indirect effect of SRL on listening through self-efficacy, which was validated in the present study but had not been fully explored in previous research, suggests that SRL can help learners gain a sense of control over their listening, thereby influencing their self-efficacy, which in turn may positively impact listening comprehension (Graham, 2011; Rahimi & Abedi, 2014). This is because people's

perception of controllability in environments can motivate them to fully exert their personal efficacy to bring about change (Bandura, 1991). For example, the influence of attribution, an important subfunction of SRL, on academic performance is believed to be mediated by self-efficacy (Bandura, 1991). Different types of attributions may cause learners to experience varying degrees of control in their learning, which in turn may affect their self-efficacy and subsequently their academic performance (Graham & Macaro, 2008; Pintrich, 2000). Specifically, external attributions, such as the speed of the listening material, or the surrounding environment, may cause learners to experience a lower sense of control over their listening, thereby reducing their listening self-efficacy, which in turn may negatively impact their listening performance. Conversely, internal attributions, such as “the degree of effort exerted or the strategies employed on a task” (Graham, 2011, p. 114) may help learners maintain or even enhance their sense of control over listening, thereby improving their listening self-efficacy and ultimately having a positive impact on their listening performance.

As was the case for the predictive power of self-efficacy on SRL, SRL also exhibited weak explanatory power ($9.7\% < 10\%$) on self-efficacy in the second hypothesised mechanism, indicating that self-efficacy was influenced by other factors. Firstly, learners' self-efficacy levels when completing the questionnaire may have been influenced by their self-perceived listening performance, similar to what was suggested for listening anxiety earlier. Secondly, considering the significant negative relationship between self-efficacy and anxiety in the current study observed in previous research (e.g., Canaran et al., 2024), it can be inferred that listening anxiety was one of the factors influencing self-efficacy in this

study. Furthermore, even though the findings of the RQ4 in the present study (see *Section 4.5.2*), along with other studies focusing on IDLE (e.g., Zadorozhnyy & Lee, 2023), found that learners' IDLE activity engagement had a significant impact on self-efficacy, it should be noted that the SEM in the present research used the pre-test data collected before the four-week IDLEL observational study. Therefore, participants' engagement in IDLEL activities could not explain their listening self-efficacy as reported in the pre-test questionnaire.

Apart from differences in the roles of self-efficacy and SRL in the joint predictive mechanisms, the two models share similarities. Firstly, the explanatory power of self-efficacy on listening anxiety within both models was at a modest level (Hair & Alamer, 2022), with between 11% and 30% of variance explained suggesting that although self-efficacy is considered an important factor in triggering anxiety (Bandura, 1990), it is not the only pathway. For instance, Efklides (2011) argued that the negative affective changes resulting from metacognitive monitoring of interruptions in cognitive processing, such as the arousal of anxiety, can be the result of unconscious, automatic, and spontaneous affective regulation. Additionally, since the listening anxiety questionnaire was administered immediately after the listening test, before the participants knew their listening test results, participants' self-perceived listening performance may have affected their listening anxiety.

The two hypothesised models shared another important similarity: both found that SRL was the strongest predictor of L2 listening, followed by self-efficacy and listening anxiety, suggesting that SRL appears to have a greater impact on listening than self-efficacy

and listening anxiety. That self-efficacy had a stronger predictive power on listening than listening anxiety did is consistent with Chen (2007) as well as with Zhang and Xu (2024), implying that enhancing learners' self-efficacy could be more effective than alleviating listening anxiety in improving listening proficiency.

The more powerful effect of SRL than self-efficacy on listening also aligns with Zhang and Xu (2024). They also found a stronger predictive effect of metacognitive awareness ($\beta = .21$), which serves as a key element of SRL, on listening than self-efficacy ($\beta = .16$). By contrast Du and Man (2023) found the opposite, namely, a lower predictive power of metacognitive awareness ($\beta = .13$) than self-efficacy ($\beta = .14$) on listening. These contradictory results suggest that the inclusion or exclusion of listening anxiety in a model may affect the strength of the predictive power of self-efficacy and SRL on listening. Specifically, self-efficacy can have a lower predictive effect on listening than SRL when anxiety is included, while its predictive power can be higher than that of SRL when anxiety is excluded. One possible explanation is that anxiety may weaken the effect of self-efficacy on listening, preventing learners from reaching their full potential. This can happen when anxiety acts as a mediator between self-efficacy and listening, as observed in this study, or when it serves as an equivalent independent variable alongside self-efficacy simultaneously predicting listening through metacognitive awareness (Zhang & Xu, 2024). The presence of anxiety also underscores the importance of SRL, which can help learners manage anxiety and reduce its negative effects on listening. Without the interference of negative affect, namely anxiety, however, self-efficacy can influence learners' listening in a more direct way, while the effect of SRL on listening may be comparatively weaker due to the lack of

motivation to regulate anxiety and self-efficacy. This finding thus reflects the significant role of listening anxiety in the joint predictive mechanism.

Taken together, the first hypothesised mechanism focused on the prediction of listening by a motivational factor (i.e., self-efficacy), including both direct prediction and indirect prediction mediated by cognitive (i.e., SRL) and affective factors (i.e., listening anxiety). The second hypothesised mechanism focused on the prediction of listening by a cognitive factor, including both direct prediction and indirect prediction mediated by motivational and affective factors.

The two validated joint predictive mechanisms complement each other, offering two essential pieces of the puzzle in understanding the interplay among motivational, affective, and cognitive factors in predicting L2 listening comprehension. The critical role of SRL in this joint predictive framework is particularly highlighted. First, the findings indicate that SRL not only exerts a direct impact on learners' listening performance, but also influences it indirectly by shaping their self-efficacy and listening anxiety. These multiple pathways through which SRL affects listening create a cumulative effect, suggesting that SRL not only enables learners to manage current listening tasks effectively, but also has a far-reaching impact on future listening performance by regulating motivational and affective factors. Second, SRL is found to play distinct roles within the joint predictive mechanisms for listening: it serves both as a crucial tool for transforming learners' belief in their ability to complete listening tasks into actual performance, and as a powerful booster of learning motivation. Taken together, the study further underscores the pivotal role of SRL in L2 learners' listening development and highlights the necessity of integrating SRL

theory into the field of L2 listening research.

5.4 Discussion for RQ3: Characteristics of IDLEL Engagement

5.4.1 The Quantity of IDLEL Engagement

The quantity of participants' ($N = 91$) IDLEL engagement in the present study was operationalised as the frequency of their engagement, that is, the number of times they engaged in IDLEL activities over a period, and the duration of their engagement in IDLEL activities.

5.4.1.1 Frequency of Engagement

Frequency in this study refers to the number of times participants engaged in IDLEL activities over a period. Descriptive statistics showed that, on average, the 91 participants engaged in IDLEL activities 10-11 times over the four-week IDLEL study, with an average weekly engagement of two to three times. Among these activities, listening to English songs was the most frequently mentioned activity, while listening to English audiobooks was the least frequent. Although previous research has not extensively explored how often learners engage in IDLE activities during a period of time, this variable is undoubtedly a crucial indicator of understanding learner engagement. As discussed in *Section 2.4.4.1*, engagement frequency can reflect L2 learners' motivation and interest in informal language learning contexts, as well as their extracurricular L2 learning habits, such as whether they have incorporated English listening practice into their daily routines. Frequent and thus sustained exposure to authentic English listening materials can not only help to maintain learners' interest and motivation in L2 listening but also enhance their sensitivity to spoken English input. Additionally, it can facilitate the reinforcement and internalisation of L2 knowledge, listening strategies, and self-regulation strategies, thereby supporting the continuity and stability of their L2 acquisition. Therefore, the

frequency of IDDEL engagement can be a factor that deserves much attention from L2 educators and researchers.

5.4.1.2 Duration of Engagement

L2 acquisition requires not only regular exposure to the target language but also contact with it of sufficient duration. The duration of engagement in informal digital language learning activities has been more widely explored in existing research. For example, in two studies conducted among Korean university EFL learners, Lee (2019a, b) found that participants (62.0% and 49.4% in two studies respectively) who spent less than one hour per day on informal digital learning of English activities (IDLE) outnumbered those (22.5%; 27.3%) who spent one to two hours per day and those (15.5%; 23.3%) who spent more than two hours per day. Moreover, Zhang and Liu's (2024) study found that among Chinese undergraduate EFL learners, the fewest participants (10.5%) spent fewer than five hours daily on IDLE activities, while the majority of participants (38.1%) spent six to seven hours. Participants in the present research invested even less time in IDDEL activities, with fewer instances of extended engagement than in the studies by Lee (2019a, b) as well as Zhang and Liu (2024): only 4.4% participants spent more than two hours daily on IDDEL activities, compared with 70.3% who spent less than one hour per day on IDDEL activities. This result may imply that participants in the present study placed a lower priority on IDDEL activities, seemingly lacking motivation and initiative to fully utilise their out-of-class time to improve their English listening skills.

One possible reason for variation in IDLE engagement across studies may be the type of informal language learning activity involved. Most existing research focuses on informal language learning activities that integrate multiple language skills (e.g., reading, writing,

speaking, listening) as a whole (Lee, 2019a, b; Zhang & Liu, 2024), whereas this study specifically investigated the duration of engagement in informal L2 listening activities. The nature and requirements of activities targeting different language skills vary, which may influence the amount of time participants invest in each activity. In other words, the frequency and duration of IDDEL engagement together shape the quantitative characteristics of participants' IDDEL experiences.

5.4.2 The Quality of IDDEL Engagement

It has been argued that when examining the impact of out-of-class learning, not only the quantity of activities but also their quality should be focused on (Lai et al., 2015). In previous studies (Lai et al., 2015; Lee, 2019a, b; Lee & Dressman, 2018), the quality of learners' IDLE engagement has been assessed by the diversity of IDLE activities they engaged in, as different activities can meet different learning needs and thus have varied effects on learning outcomes. High-quality IDLE experiences can therefore be considered to involve a network of diversified IDLE activities that help satisfy learners' diverse language learning needs and achieve balanced learning (Lai et al., 2015). Similarly, engaging in various types of IDDEL activities means that participants can access different types of English listening input, thereby broadening the scope of their L2 listening learning. Thus, the diversity of IDDEL activities was used in this study as an important indicator to evaluate the quality of participants' IDDEL engagement.

Additionally, the quality of participants' IDDEL engagement in the current study was also reflected by the diversity of SRL strategies they used in IDDEL activities. The use of SRL strategies can help enhance the planning and goal-oriented nature of participants' informal language learning, enabling them to tackle challenges more effectively in IDDEL activities,

flexibly adjust their learning strategies, and thus improve the development of their L2 listening outside the classroom. The diversity of learning strategies, furthermore, can reflect learners' adaptability to different task demands and learning contexts, and can help researchers evaluate whether learners have achieved a balanced self-regulation across cognitive, motivational, and affective dimensions. Therefore, the diversity of strategy use was employed in the current study to investigate the quality of participants' IDLEL engagement.

5.4.2.1 Diversity of IDLEL Activities

Descriptive statistics showed that participants engaged in a total of 10 types of IDLEL activities over the four-week study, with an average of one to two types of IDLEL activities per week. Among the 10 types of activities, the activity chosen by the most participants was listening to English songs, while the activity with the fewest participants was listening to English audiobooks. However, as Lai et al. (2018) pointed out, focusing solely on identifying the technological resources learners use for language learning outside the classroom may not provide a comprehensive and detailed description of the nature of their informal language learning; what matters more is understanding how learners perceive and utilise these technological resources to achieve various goals in their out-of-class language learning. Therefore, when discussing the diversity of informal language learning activities, existing research (Lee, 2019a, b; Lee & Dressman, 2018) typically adopts a learner-perspective framework for categorizing informal language learning activities, such as the classification proposed by Lai et al. (2015).

According to Lai et al.'s (2015) classification and as indicated by participants' IDLEL E-logs, IDLEL activities in the present study such as using English learning apps and practising

listening for exams can be considered to be form-oriented activities, which focus more on the target language system, formal elements of the language, and linguistic accuracy within the informal digital language learning environment. The other eight activities, such as watching English movies, TV shows, talk shows, and listening to English songs, seem more aligned with the definition of meaning-oriented activities, which emphasise the provision of naturalistic target language exposure in a real, informal digital environment, with a primary focus on meaning and communication. Consistent with previous research (Lai et al., 2015; Lee, 2019a, b; Lee & Dressman, 2018), participants in this study engaged more frequently in meaning-oriented informal learning activities than in form-oriented ones. On the one hand, the results may suggest that participants' IDDEL activity diversity was somewhat lacking. In language learning, diversity is expected to manifest as a balanced focus on form and meaning (Lai et al., 2015), and it has been found to positively impact learners' language learning outcomes (e.g., vocabulary, speaking) (Lee, 2019a, b; Lee & Dressman, 2018), L2 anxiety and willingness to communicate (Lee, 2019b; Lee & Dressman, 2018). On the other hand, this result may be related to the participants' purposes for engaging in IDDEL activities. Participants stated in their E-logs that sometimes their main purpose for participating in IDDEL activities was for entertainment or relaxation. Therefore, they tended to choose more meaning-oriented activities, such as listening to English songs or watching English movies and TV shows, rather than form-oriented activities, such as using English learning Apps. However, it should be noted that participating in meaning-oriented activities does not mean that participants focus only on comprehending the meaning of the listening material without paying attention to any linguistic elements. In the process of attempting to understand the meaning of the listening input, participants may pay

attention to unfamiliar, complex, or difficult language elements, such as vocabulary and grammar, to help them achieve better comprehension. For instance, participants' E-logs showed that although they initially participated in some IDLEL activities, such as watching English movies, with the purpose of entertainment or relaxation, they often gained new language knowledge, such as new vocabulary and grammar, during the activity, which resonates with Lai et al.'s (2018) study on informal digital language learning among university foreign language learners.

5.4.2.2 Diversity of SRL Strategy-Use

It is rare in previous research to use participants' SRL strategy use in informal digital language learning activities as an indicator of the quality of their IDLEL engagement. In this study, thematic analysis of participants' E-logs revealed that participants used a total of 54 self-regulated listening strategies in their IDLEL engagement, including both task-level strategies and motivation/affective strategies. Participants were found to use a wider range of task-level strategies than motivation/affective strategies, perhaps indicating their lower levels of self-regulation of a motivational/affective nature.

The types of task-level strategies used by participants covered all phases of self-regulated L2 listening proposed in this study (i.e., Task Representation, Goal Setting & Strategy Planning, Performance, Attribution & Adjustment, and Monitoring & Control). Nevertheless, the number of participants who reported using some of them was low, another possible indication of a lack of self-regulation among the group as a whole. For example, few participants mentioned in their E-logs monitoring their listening comprehension and listening difficulties during IDLEL activities, reflecting on their strategy use and making external or internal attributions for their

listening performance. These findings may suggest a lack of self-regulatory awareness among participants in terms of self-monitoring and self-reflection on strategy use and listening performance. This finding can further be supported by the fact that among the five listening comprehension monitoring strategies included in the codebook, three were not mentioned by the participants. Moreover, as for listening strategies used to aid comprehension in IDLEL activities, using electronic dictionaries and translators, replaying listening materials, and relying on subtitles or scripts were the most popular. These strategies highlight the critical role of technology in assisting learners' L2 listening within the IDLEL context.

Regarding motivation/affect-level self-regulation, participants mentioned motivation and affect awareness, goal clarification, motivation and affect-level strategy employment, and motivation and affect adjustment. However, there was a lack of monitoring of motivation and affective states during activities, as well as a lack of reflection and attribution regarding the outcomes of motivation and affect management and the strategies used. Interestingly, participants tended to focus on their motivation and affect before listening, but fewer reported doing so after the listening began. One major reason for this may be the lack of continuous monitoring of motivation and affective states during the activity, as mentioned earlier. Once listening begins, participants' attention may mostly be focused on understanding the listening input, causing them to overlook monitoring and regulating their motivation and affect. Moreover, the effectiveness of the motivation and affect regulation strategies that participants employed before listening, as well as their progress in listening comprehension, could also explain why they placed less emphasis on motivation and affect regulation during listening. For instance, if participants believed that the strategies used before the activity successfully boosted

their confidence or reduced listening anxiety, they may not have felt the need to pay as much attention to motivation and affect regulation during listening as they did before it. Additionally, if participants achieved good comprehension in the IDLEL activity without encountering significant listening difficulties that would affect their motivation and affect, they might have reduced their motivation and affect regulation during the activity.

Another interesting finding is that, before starting the listening activity, the majority of participants mentioned being aware of their listening anxiety, whereas only a few participants mentioned being aware of their listening self-efficacy. This result, on the one hand, may indicate that participants lacked the awareness to perceive their listening self-efficacy. Successful experiences can be considered as a key source of self-efficacy (Bandura, 1997). In formal listening contexts, learners can build their self-efficacy through teacher feedback and standardised listening test results. However, in informal listening contexts, participants lack clear external feedback to help them evaluate their listening performance and also lack the awareness to reflect on their own listening performance, as mentioned earlier. Consequently, there may be difficulty for them to develop a clear understanding of their previous informal listening experiences, which may thus impede them from perceiving their listening self-efficacy in informal listening contexts. On the other hand, this result may also indicate that, from the participants' perspective, listening anxiety may have a more profound influence on their L2 listening than self-efficacy. Hence, they tended to give more attention to listening anxiety than self-efficacy.

Moreover, among the 54 identified strategies, participants were found to use an average of 10 to 11 strategies over the four-week study period, with the highest number of strategies used

by a participant being nine and the lowest being one. Due to the lack of existing research data on the diversity of strategy use in informal language learning contexts, the performance of participants in the present study on this measure cannot yet be directly compared with learners in other research contexts. This highlights a potential direction for future research.

5.4.3 Purposes and Benefits of IDLEL Engagement

Apart from exploring the quantity and quality of participants' engagement in IDLEL activities as other studies have done, this study also examined their purposes for and gains from participating in IDLEL activities. Participants often engaged in various IDLEL activities with different goals. The primary reason participants watched English movies, TV series, talk shows, videos on social media platforms, and listened to English songs was for entertainment, corresponding to the entertainment-oriented technological experiences proposed by Lai et al. (2018), where learners use technology to access target language resources to meet entertainment and other daily life needs. In this case, "relaxation," "entertainment," and "personal interest" were the primary incentives participants mentioned in their E-logs for engaging in such IDLEL activities, and the main benefits they reported from these activities were reduced English listening anxiety, followed by vocabulary expansion. This aligns with Lai et al.'s (2018) findings, where foreign language learners at a university in Hong Kong reported that entertainment-oriented technological experiences improved their understanding of colloquial language use in the target language and increased their motivation for language learning.

Additionally, similar to Lai et al. (2018)'s finding that learners participated in instruction-oriented technological experiences primarily to expand their knowledge of the target language, participants in the present study also engaged in IDLEL activities for reasons related

to formal language learning or to improve their language skills and expand their language knowledge. For example, participants reported that the main reasons for listening to radio programmes and audiobooks, as well as engaging in test-related listening practices, were the teacher's recommendations, while the primary purpose of using English learning mobile applications and watching English presentations was to improve their English skills. The main benefit participants reported from these IDLEL activities was vocabulary expansion, which is consistent with the findings of Lai et al. (2018).

5.4.4 IDLEL Engagement Patterns

To better understand the characteristics of participants' IDLEL engagement, cluster analysis was used to examine whether participants with varying L2 listening proficiency, self-regulation, self-efficacy, and listening anxiety differed in their IDLEL engagement. Three IDLEL engagement patterns among participants were identified. Specifically, participants with high listening self-efficacy, strong self-regulation in listening, and low listening anxiety were the ones most willing to spend substantial time on IDLEL activities. On the one hand, high self-regulation, strong motivation, and a positive affective state may have given these participants powerful intrinsic drive for IDLEL engagement, motivating them to invest considerable extracurricular time in IDLEL activities. On the other hand, prolonged exposure to authentic listening materials in the target language may have exerted a positive impact on their listening self-regulation, listening anxiety, and self-efficacy (Lee, 2019b), thereby creating a positive cycle.

Moreover, there were two types of participants who were unwilling to invest much time in IDLEL activities. One type included participants with low self-regulation and self-efficacy but

high listening anxiety. On the one hand, the lack of autonomy in learning, along with negative motivation and affective states, may have resulted in a lack of intrinsic drive to engage in IDLEL activities. Therefore, participating in IDLEL activities may not have been a priority for them during their free time. On the other hand, the lack of sufficient informal listening practice may have made it difficult for these participants to improve their suboptimal self-regulation, self-efficacy, and listening anxiety, thus forming a negative cycle.

Another type of participant unwilling to spend significant time on IDLEL activities included those with moderate self-efficacy and listening anxiety but low self-regulation. Moderate levels of self-efficacy and anxiety may imply that their influence (positive or negative) on participants' IDLEL engagement was limited. Self-regulation, therefore, could be the key reason these participants were unwilling to engage in IDLEL activities extensively. Low self-regulation could suggest that these participants lacked autonomy and self-management for their L2 listening, along with a lack of planning, goal orientation, and the ability to effectively face challenges and adjust strategies during listening. As a result, they may have been unwilling to spend much time on activities they could not effectively control without external assistance from their teachers and peers. This result may thus demonstrate the importance of self-regulation in IDLEL engagement.

In summary, the cluster analysis results discussed above indicate that there may be a close relationship between self-regulation, self-efficacy, listening anxiety, and IDLEL engagement. Therefore, the following sections will provide a detailed and comprehensive exploration of their relationships.

5.5 Discussion for RQ4: The Prediction of IDDEL Engagement and SRL on L2

Listening, Self-Efficacy, and Listening Anxiety

5.5.1 The Predictive Effects of IDDEL Engagement and SRL on L2 Listening

After the 91 participants engaged in the four-week IDDEL study, their listening test scores showed an overall upward trend from the pre-test (T1) to the post-test (T2). However, the degree of improvement varied depending on the participants' T1 listening self-efficacy, T2 listening anxiety, T2 SRL, and the duration of weekly IDDEL activity engagement. Specifically, participants who were more confident in their English listening before the IDDEL study (i.e., higher T1 self-efficacy) showed smaller improvements in listening scores from T1 to T2. This result may be because participants with higher T1 self-efficacy may have already possessed listening proficiency close to the upper limit of their current potential before participating in the IDDEL study, and the observation period may be not long enough to trigger further changes. Therefore, their room for listening improvement might have been limited, even though they devoted comparable effort as other participants over the four weeks. The finding that participants with higher T1 self-efficacy scored higher on both the T1 and T2 listening tests can provide evidence for this explanation and is consistent with other studies that have confirmed a significant positive correlation between self-efficacy and L2 listening (Canaran et al., 2024; Zhang & Xu, 2024). Conversely, participants who were less confident in their English listening before the IDDEL study, namely those with lower T1 self-efficacy, showed greater improvement in listening from T1 to T2. This may be because participants with lower T1 self-efficacy tended to have lower T1 listening test scores, indicating they had more room for improvement, although their

overall listening test scores (i.e., T1 and T2) were lower than those of participants with higher self-efficacy.

Moreover, participants with lower T2 listening anxiety showed smaller improvements in listening from T1 to T2, as they scored higher on both the T1 and T2 listening tests, leading to a limited room for the improvement of their listening. By contrast, participants with higher T2 listening anxiety demonstrated larger improvements in listening from T1 to T2. Their lower T1 listening test scores implied that they had more room for improvement in their listening proficiency. Additionally, the significant negative relationship between listening anxiety and L2 listening reflected in this result is consistent with previous research (Golchi, 2012; Xu & Huang, 2018; Zhang & Xu, 2024).

These results may suggest that the IDLEL study seems to be more beneficial for low-proficiency L2 listeners with low self-efficacy and high listening anxiety, as they can achieve greater listening improvement through IDLEL engagement. For confident, low-anxiety, high-proficiency L2 listeners, however, breaking through their listening proficiency ceiling and significantly improving their listening through IDLEL engagement may require investing more effort and time in IDLEL activities and enhancing their listening self-regulation.

According to the previous discussion, participants' IDLEL activity engagement (i.e., duration) and T2 (post-test) SRL exerted a significant predictive effect on their T1 (pre-test) and T2 (post-test) L2 listening. However, these results can only reflect the immediate impact of IDLEL engagement and SRL on listening. To understand whether IDLEL engagement and SRL can have a long-term effect on listening, a delayed post-test (T3

listening test) was conducted three months after the IDLEL study. The T3 listening test results from 60 participants showed that their listening scores improved from T1 to T3, but the improvement was less than that from T1 to T2. Additionally, IDLEL activity engagement and SRL did not significantly predict participants' listening progress from T1 to T3. This may be because the four-week IDLEL study allowed participants to rapidly improve their English listening in the short term through intensive IDLEL activity engagement and enhanced SRL awareness by recording their IDLEL engagement behaviours and feelings in E-logs. However, after the study ended, the intensity of IDLEL engagement and the SRL awareness enhanced during the study may not have been maintained. Consequently, the rapid improvement in listening achieved during the study may have regressed. Moreover, the finding that IDLEL engagement and SRL were not significant predictors of participants' listening progress from T1 to T3 may also imply that other factors, such as self-efficacy and listening anxiety examined in this study, or factors not addressed in the current research, such as the impact of formal language education, had a more significant influence on participants' listening improvement.

However, it should be noted that the relationship between IDLEL engagement (i.e., duration) and listening is correlational rather than causal (Lee, 2019b). That is, it cannot be definitively concluded that IDLEL engagement is the cause of changes in participants' listening. It could be that those who had higher listening proficiency tended to spend more time engaging in IDLEL activities.

5.5.2 The Predictive Effects of IDLEL Engagement and SRL on Self-Efficacy

Overall, participants' listening self-efficacy showed an upward trend from T1 to T2,

suggesting that they became more confident in listening after engaging in the IDLEL study. Listening anxiety and listening test scores were found to be the predictors of the changes of participants' self-efficacy from T1 to T2, while listening self-regulation, as well as the frequency and duration of IDLEL activity engagement, were all significant predictors of overall listening self-efficacy. Specifically, participants who felt more anxious about English listening before the IDLEL study, namely those with higher T1 listening anxiety, showed a greater increase in listening confidence from T1 to T2; conversely, participants with lower T1 listening anxiety, namely those who were less anxious about listening before the IDLEL study, showed a smaller increase in listening confidence from T1 to T2. Consistent with other studies confirming the negative correlation between listening anxiety and self-efficacy (Canaran et al., 2024), this study found that participants with higher T1 listening anxiety had lower T1 self-efficacy while participants with lower T1 listening anxiety had higher T1 self-efficacy. Participants with lower initial self-efficacy, therefore, had more room for improvement in self-efficacy through IDLEL engagement compared to those with higher self-efficacy. Additionally, participants with lower listening anxiety after the IDLEL study (i.e., lower T2 listening anxiety) tended to have lower confidence in listening before the study (i.e., lower T1 self-efficacy) but higher confidence after the study (i.e., higher T2 self-efficacy). Such notable improvement in self-efficacy can serve as a confirmation of the positive impact of IDLEL engagement on listening self-efficacy.

Regarding the predictive role of L2 listening on self-efficacy, participants with lower T1 listening test scores tended to have lower T1 listening self-efficacy but showed greater progress in self-efficacy from T1 to T2. Their lower initial self-efficacy before the study

(i.e., lower T1 self-efficacy) and their engagement in IDLEL activities over the four weeks may have created conditions for a substantial increase in self-efficacy. By contrast, participants who scored higher on the T1 listening test tended to have higher T1 self-efficacy but showed smaller progress in self-efficacy from T1 to T2 compared to participants with lower T1 listening test scores. On the one hand, higher T1 self-efficacy may imply limited room for improvement in self-efficacy. On the other hand, the four-week IDLEL study may not have significantly helped these participants, who were already confident in their listening, to further increase their listening self-efficacy. This result, therefore, demonstrates the positive influence of IDLEL engagement on the self-efficacy of low-proficiency L2 listeners.

Moreover, participants with higher T2 listening test scores tended to have lower T1 listening self-efficacy but showed greater self-efficacy progress from T1 to T2. Conversely, participants with lower T2 listening test scores tended to have higher T1 self-efficacy but showed smaller improvements in self-efficacy from T1 to T2. This result thus further confirms that IDLEL engagement seems to be more beneficial for participants with lower self-efficacy than for those with higher self-efficacy, as discussed above. Furthermore, participants with lower T2 listening test scores tended to have higher T1 and T2 self-efficacy, suggesting that these participants may have been overly confident in their listening proficiency. By contrast, participants with higher T2 listening test scores tended to have lower T1 and T2 self-efficacy, indicating that these participants may have underestimated their listening proficiency.

5.5.3 The Predictive Effects of IDLEL Engagement and SRL on Listening Anxiety

Results showed that participants' listening anxiety decreased from T1 to T2, indicating that they became less anxious about L2 listening after participating in the IDLEL research. T2 listening self-regulation and T1 listening test scores were found to have a significant negative predictive effect on listening anxiety. Specifically, participants with lower T2 self-regulation or lower T1 listening test scores had higher listening anxiety at both T1 and T2, whereas participants with higher T2 self-regulation or higher T1 listening test scores had lower listening anxiety at both T1 and T2. This finding supports the negative relationship between self-regulation (or metacognitive awareness, one of its key components) and listening anxiety, as well as between L2 listening and listening anxiety, as found in previous studies (Golchi, 2012; Xu & Huang, 2018; Zhang & Xu, 2024).

5.6 Discussion for RQ5: The Moderating Effect of SRL on the IDLEL Engagement and L2 Listening, Anxiety, and Self-Efficacy

The fifth question of this study explored the potential moderating role of SRL in the relationship between IDLEL engagement (i.e., duration, frequency, diversity, and strategy use) and listening, self-efficacy, as well as listening anxiety. Based on the LMM results discussed earlier, IDLEL engagement (i.e., duration and frequency) significantly predicted participants' L2 listening and listening self-efficacy, but not listening anxiety. Therefore, the present study examined the moderating effect of SRL on the relationship between IDLEL engagement and listening, as well as between IDLEL engagement and self-efficacy.

Surprisingly, the results showed that SRL did not have significant moderating effects on the relationship between IDLEL engagement and listening, nor between IDLEL engagement and self-efficacy. In other words, SRL did not significantly influence the relationship between

IDLEL engagement and listening, nor between IDLEL engagement and self-efficacy. This result firstly indicates that the impact of IDLEL on listening and self-efficacy can be independent, as IDLEL engagement itself, by providing ample language input opportunities for learners, can directly and effectively enhance their L2 listening and self-efficacy without relying on the moderating effect of SRL. For example, participants who frequently engaged with authentic English materials in IDLEL activities may have gradually become accustomed to the target language environment and accumulated extensive listening experience, thereby naturally improving their L2 listening, and boosting their listening self-efficacy. Therefore, this result may suggest that the effectiveness of IDLEL engagement can be stable and universally applicable to learners. Even learners with lower SRL levels may also experience improvements in their L2 listening and self-efficacy through engagement in IDLEL activities.

Secondly, the fact that SRL did not strengthen or weaken the impact of IDLEL engagement on listening and self-efficacy may indicate that self-regulation, as an intrinsic learning driver, might have influenced listening and self-efficacy independently of IDLEL engagement. This implies that within any learning context, SRL may affect learners' listening and self-efficacy through certain mechanisms. Therefore, for learners with high self-regulation, even if they do not engage in IDLEL activities, they may still achieve improvement in their L2 listening and self-efficacy.

Taken together, the finding that SRL did not exert moderating effects on the relationship between IDLEL engagement and listening, nor between IDLEL engagement and self-efficacy, suggests, on the one hand, the stable and universally applicable benefits of IDLEL engagement for L2 learners' listening and self-efficacy. On the other hand, it indicates that SRL, as an

intrinsic learning driver, can exert an influence on listening and self-efficacy that is independent of specific learning contexts.

At this point, this chapter has completed the discussion of findings for all five research questions. The subsequent chapter will provide a comprehensive summary of the entire study and discuss the implications of the research results.

Chapter 6. CONCLUSION

6.1 Introduction

This chapter will begin with a brief review of the whole study, followed by a summary of the main findings for each research question. It will also outline the potential contributions of this study from the empirical, theoretical, methodological, and pedagogical perspectives. Finally, the chapter will summarise the limitations of the current study and offer potential implications for future research.

6.2 Summary of the Study

This study firstly constructed and validated a self-regulated L2 listening model and then used it to further explore the joint predictive effect of self-regulated learning (SRL), listening self-efficacy and anxiety on L2 listening comprehension. It also used the new model to investigate the role of informal digital learning of English listening (IDLEL) activity engagement (i.e. frequency, diversity, duration, and strategy-use), as well as self-regulation in the development of Chinese undergraduates' L2 listening comprehension, listening self-efficacy, and listening anxiety within the IDLEL context.

To validate the hypothesised structure of the self-regulated L2 listening model (RQ1) and the questionnaire specifically developed for this study, 582 EFL learners (aged 17 to 21) from five universities in China were invited to complete the Self-Regulated L2 Listening Questionnaire (SRLLQ), developed based on the hypothesised model structure; the Listening Self-Efficacy Questionnaire (LSEQ), designed to measure learners' listening self-efficacy; and the L2 Listening Anxiety Scale (LLAS), used to assess listening anxiety. The participants came from five different majors, with Chinese as their first language and

English as their second language. Additionally, to address RQ2 to RQ5, 130 English major students (aged 17 to 21) from two of the five universities in China were invited to participate in a four-week IDLEL observational study. Before the study commenced, participants completed a pretest consisting of an English listening comprehension test and three questionnaires (i.e., SRLLQ, LSEQ, and LLAS). During the four-week IDLEL study, participants engaged in IDLEL activities without intervention and completed four weekly E-logs to record their IDLEL experiences. Immediately after the IDLEL study, participants were required to complete a second listening test and the same three questionnaires as the post-test. Three months after the post-test, participants completed the delayed post listening comprehension test.

6.3 Summary of the Findings

6.3.1 Validation of the Hypothesised Self-Regulated L2 Listening Model and the Joint Predictive Mechanisms

The overarching aim of the first two research questions was: 1) to validate the five-phase construct of a newly proposed self-regulated L2 listening model (RQ1), and 2) to use the new model to further explore the joint predictive mechanisms of SRL, listening self-efficacy, and listening anxiety on L2 listening (RQ2). The newly proposed self-regulated L2 listening model hypothesised that self-regulated L2 listening comprises five phases: *Task Representation, Goal Setting & Strategy Planning, Performance, Attribution & Adjustment*, and *Monitoring & Control*. Each phase involved both cognitive regulation as well as motivational and affective regulation. The present study ultimately confirmed the dual-level, five-construct structure of the self-regulated L2 listening model,

thereby addressing RQ1.

Regarding RQ2, the current study confirmed the two hypothesised joint predictive mechanisms of SRL, self-efficacy, and listening anxiety on L2 listening. Grounded in Social Cognitive Theory (SCT), the first hypothesis depicted both the direct predictive effect of self-efficacy on listening as well as its indirect effect with SRL and listening anxiety being the mediators. The second hypothesis focused on both the direct predictive effect of SRL on listening and its indirect effect via self-efficacy and listening anxiety as mediators. The two models complement each other, together depicting a positive cycle that supports L2 learners' listening development: effective SRL can boost learners' self-efficacy in managing current and/or future listening tasks (Hypothesis 2); in turn, boosted self-efficacy may promote more efficient self-regulation of listening and suppress listening anxiety in current and/or future tasks, ultimately enhancing learners' listening performance (Hypothesis 1).

Moreover, the distinct roles of SRL in the two joint prediction mechanisms were confirmed. Specifically, SRL can serve not only as a crucial tool for transforming learners' beliefs in their ability to complete listening tasks into actual listening performance by mediating the relationship between self-efficacy and listening, but can also act as a motivation booster, indirectly influencing listening by enhancing learners' self-efficacy. Thus, it can be inferred that the impact of SRL on listening is not only immediate but also potentially long-term. Overall, the validation of these two hypothetical mechanisms illustrates the comprehensive relationship among motivational factors, affective responses, cognitive factors, and listening comprehension, while highlighting the critical role of SRL

in the joint predictive mechanisms for listening.

6.3.2 The Characteristics of IDLEL Engagement

RQ3 explored the characteristics of participants' IDLEL engagement in terms of the quantity of IDLEL engagement (i.e., frequency and duration of IDLEL engagement), and the quality of IDLEL engagement (i.e., diversity of activities engaged and strategy-use).

6.3.2.1 The Quantity of IDLEL Engagement: Frequency and Duration

Participants engaged in IDLEL activities 10-11 times over the four-week IDLEL study, with an average weekly engagement of two to three times. Additionally, only 4.4% participants in the present study spent more than two hours daily on IDLEL activities, compared with 70.3% who spent less than one hour per day on IDLEL activities, indicating that most participants in the current research assigned less importance to IDLEL activities, appearing to lack the motivation and initiative to make full use of their out-of-class time to enhance their English listening skills.

6.3.2.2 The Quality of IDLEL Engagement: Diversity of Activities Engaged

Participants engaged in a total of ten types of IDLEL activities over the four-week study, with an average of one to two types of IDLEL activities per week. Among the ten types of activities, the one chosen by the most participants was listening to English songs, while the activity with the fewest participants was listening to English audiobooks.

Additionally, the research also found that participants engaged more frequently in meaning-oriented informal learning activities (e.g., watching English movies and TV shows) which emphasise communication and meaning within authentic informal digital environments, than in form-oriented ones (e.g., using English learning apps and practicing

listening for exams), which focus more on the target language system, formal elements of the language, and linguistic accuracy. This result may indicate that the diversity of participants' IDLEL activities was somewhat lacking, as their engagement in IDLEL activities did not achieve a balanced focus on form and meaning.

6.3.2.3 The Quality of IDLEL Engagement: Diversity of Strategy-Use

The results of the thematic analysis of participants' E-logs revealed that participants used a total of 54 self-regulated listening strategies in their IDLEL engagement, including both task-level strategies and motivation/affective strategies. However, participants used a wider range of task-level strategies than motivation/affective strategies, perhaps indicating their lower levels of self-regulation of a motivational/affective nature. Additionally, among the 54 identified strategies, participants were found to use an average of 10 to 11 strategies over the four-week IDLEL study, with the highest number of strategies used by a participant being nine and the lowest being one.

At the task level, interestingly, only a small number of participants mentioned monitoring their listening comprehension and difficulties, reflecting on their strategy use, or making external or internal attributions for their listening performance during IDLEL activities. This may indicate a lack of self-regulatory awareness among participants in terms of self-monitoring and reflecting on their strategy use and listening performance. Additionally, the most popular listening strategies among participants included using electronic dictionaries and translators, replaying listening materials, and relying on subtitles or scripts. These commonly used strategies highlight the critical role of technology in supporting learners' L2 listening in the IDLEL context.

As for motivation/affect-level self-regulation, participants were found lacking in the monitoring of their motivational and affective states during activities, as well as reflection and attribution concerning the outcomes of motivation and affect management and the strategies used. Moreover, most participants reported being aware of their listening anxiety before starting listening activities, whereas only a few mentioned awareness of their listening self-efficacy. This result may suggest that participants lack awareness of their listening self-efficacy. Alternatively, it may indicate that, from the participants' perspective, listening anxiety can have a more significant impact on their L2 listening performance than self-efficacy.

6.3.2.4 Profiles of IDLEL Engagement Patterns

To better understand the characteristics of participants' IDLEL engagement, cluster analysis was deployed to reveal participants' different IDLEL engagement patterns. Three distinct IDLEL engagement patterns were then identified. The first group consisted of participants with high listening self-efficacy, strong self-regulation in listening, and low listening anxiety, who were the most willing to invest significant time in IDLEL activities. The second group included participants who were reluctant to spend much time on IDLEL activities, characterised by low self-regulation and self-efficacy but high listening anxiety. The third group of participants who also displayed limited IDLEL engagement comprised those with moderate levels of listening self-efficacy and anxiety but low self-regulation. These cluster analysis results suggest a close relationship between self-regulation, self-efficacy, listening anxiety, and IDLEL engagement. Accordingly, RQ4 investigated the relationships among these factors in greater depth.

6.3.3 The Predictive Effects of IDLEL Engagement and SRL on L2 Listening,

Self-Efficacy, and Listening Anxiety

RQ4 investigated the predictive ability of participants' IDLEL engagement and listening self-regulation for their L2 listening proficiency, self-efficacy, and listening anxiety.

6.3.3.1 The Predictive Effects of IDLEL Engagement and SRL on L2 Listening

In terms of L2 listening, participants ($N = 91$) who took part in the four-week IDLEL study demonstrated an overall improvement in their listening test scores from the pretest to the post-test. Moreover, the present study showed that participants' SRL (post-test) and the duration of their IDLEL activity engagement significantly and positively predicted their listening improvement from pre-test to post-test. Specifically, participants who dedicated more time (i.e., higher duration) to IDLEL activities weekly or exhibited higher listening self-regulation on the post-test showed greater progress in listening from the pretest to the post-test; conversely, those who spent less time (i.e., lower duration) on IDLEL activities or had lower listening self-regulation on the post-test experienced smaller improvements in their listening proficiency from the pretest to the post-test. These results thus highlight the importance of the quality of IDLEL engagement, particularly the duration of IDLEL engagement, and listening self-regulation for learners' L2 listening development.

To investigate whether IDLEL engagement can have a long-term effect on learners' listening development, a delayed post-test was conducted three months after the IDLEL study. The current research found that participants' ($N = 60$) listening scores improved from the pretest to the delayed post-test, but the improvement was smaller compared to the

progress from the pretest to the post-test. This result may indicate that the intensive effect of the IDLEL study on listening development was primarily short-term. If participants did not maintain the same level of IDLEL engagement after the study, the long-term impact of the IDLEL study could be limited. Additionally, the finding that IDLEL engagement did not significantly predict participants' listening progress from the pretest to the delayed post-test further suggests that the long-term influence of IDLEL engagement on listening development could be limited.

6.3.3.2 The Predictive Effects of IDLEL Engagement and SRL on Self-Efficacy

Participants' ($N = 91$) listening self-efficacy showed an overall upward trend from the pre-test to the post-test, indicating that they became more confident in their listening abilities after participating in the IDLEL study. Moreover, the study revealed that participants' SRL (post-test) and the duration of their IDLEL activity engagement significantly and positively predicted their listening self-efficacy at both the pre-test and post-test, while the frequency of IDLEL engagement had a significant negative predictive effect on their listening self-efficacy at both time points. Specifically, participants with higher listening self-regulation at the post-test were more likely to have higher listening self-efficacy at both pretest and post-test, highlighting the significant influence of self-regulation on self-efficacy. Regarding the predictive effect of IDLEL engagement on self-efficacy, the findings showed that participants who spent more time (i.e., higher duration) on IDLEL activities generally exhibited higher listening self-efficacy on both pretest and post-test, whereas those who engaged in IDLEL activities more frequently (i.e., higher frequency) tended to have lower listening self-efficacy on both tests. These results

suggest that the duration of IDLEL engagement could be more beneficial than the frequency of engagement for enhancing learners' listening self-efficacy.

6.3.3.3 The Predictive Effects of IDLEL Engagement and SRL on Listening Anxiety

Participants' ($N = 91$) listening anxiety decreased from the pretest to the post-test, indicating that their anxiety about L2 listening was alleviated after participating in the IDLEL study. Additionally, participants with lower self-regulation at the post-test exhibited higher listening anxiety at both pretest and post-test, whereas those with higher self-regulation at the post-test showed lower listening anxiety on both tests. These findings thus demonstrate the negative relationship between self-regulation and listening anxiety. Surprisingly, IDLEL engagement was not found to have a predictive effect on listening anxiety.

6.3.4 The Moderating Role of SRL Between IDLEL Engagement and L2 Listening, Self-efficacy, and Listening Anxiety

RQ5 investigated the potential moderating effect of SRL on the relationship between IDLEL engagement and L2 listening, self-efficacy, as well as listening anxiety. According to the LMM results discussed earlier, the duration of IDLEL engagement significantly predicted participants' listening improvement from the pretest to the post-test, while both duration and frequency of IDLEL engagement significantly predicted participants' overall listening self-efficacy (i.e., pre- and post-test). However, IDLEL engagement did not have a significant predictive effect on listening anxiety. Consequently, this study specifically examined whether SRL moderated the relationship between IDLEL engagement and listening, as well as between IDLEL engagement and self-efficacy. The results revealed that

SRL did not significantly moderate either of these relationships. In other words, SRL showed no significant effect on the link between IDDEL engagement and listening or between IDDEL engagement and self-efficacy. These findings, on the one hand, imply that the benefits of IDDEL engagement can be consistent and widely applicable across learners. Even those with lower SRL levels may experience gains in their L2 listening and self-efficacy through engagement in IDDEL activities. On the other hand, the findings suggest that SRL, as an intrinsic driver of learning, may influence listening and self-efficacy independently of specific learning contexts. Consequently, learners with strong self-regulation skills may still achieve gains in their L2 listening and self-efficacy even without engaging in IDDEL activities.

6.4 Contributions of the Study

6.4.1 Empirical Contributions

First, this study contributes empirical evidence to the limited research on L2 learners' SRL within informal learning contexts. As mentioned earlier, most existing studies primarily focus on the impact of self-regulation in formal learning contexts for L2 listening (Fatemi et al., 2014; Nasrollahi-Mouziraji & Birjandi, 2016; Xu & Luo, 2024). A few studies have investigated SRL in informal learning contexts, but their generalisability may be limited due to small sample sizes (Zeng & Goh, 2018) or the specific applicability of the research instrument (Yabukoshi, 2024). Compared to these studies, this research had a more appropriate sample size and a more generally applicable research instruments, which may lead to greater generalisability of the findings.

Secondly, this study fills a gap in the limited research that has comprehensively

explored the joint predictive effects of metacognitive factors (e.g., SRL), motivational factors (e.g., self-efficacy), and affective factors (e.g., listening anxiety) on L2 listening within a single research framework. As one of the few studies attempting to investigate this area, Zhang and Xu (2024) focused on the joint predictive effects of listening anxiety, self-efficacy, and metacognitive awareness on L2 listening. Although metacognitive awareness is a key component of SRL (Efklides, 2011), it is not synonymous with SRL, which also depends on motivational and affective factors. Therefore, this study is the first to explicitly explore the joint predictive mechanisms of SRL, self-efficacy, and listening anxiety on L2 listening. Moreover, unlike Zhang and Xu (2024), who focused solely on the mediating role of metacognitive factors between motivational and listening as well as between affective factors and listening, this study also investigated the predictive effect of metacognitive factors on listening with motivational and emotional factors being mediators. As such, this research provides novel empirical evidence on the joint predictive mechanisms of these three factors on listening.

Thirdly, there has been almost no research focusing on the L2 listening development of Chinese university EFL learners in informal digital learning of English (IDLE) contexts, and this study fills that gap. On the one hand, research exploring the L2 listening development of Chinese university EFL learners in informal contexts is very limited (Zeng & Goh, 2018), so this study can enrich and expand the research in this area. On the other hand, listening is one of the language skills that has been neglected in IDLE research. As one of the few empirical studies focusing on the relationship between IDLE engagement and listening, Sylvén and Sundqvist (2012) confirmed the significant role of IDLE

engagement in the L2 listening development of young learners. However, this study provides empirical evidence for the relationship between IDLE engagement and L2 listening among university-level EFL learners. Additionally, this study is the first to investigate the relationship between L2 learners' IDLE engagement and their listening self-efficacy and listening anxiety, thus filling a corresponding research gap.

Lastly, this study is the first to incorporate the number of IDLEL activities undertaken over a period of time (i.e., frequency) and the use of strategies in IDLEL activities (i.e., strategy-use) as indicators to assess the quantity and quality of learners' IDLEL engagement. Engagement frequency can serve as a key indicator of L2 learners' motivation and interest in L2 learning, as well as reflect their habits in extracurricular L2 learning. Therefore, it can be used alongside the duration of IDLEL activity engagement to evaluate the quantity of learners' IDLEL engagement, reflecting their investment in informal language learning and indicating the depth and intensity of their engagement. Additionally, the diversity of learning strategies used can, to some extent, reflect learners' adaptability to different task demands and learning contexts, as well as whether they have achieved a balance in self-regulation across cognitive, motivational, and affective aspects. Thus, it can be used, along with the diversity of activities participated, as an indicator to assess the quality of learners' IDLEL engagement, revealing the breadth of their informal language learning experiences. No previous research has explored the characteristics of learners' IDLE engagement in these two aspects, so this study fills this gap.

6.4.2 Theoretical Contributions

From a theoretical perspective, firstly, the self-regulated L2 listening model extends

and enriches existing SRL theories. On the one hand, the new model draws on the strengths of classical SRL models by, for instance, integrating key phases from other models and reflecting their cyclical nature. On the other hand, this new model also overcomes several of their limitations. For example, while traditional SRL models offer wide applicability, they often lack specificity and provide limited insights into the regulation of motivation and affect. Consequently, this study not only advances current SRL models but also expands the application of SRL theory to the domain of L2 learning, especially L2 listening.

Secondly, the two validated joint predictive mechanisms complement each other, offering a comprehensive depiction of the interplay between motivational, affective, and cognitive factors in predicting L2 listening comprehension. On the one hand, the study revealed distinct predictive pathways of SRL within the joint mechanism: SRL can influence learners' listening not only directly but also indirectly by shaping their motivational and affective factors. The diverse pathways through which SRL influences listening create a cumulative effect, indicating that SRL can not only serve as a tool for learners to effectively manage their current listening tasks; it can also exert a profound influence on their performance in future listening tasks by influencing their motivation and affect. On the other hand, the study also highlighted the different roles that SRL plays in the joint predictive mechanisms for listening: it acts both as a critical tool for transforming learners' belief in their ability to complete listening tasks into actual performance and as a motivation booster. Taken together, the current research further underscores the pivotal role of SRL in L2 learners' listening development and emphasises the necessity of integrating SRL theory into the domain of L2 listening.

Thirdly, this study highlights the significance of IDLE in L2 listening development, thus underscoring the necessity of integrating IDLE theories into the field of L2 listening. Additionally, the significant predictive role of self-regulation in listening development within the IDLEL context emphasises the importance of extending SRL theories to the domain of IDLE. Furthermore, exploring the interplay between L2 listening, self-regulation, self-efficacy, and listening anxiety within the IDLE context can not only enrich and refine the theoretical framework of IDLE but also provide new insights for advancing SLA in informal learning contexts.

6.4.3 Methodological Contributions

Most studies investigating the relationship between learners' IDLE engagement and language proficiency have used questionnaires to gather information about the quantity (i.e., duration) of learners' IDLE engagement (Lee, 2019 a,b), using questions like "On average, how many hours each day did you spend engaging in IDLE activities in the past six months?" (Lee, 2019a, p.770). Additionally, information about the quality (i.e., diversity) of learners' IDLE engagement is primarily obtained through semi-structured interviews (Lee, 2019 a,b; Lee & Dressman, 2018), using questions like "What types of IDLE activities do you engage in?" (Lee, 2019a, p.771). The IDLE engagement data collected through such questionnaires and interviews, however, mainly rely on learners' recall of their IDLE experiences over a long period, which may significantly impact the accuracy and reliability of the data, potentially leading to doubts about the research findings. The research design of the present study addresses these methodological limitations of previous studies. Specifically, this study employed E-logs to collect data on learners' weekly IDLEL

engagement in terms of both quantity (i.e., duration and frequency) and quality (i.e., diversity and strategy-use). Since weekly E-logs allow learners to record their IDLEL engagement experiences in real time, the accuracy of the data and the reliability of the research findings are likely to be improved.

Like the current research, a few studies, such as Sylvén and Sundqvist's (2012) research, used diaries to collect data on learners' IDLE engagement duration. However, their study lasted only one week and limited the types of IDLE activities learners could engage in. On the one hand, the learning behaviours displayed by learners in a short period may not represent their long-term learning habits and patterns, but rather may be a response to external interventions. Therefore, short-term studies may either overestimate or underestimate the actual impact of activities on learning. On the other hand, restricting the types of IDLE activities learners can participate in may lead to choices that do not reflect their real interests and needs, but rather behaviours influenced by expectations for the study or external requirements, thus interfering with the research findings.

This study addresses the limitations in existing research concerning the research duration and external intervention. Firstly, this study, designed to be longitudinal, lasted for four weeks. A longer duration of research can allow for a more comprehensive and detailed observation of learners' consistent and stable learning behaviours in the IDLEL context, capturing changes in their self-regulation, self-efficacy, and listening anxiety. Secondly, this study was observational and did not impose any restrictions on learners' IDLEL engagement. This naturalistic observation approach can enable a more authentic reflection of learners' behaviours and responses, making the results more generalisable to real

informal learning contexts and thus improving the external validity of the research findings.

Another important methodological contribution of the present study is the development of new research instruments, including the *Self-Regulated L2 Listening Questionnaire (SRLLQ)* for measuring learners' self-regulated L2 listening, the *Listening Self-Efficacy Questionnaire (LSEQ)* for measuring listening self-efficacy, and the *L2 Listening Anxiety Scale (LLAS)* for measuring listening anxiety. The three questionnaires demonstrated reliability ranging from acceptable to good, indicating their potential applicability in other research and learning contexts.

6.4.4 Pedagogical Contributions and Implications

Firstly, this study can be particularly important for Chinese university EFL learners as it not only reveals that these learners face the challenge of limited support for L2 listening development within the classroom context but also suggests an effective solution to this dilemma, that is, engaging in IDLEL activities during their extracurricular time. Therefore, this research should draw the attention of Chinese educational authorities, universities, and L2 teachers. On the one hand, all stakeholders should make efforts to address the various issues hindering Chinese EFL learners' L2 listening development in the classroom, such as limited listening instruction time, a lack of diverse teaching methods, exam-oriented teaching content, and the absence of authentic listening materials. On the other hand, attention should also be paid to the positive impact of informal language learning on learners' L2 listening development, providing support for their listening practice in IDLE contexts. For example, universities can not only offer more diverse IDLEL learning resources but also provide relevant training for teachers, guiding them in integrating and

developing IDLEL resource packages that complement classroom teaching. Teachers can combine classroom tasks with extracurricular autonomous listening activities, helping and encouraging students to develop regular IDLEL learning habits, guiding them in selecting IDLEL activities that are suitable for their proficiency level and conducive to long-term engagement, ensuring that students can continue participating in informal listening practice. In addition, teachers should focus on fostering students' SRL abilities, as well as their capacity to regulate self-efficacy and listening anxiety in class, and guide students to transfer these skills into informal language learning to support the development of their L2 listening abilities.

Secondly, the self-regulated L2 listening model constructed in the present study can provide a theoretical foundation and guidance for both teachers' and learners' SRL educational practices. Teachers can design phase-specific SRL teaching activities based on the five-phase structure of the model. For example, at the *Task Representation* phase, teachers can guide students to understand the specific requirements of the listening task and help them clarify learning objectives. At the *Goal Setting & Strategy Planning* phase, teachers can guide students how to set specific, achievable learning goals and strategies. Additionally, teachers can develop targeted SRL teaching resources based on the core objectives of each phase. For instance, one of the core objectives of the *Attribution & Adjustment* phase is to reflect on learners' listening performance and motivation/affect-regulation, attributing the outcomes to either internal (e.g., strategy-use, effort) or external factors (e.g., speed of the listening material). For SRL teaching in this phase, teachers can design self-evaluate questionnaires or checklists to help students reflect

on and attribute their listening performance and strategy use.

Thirdly, the SRLLQ, LSEQ, and LLAS developed in the current study can provide data support for personalised interventions. Specifically, teachers can use these tools to regularly assess students' SRL abilities, self-efficacy and listening anxiety levels, and based on the assessment results, implement data-driven personalised interventions. For example, for students with high listening anxiety, teachers could adopt a phased approach to gradually reduce the difficulty of tasks. Moreover, these tools can serve as self-assessment instruments for students, helping them better understand their cognitive, motivational and affective status and make targeted improvements to their learning strategies.

Fourthly, the present study revealed the interplay between self-efficacy, listening anxiety, and self-regulation, as well as their joint impact on learners' L2 listening. This result suggests that teachers need to actively coordinate these three factors through systematic design in L2 listening instruction. Teachers should avoid the limitations of isolated interventions (e.g., simply enhancing self-efficacy may lead to a rebound of anxiety due to a lack of effective strategies). They should also avoid using fixed teaching strategies and instead adjust their focus flexibly based on students' feedback and learning progress. For example, when an imbalance is detected (e.g., an increase in anxiety), teachers should prioritise targeted interventions (e.g., helping students alleviate anxiety). Furthermore, this study showed that self-regulation could not only serve as a tool to transform learners' beliefs into actual performance, impacting their immediate listening performance, but also act as a motivation booster, influencing their long-term listening development. Therefore, teachers should place particular emphasis on cultivating learners' SRL abilities in L2

listening instruction, and the previously mentioned self-regulated L2 listening model and SRLLQ questionnaire can both serve as effective tools for teachers to develop learners' self-regulated L2 listening abilities.

6.5 Limitations and Future Research

The first limitation stemmed from issues with missing data. There are two likely reasons for this. First, this was a longitudinal study designed to examine the same group of participants over different time points. Similar to many other longitudinal studies, sampling attrition occurred, and the loss of participants intensified as the study duration increased (Ahern & Le Brocque, 2005). For instance, among the 130 participants who completed valid pre- and post-listening tests, only 83 completed the delayed post-test three months after the post-test, as 31 of them from University A were reassigned to different classes that were not involved in the study in the new semester. 16 of them from University B did not provide an explanation for their absence from the third listening test. Second, this was designed as an observational study, which requires the researcher to refrain from intervening or making specific requests to ensure the data can reflect participants' natural states and behaviours outside the research context. However, this may lead to selective engagement in tasks based on participants' preferences and habits, resulting in missing (Ahern & Le Brocque, 2005). For example, among the 130 participants who completed valid pre- and post-listening tests and questionnaires, only 91 completed at least one valid weekly IDLEL E-log. Therefore, when conducting future research, researchers should be aware of various potential risks such as these when designing studies and plan appropriate measures to maximise data retention

Secondly, as a confirmatory study, this research validated two hypothesised predictive mechanisms for listening based on the framework of Social Cognitive Theory. The two hypothesised mechanisms used self-efficacy and self-regulated learning (SRL) as independent variables respectively, with anxiety being the mediator. Since no adequate theoretical support was found for the hypothesis that anxiety, as an independent variable, predicts L2 listening through self-efficacy and SRL, this relationship was not explored or validated in this study. This, however, could potentially serve as a research direction for future exploratory studies. Additionally, SRL, self-efficacy, and listening anxiety together accounted for only about 23% of the variance in listening comprehension across the two hypothesised predictive mechanisms in the present study. This indicates that other factors may also play a role in influencing listening comprehension. Future research could explore incorporating additional variables not addressed in this study, such as vocabulary and working memory capacity, into the joint predictive mechanisms for further understanding.

Thirdly, as mentioned in *Section 3.2.1*, one limitation of this study is the uneven gender distribution in the sample, which also reflected the reality of language classrooms in many Chinese universities, where there are usually more females than males (Yu, 2015). Many studies have explored the role of gender in various aspects of L2 learning but the results are mixed. Some researchers found significant gender difference in either English listening confidence (Xu & Qiu, 2022), reading anxiety (Shi & Liu, 2006), language learning strategy use (Rao, 2005) or language proficiency (Gu, 2002). However, some researchers reported no significant gender differences in variables, such as learners' self-reported language proficiency (Liu et al., 2016), SRL strategy use (Ting & Chao, 2013)

and foreign language classroom anxiety (Shi & Liu, 2006).

Despite ongoing debates about the role of gender in various aspects of L2 learning, the gender imbalance in the sample remains a limitation of this study. Future research in similar areas could consider addressing this limitation.

Fourthly, in Section 2.2.3, the researcher discussed how multimodal input can, to some extent, reduce the influence of vocabulary knowledge and working memory capacity on listening comprehension. However, the listening comprehension test in the present study did not incorporate multimodal materials to echo this point, which constitutes a limitation, although questionnaire items did refer to multimodality. Although the listening test used in this study was characterised by the use of authentic listening materials drawn from real-life contexts, such as movie dialogues and presentations, only the audio components were used, and the inherently multimodal nature of these materials was not preserved. This, to some degree, weakened the authenticity of the listening input. Future studies of a similar nature may consider incorporating authentic multimodal listening materials in the design of listening tests.

Fifthly, when a study addresses multiple research questions around a topic and employs various analytical methods, connecting different findings for further analysis may yield more valuable insights. In this study, it was not possible to link the results of the cluster analysis, which identified participants' patterns of IDLEL engagement, with the LMM findings that revealed the predictive roles of IDLEL engagement and SRL in learners' listening performance, self-efficacy, and listening anxiety. This is because the two sets of results could not be directly compared or integrated because of differences in the data used:

the cluster analysis was conducted using only pre-test data, while the LMM analysis used pre-test, post-test, and delayed post-test data. Such inconsistency in data sources may explain the different results that emerged from the cluster analysis and the LMM analysis. That such findings could not be meaningfully connected is acknowledged as a limitation. In future research, researchers are encouraged to consider designing studies that enable more integrated analyses across different types of sources of data and findings.

6.6 Conclusion

As the final section of the thesis, this chapter provided a summary of the main findings of the study, outlining its contributions from the empirical, theoretical, methodological, and pedagogical perspectives. It also highlighted the main limitations of the study while offering suggestions for future research.

The present thesis now concludes the presentation and discussion of the research. Motivated by the need to address the insufficient in-class L2 listening learning support faced by Chinese university EFL learners, this study was conducted to validate the viability of informal listening activities beyond the classroom as a significant approach to enhancing their L2 listening proficiency. By investigating the predictive effects of learners' IDDEL engagement and SRL on their L2 listening, listening self-efficacy, and listening anxiety, the study confirms the rationale for incorporating IDDEL as a crucial supplement for Chinese EFL learners. Moreover, the research validates the newly proposed self-regulated L2 listening model and two joint predictive mechanisms of self-regulation, self-efficacy, and listening anxiety on L2 listening, thereby enriching and expanding the SRL theory and L2 listening theory. Taken together, this study has achieved all its predefined objectives.

References

Ahern, K., & Le Brocq, R. (2005). Methodological issues in the effects of attrition: Simple solutions for social scientists. *Field Methods*, 17(1), 53-69.
<https://doi.org/10.1177/1525822X04271006>

Alshahrani, T. M., & Almanea, M. M. (2023). Does the use of web-based extensive listening activities affect foreign language listening anxiety and foreign language anxiety?. *Journal of Language Teaching and Research*, 14(6), 1710-1721.
<https://doi.org/10.17507/jltr.1406.30>

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). Freeman.

Asamoah, M. K. (2014). Re-examination of the limitations associated with correlational research. *Journal of Educational Research and Reviews*, 2(4), 45-52.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36(3), 189-208. [https://doi.org/10.1016/S0021-9924\(03\)00019-4](https://doi.org/10.1016/S0021-9924(03)00019-4)

Bailey, K. M., & Ochsner, R. (1983). A methodological review of the diary studies: Windmill tilting or social science. In K.M. Bailey, M.H. Long, & S. Peck (eds.), *Second language acquisition studies* (pp. 188-198). Newbury House.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bandura, A. (1988). Self-efficacy conception of anxiety. *Anxiety Research*, 1, 77-98.
<http://dx.doi.org/10.1080/10615808808248222>

Bandura, A. (1990). Perceived self-efficacy in the exercise of personal agency. *Journal of Applied Sport Psychology*, 2(2), 397-427.
<https://doi.org/10.1080/10413209008406426>

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, 50(2), 248-287.
[https://doi.org/10.1016/0749-5978\(91\)90022-L](https://doi.org/10.1016/0749-5978(91)90022-L)

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.
https://doi.org/10.1207/s15326985ep2802_3

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan, & F. Pajares (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). IAP.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278. <https://doi.org/10.1016/j.jml.2012.11.001>

Berkemeyer, V. C. (1989). Qualitative analysis of immediate recall protocol data: Some classroom implications. *Teaching German*, 22(2), 131-137.
<https://doi.org/10.2307/3530185>

Bhattacherjee, A. (2012). *Social science research: Principles, methods, and practices* (2nd ed.). Anol Bhattacherjee.

Bloomfield, A., Wayland, S. C., Rhoades, E., Blodgett, A., Linck, J., & Ross, S. (2010). *What makes listening difficult? Factors affecting second language listening*

comprehension. University of Maryland Center for Advanced Study of Language.

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review*, 18(3), 199-210. <https://doi.org/10.1007/s10648-006-9013-4>

Boekaerts, M., & Minnaert, A. (1999). Self-regulation with respect to informal learning. *International Journal of Educational Research*, 31(6), 533-544. [https://doi.org/10.1016/S0883-0355\(99\)00020-8](https://doi.org/10.1016/S0883-0355(99)00020-8)

Boileau, T. (2018). Informal learning. In R. E. West (Ed.), *Foundations of learning and instructional design technology: Historical roots and current trends*. EdTech Books. Retrieved from https://edtechbooks.org/lidtfoundations/informal_learning

Borkowski, J. G. (1996). Metacognition: Theory or chapter heading? *Learning and Individual Differences*, 8(4), 391-402. [https://doi.org/10.1016/S1041-6080\(96\)90025-4](https://doi.org/10.1016/S1041-6080(96)90025-4)

Bostancioğlu, A., & Handley, Z. (2018). Developing and validating a questionnaire for evaluating the EFL “Total PACKage”: Technological Pedagogical Content Knowledge (TPACK) for English as a Foreign Language (EFL). *Computer Assisted Language Learning*, 31(5-6), 572-598. <https://doi.org/10.1080/09588221.2017.1422524>

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

Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing*, 19(4), 369-394. <https://doi.org/10.1191/0265532202lt236oa>

Brunfaut, T., & Revesz, A. (2015). The role of task and listener characteristics in second language listening. *TESOL Quarterly*, 49(1), 141-168. <https://doi.org/10.1002/tesq.168>

Canaran, Ö., Bayram, İ., Doğan, M., & Baturay, M. H. (2024). Causal relationship among the sources of anxiety, self-efficacy, and proficiency in L2 listening. *International Journal of Listening*, 38(2), 105-117. <https://doi.org/10.1080/10904018.2020.1793676>

Chen, H-Y. (2007). *The relationship between EFL learners' self-efficacy beliefs and English performance* [Doctoral dissertation, Florida State University]. FSU Digital Library. http://purl.flvc.org/fsu/fd/FSU_migr_etd-3846

Cohen, A. D. (1998). *Strategies in learning and using a second language*. Longman.

Creswell, J. W. (1999). Mixed-method research: Introduction and application. In G. J. Cizek (Ed.), *Handbook of educational policy* (pp. 455-472). Academic Press.

Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint of the listener. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 123-166). OUP.

Dammak, A. (2015). Research paradigms: Methodologies and compatible methods. *Veritas*, 6(2), 1-5.

Dash, G., & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change*, 173, 121092. <https://doi.org/10.1016/j.techfore.2021.121092>

Dehnad, A., Nasser, H., & Hosseini, A. F. (2014). A comparison between three-and

four-option multiple choice questions. *Procedia-Social and Behavioral Sciences*, 98, 398-403. <https://doi.org/10.1016/j.sbspro.2014.03.432>

DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). Sage

Doody, O., & Doody, C. M. (2015). Conducting a pilot study: Case study of a novice researcher. *British Journal of Nursing*, 24(21), 1074-1078. <https://doi.org/10.12968/bjon.2015.24.21.1074>

Dressman, M. (2020). Introduction. In M. Dressman & R. Sadler (Eds.), *The handbook of informal language learning* (pp. 1-12). Wiley-Blackwell.

Du, G., & Man, D. (2022). Person factors and strategic processing in L2 listening comprehension: Examining the role of vocabulary size, metacognitive knowledge, self-efficacy, and strategy use. *System*, 107, 102801. <https://doi.org/10.1016/j.system.2022.102801>

Du, G., & Man, D. (2023). Mediating effect of metacognition on the relationship between listening self-efficacy and L2 listening comprehension. *The Asia-Pacific Education Researcher*, 32(5), 655-664. <https://doi.org/10.1007/s40299-022-00684-z>

Du, G., Hasim, Z., & Chew, F. P. (2022). Contribution of English aural vocabulary size levels to L2 listening comprehension. *International Review of Applied Linguistics in Language Teaching*, 60(4), 937-956. <https://doi.org/10.1515/iral-2020-0004>

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, 46(1), 6-25. <https://doi.org/10.1080/00461520.2011.538645>

Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *The Modern Language Journal*, 89(2), 206-220. <https://doi.org/10.1111/j.1540-4781.2005.00275.x>

Engin, M. (2011). Research diary: A tool for scaffolding. *International Journal of Qualitative Methods*, 10(3), 296-306. <https://doi.org/10.1177/160940691101000308>

Fatemi, M. A., Alishahi, M., Khorasani, M. N., & Seifi, M. (2014). The relationship between EFL learners' self-regulation and their listening comprehension. *Advances in Language and Literary Studies*, 5(4), 198-201. <https://doi.org/10.7575/aiac.all.v.5n.4p.198>

Fathi, J., Derakhshan, A., & Torabi, S. (2020). The effect of listening strategy instruction on second language listening anxiety and self-efficacy of Iranian EFL learners. *Sage Open*, 10(2), 2158244020933878. <https://doi.org/10.1177/2158244020933878>

Field, J. (2013). Cognitive validity. In A. Geranpayeh & L. Taylor (Eds.), *Examining listening: Research and practice in assessing second language listening* (pp. 77-151). Cambridge University Press.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>

Franklin, C., & Ballan, M. (2001). Reliability and validity in qualitative research. In B.A. Thyer (Ed.), *The handbook of social work research methods* (pp. 273-392). Sage Publications.

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A

comparative analysis of various methods and student self-report instruments. In S. Christenson, A. L. Reschly, & C. Wylie (Ed.), *Handbook of research on student engagement* (pp. 763-782). Springer.

Fry, J. (1988). Diary studies in classroom SLA research: Problems and prospects. *Jalt Journal*, 9(2), 158-167.

Gao, X. (2007). Has language learning strategy research come to an end? A response to Tseng et al. (2006). *Applied Linguistics*, 28(4), 615-620.
<https://doi.org/10.1093/applin/amm034>

Garcia, T., & Pintrich, P. (1996). Assessing students' motivation and learning strategies in the classroom context: The motivation and strategies in learning questionnaire. In M. Birenbaum & F. J. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes, and prior knowledge* (pp. 319-339). Springer.

Gellert, A. S., Arnbak, E., Wischmann, S., & Elbro, C. (2021). Morphological intervention for students with limited vocabulary knowledge: Short-and long-term transfer effects. *Reading Research Quarterly*, 56(3), 583-601.
<https://doi.org/10.1002/rrq.325>

George, D., & Mallory, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.

Gilakjani, A. P. (2016). The significance of listening comprehension in English language teaching. *Theory and Practice in Language Studies*, 6(8), 1670-1677.
<https://doi.org/10.17507/tpls.0608.22>

Gilakjani, A. P., & Ahmadi, M. R. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement. *Journal of Language Teaching and Research*, 2(5), 977-988.
<https://doi.org/10.4304/jltr.2.5.977-988>

Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal*, 39(2), 188-213.
<https://doi.org/10.1177/0033688208092184>

Goh, C. C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1), 55-75. [https://doi.org/10.1016/S0346-251X\(99\)00060-3](https://doi.org/10.1016/S0346-251X(99)00060-3)

Goh, C. C. (2002). Exploring listening comprehension tactics and their interaction patterns. *System*, 30(2), 185-206.
[https://doi.org/10.1016/S0346-251X\(02\)00004-0](https://doi.org/10.1016/S0346-251X(02)00004-0)

Goh, C. C., & Vandergrift, L. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, 8(4), 597-607.
<https://doi.org/10.46743/2160-3715/2003.1870>

Golchi, M. M. (2012). Listening anxiety and its relationship with listening strategy use and listening comprehension among Iranian IELTS learners. *International Journal of English Linguistics*, 2(4), 115-128. <http://dx.doi.org/10.5539/ijel.v2n4p115>

Goleman, D. (2020). *Emotional intelligence: Why it can matter more than IQ*. Bloomsbury Publishing.

Graham, S. (2011). Self-efficacy and academic listening. *Journal of English for Academic*

Purposes, 10(2), 113-117. <https://doi.org/10.1016/j.jeap.2011.04.001>

Graham, S. (2018). Teaching listening strategies. In J. Lontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1-5). John Wiley & Sons.

Graham, S. (2022). Self-efficacy and language learning—what it is and what it isn't. *The Language Learning Journal*, 50(2), 186-207.
<https://doi.org/10.1080/09571736.2022.2045679>

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning*, 58(4), 747-783.
<https://doi.org/10.1111/j.1467-9922.2008.00478.x>

Graham, S., Santos, D., & Vanderplank, R. (2008). Listening comprehension and strategy use: A longitudinal exploration. *System*, 36(1), 52-68.
<https://doi.org/10.1016/j.system.2007.11.001>

Griffiths, R. (1990). Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning*, 40(3), 311-336.
<https://doi.org/10.1111/j.1467-1770.1990.tb00666.x>

Gu, M. (2019). An investigation into the listening anxiety of English major university students and coping strategies. *Language Education*, 7(3), 32-38.

Gu, Y. (2002). Gender, academic major, and vocabulary learning strategies of Chinese EFL learners. *RELC Journal*, 33(1), 35-54.
<https://doi.org/10.1177/003368820203300102>

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Sage.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017). Mirror, mirror on the wall: A comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science*, 45, 616-632.
<https://doi.org/10.1007/s11747-017-0517-x>

Hair, J. F., Ringle, C. M., Gudergan, S. P., Fischer, A., Nitzl, C., & Menictas, C. (2019). Partial least squares structural equation modeling-based discrete choice modeling: An illustration in modeling retailer choice. *Business Research*, 12(1), 115-142.
<https://doi.org/10.1007/s40685-018-0072-4>

Hair, J., & Alamer, A. (2022). Partial Least Squares Structural Equation Modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics*, 1(3), 100027.
<https://doi.org/10.1016/j.rmal.2022.100027>

Hamat, A., & Hassan, H. A. (2019). Use of social media for informal language learning by Malaysian university students. *3L: Language, Linguistics, Literature*, 25(4), 68-83. <https://doi.org/10.17576/3L-2019-2504-05>

Han, S., Zha, Q., & Xie, J. (2023). Understanding Chinese universities' strategic responses to the 'Double First-class' initiative and the potential outcome on higher education diversity: An exploratory study. *Studies in Higher Education*, 48(11), 1624-1638.
<https://doi.org/10.1080/03075079.2023.2206426>

Hayati, A. (2010). The effect of speech rate on listening comprehension of EFL learners. *Creative Education*, 1(2), 107-114. <https://doi.org/10.4236/ce.2010.12016>

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based Nursing*, 18(3), 66-67. <https://doi.org/10.1136/eb-2015-102129>

Healy, M., & Perry, C. (2000). Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. *Qualitative Market Research: An International Journal*, 3(3), 118-126. <https://doi.org/10.1108/13522750010333861>

Hemmati, F., & Ghaderi, E. (2014). The effect of four formats of multiple-choice questions on the listening comprehension of EFL learners. *Procedia-Social and Behavioral Sciences*, 98, 637-644. <https://doi.org/10.1016/j.sbspro.2014.03.462>

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132. <https://doi.org/10.2307/327317>

Hu, G. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development*, 24(4), 290-318. <https://doi.org/10.1080/01434630308666503>

Indrasari, N. (2019). Designing listening tasks for English students. *English Education: Jurnal Tadris Bahasa Inggris*, 12(1), 11-21. <https://doi.org/10.24042/ee-jtbi.v12i1.4423>

In'nami, Y. (2006). The effects of test anxiety on listening test performance. *System*, 34(3), 317-340. <https://doi.org/10.1016/j.system.2006.04.005>

In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing*, 26(2), 219-244. <https://doi.org/10.1177/0265532208101006>

Jin, Y., & Fan, J. (2011). Test for English majors (TEM) in China. *Language Testing*, 28(4), 589-596. <https://doi.org/10.1177/0265532211414852>

Johnson, R. B. (1997). Examining the validity structure of qualitative research. *Education*, 118(2), 282-292.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling. *R package version 0.5-6*. Retrieved from <https://CRAN.R-project.org/package=semTools>

Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social Sciences*, 8(9), 255. <https://doi.org/10.3390/socsci8090255>

Kim, M., Nam, Y., & Crossley, S. A. (2022). Roles of working memory, syllogistic inferencing ability, and linguistic knowledge on second language listening comprehension for passages of different lengths. *Language Testing*, 39(4), 593-617. <https://doi.org/10.1177/026553222110600>

Kivunja, C., & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, 6(5), 26-41. <https://doi.org/10.5430/ijhe.v6n5p26>

Korkmaz, S., Göksüyük, D., & Zararsız, G. Ö. K. M. E. N. (2014). MVN: An R package for assessing multivariate normality. *R JOURNAL*, 6(2), 151-162. <https://journal.r-project.org/archive/2014-2/korkmaz-goksuluk-zararsiz.pdf>.

Krashen, S. D., Terrell, T. D., Ehrman, M. E., & Herzog, M. (1984). A theoretical basis for teaching the receptive skills. *Foreign Language Annals*, 17(4), 261.

https://doi.org/10.1111/j.1944-9720.1984.tb03226.x

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1-26. https://doi.org/10.18637/jss.v082.i13

Lai, C., Hu, X., & Lyu, B. (2018). Understanding the nature of learners' out-of-class language learning experience with technology. *Computer Assisted Language Learning*, 31(1-2), 114-143. https://doi.org/10.1080/09588221.2017.1391293

Lai, C., Zhu, W., & Gong, G. (2015). Understanding the quality of out-of-class English learning. *TESOL Quarterly*, 49(2), 278-308. https://doi.org/10.1002/tesq.171

Lau, K., & Gardner, D. (2019). Disciplinary variations in learning styles and preferences: Implications for the provision of academic English. *System*, 80, 257-268. https://doi.org/10.1016/j.system.2018.12.010

Lee, J. C. K., Yin, H., & Zhang, Z. (2009). Exploring the influence of the classroom environment on students' motivation and self-regulated learning in Hong Kong. *The Asia-Pacific Education Researcher*, 18(2), 219-232. https://doi.org/10.3860/taper.v18i2.1324

Lee, J. S. (2019a). Informal digital learning of English and second language vocabulary outcomes: Can quantity conquer quality? *British Journal of Educational Technology*, 50(2), 767-778. https://doi.org/10.1111/bjet.12599

Lee, J. S. (2019b). Quantity and diversity of informal digital learning of English. *Language Learning & Technology*, 23(1), 114-126. https://doi.org/10125/44675

Lee, J. S., & Dressman, M. (2018). When IDLE hands make an English workshop: Informal digital learning of English and language proficiency. *TESOL Quarterly*, 52(2), 435-445. https://doi.org/10.1002/tesq.422

Lee, J. S., & Lee, K. (2021). The role of informal digital learning of English and L2 motivational self-system in foreign language enjoyment. *British Journal of Educational Technology*, 52(1), 358-373. https://doi.org/10.1111/bjet.12955

Leeser, M. J. (2004). The effects of topic familiarity, mode and pausing on second language learners' comprehension and focus on form. *Studies in Second Language Acquisition*, 26(4), 587-615. https://doi.org/10.1017/S0272263104040033

Lei, J., & Hu, G. (2014). Is English-medium instruction effective in improving Chinese undergraduate students' English competence? *International Review of Applied Linguistics in Language Teaching*, 52(2), 99-126. https://doi.org/10.1515/iral-2014-0005

Li, C., & Zhou, T. (2017). A questionnaire-based study on Chinese university students' demotivation to learn English. *English Language Teaching*, 10(3), 128-135. http://doi.org/10.5539/elt.v10n3p128

Li, K., Qin, X., Ji, S., & Zou, L. (2023). The effects of general listening anxiety and listening test anxiety on self-perceived listening performance among Chinese English learners. *Psychology Research and Behavior Management*, 16, 3541-3560. https://doi.org/10.2147/PRBM.S422030

Liu, M., & Xu, H. (2021). Testing effects of foreign language listening anxiety on Chinese university students' English listening test performance. *Frontiers in Psychology*, 12, 701926. https://doi.org/10.3389/fpsyg.2021.701926

Liu, N., Lin, C. K., & Wiley, T. G. (2016). Learner views on English and English language teaching in China. *International Multilingual Research Journal*, 10(2), 137-157. <https://doi.org/10.1080/19313152.2016.1147308>

Lowry, P. B., & Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioural causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication*, 57(2), 123-146. doi: 10.1109/TPC.2014.2312452

Macaro, E., & Erler, L. (2008). Raising the achievement of young-beginner readers of French through strategy instruction. *Applied Linguistics*, 29(1), 90-119. <https://doi.org/10.1093/applin/amm023>

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519-530. <https://doi.org/10.1093/biomet/57.3.519>

Masrai, A. (2020). Exploring the impact of individual differences in aural vocabulary knowledge, written vocabulary knowledge and working memory capacity on explaining L2 learners' listening comprehension. *Applied Linguistics Review*, 11(3), 423-447. <https://doi.org/10.1515/applrev-2018-0106>

Mayer, R. E. (2014). Research-based principles for designing multimedia instruction. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (pp. 59-70). Society for the Teaching of Psychology. <http://teachpsych.org/ebooks/asle2014/index.php>

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270. doi: 10.1037/a0028228.

Mendelsohn, D. J. (1994). *Learning to listen: A strategy-based approach for the second language learner*. Dominie Press.

Menezes, V. (2011). Affordances for language learning beyond the classroom. In P. Milliner, B., & Dimoski, B. (2024). The effects of a metacognitive intervention on lower-proficiency EFL learners' listening comprehension and listening self-efficacy. *Language Teaching Research*, 28(2), 679-713. <https://doi.org/10.1177/13621688211004646>

Mohamadi, Z. (2013). Determining the difficulty level of listening tasks. *Theory and Practice in Language Studies*, 3(6), 987-994. <https://doi.org/10.4304/tpls.3.6.987-994>

Mohammed, G. M. S., & Ali, J. K. M. (2021). Informal digital learning of English vocabulary: Saudi EFL learners' attitudes and practices. *Arab World English Journal*, 7, 345-358. <https://doi.org/10.24093/awej/call7.24>

Morshedian, M., Hemmati, F., & Sotoudehnama, E. (2017). Training EFL learners in self-regulation of reading: Implementing an SRL model. *Reading & Writing Quarterly*, 33(3), 290-303. <https://doi.org/10.1080/10573569.2016.1213147>

Naderifar, M., Goli, H., & Ghaljaie, F. (2017). Snowball sampling: A purposeful method of sampling in qualitative research. *Strides in Development of Medical Education*, 14(3), 1-4. <https://doi.org/10.5812/sdme.67670>

Nasrollahi-Mouziraji, A., & Birjandi, P. (2016). Motivational beliefs, self-regulation and EFL listening achievement: A path analysis. *Teaching English as a Second Language*

Quarterly (Formerly Journal of Teaching Language Skills), 35(3), 91-118. doi: 10.22099/jtls.2016.3923

Newmann, F. M. (1992). *Student engagement and achievement in American secondary schools* (Ed.). Teachers College Press.

Nowrouzi, S., Tam, S. S., Zareian, G., & Nimehchisalem, V. (2015). Iranian EFL students' listening comprehension problems. *Theory and Practice in Language Studies*, 5(2), 263-269. <https://doi.org/10.17507/tpls.0502.05>

Oppong, F. B., & Agbedra, S. Y. (2016). Assessing univariate and multivariate normality. A guide for non-statisticians. *Mathematical Theory and Modeling*, 6(2), 26-33.

Oxford, R. L. (1993). Research update on teaching L2 listening. *System*, 21(2), 205-211. [https://doi.org/10.1016/0346-251X\(93\)90042-F](https://doi.org/10.1016/0346-251X(93)90042-F)

Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context*. Routledge.

O'Malley, J. M., Chamot, A. U., & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, 10(4), 418-437. <https://doi.org/10.1093/applin/10.4.418>

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422. <https://doi.org/10.3389/fpsyg.2017.00422>

Passiatore, Y., Pirchio, S., Oliva, C., Panno, A., & Carrus, G. (2019). Self-efficacy and anxiety in learning English as a Foreign language: Singing in class helps speaking performance. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, 20, 121-138. <https://doi.org/10.7358/ecps-2019-020-passi>

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (3rd ed.). SAGE Publications, inc.

Payaprom, S. (2023). Investigating the relationship among metacognitive awareness, self-efficacy, and EFL learners' listening comprehension performance. *Theory and Practice in Language Studies*, 13(1), 27-34. <https://doi.org/10.17507/tpls.1301.04>

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). Academic Press.

Purdy, M., & Borisoff, D. (Eds.). (1997). *Listening in everyday life: A personal and professional approach*. University Press of America.

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research*, 45(3), 269-286. <https://doi.org/10.1080/00313830120074206>

Rahi, S. (2017). Research design and methods: A systematic review of research paradigms, sampling issues and instruments development. *International Journal of Economics & Management Sciences*, 6(2), 1-5. <https://doi.org/10.4172/2162-6359.1000403>

Rahimi, M., & Abedi, S. (2014). The relationship between listening self-efficacy and metacognitive awareness of listening strategies. *Procedia-Social and behavioural*

Sciences, 98, 1454-1460. <https://doi.org/10.1016/j.sbspro.2014.03.565>

Ranalli, J. M. (2012). *The VVT Project: A web-based platform for strategy instruction and research into self-regulated learning of L2 vocabulary* [Doctoral dissertation, Iowa State University]. Iowa State University Digital Repository. <https://doi.org/10.31274/etd-180810-1093>

Rao, Z. H. (2005). Gender, academic major, and Chinese students' use of language learning strategies: Social and educational perspectives. *Journal of Asia TEFL*, 2(3), 115-138.

Reinders, H., & Nakamura, S. (2021). Engagement. In Tammy, G., & Sarah, M (Ed.), *The Routledge handbook of the psychology of language learning and teaching* (pp. 137-148). Routledge.

Renanya, W.A., Hu, G. (2018). L2 listening in China: An examination of current practice. In: Burns, A., Siegel, J. (eds) *International perspectives on teaching the four skills in ELT*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-63444-9_3

Riley, G. L., & Lee, J. F. (1996). A comparison of recall and summary protocols as measures of second language reading comprehension. *Language Testing*, 13(2), 173-189. <https://doi.org/10.1177/026553229601300203>

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3-13. <https://doi.org/10.1111/j.1745-3992.2005.00006.x>

Rose, H., Briggs, J.G., Boggs J.A., Sergio, L., & Ivanova-Slavianskaia, N. (2017). A systematic review of language learner strategy research in the face of self-regulation. *System*, 72, 151-163. <https://doi.org/10.1016/j.system.2017.12.002>

Rost, M. (2005). L2 Listening. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 527-552). Routledge.

Rost, M. (2013). *Teaching and researching: Listening*. Routledge.

Saito, Y., Garza, T. J., & Horwitz, E. K. (1999). Foreign language reading anxiety. *The Modern Language Journal*, 83(2), 202-218. <https://doi.org/10.1111/0026-7902.00016>

Sakai, H. (2005). An examination of free written recall tasks as listening comprehension tests [Doctoral dissertation, Shinshu University Library].

Salmani, S., & Rahimi, M. (2024). Does multimodality add to the value of L2 listening homework? A mixed-methods study of comprehension and cognitive load. *English Teaching & Learning*, 1-28. <https://doi.org/10.1007/s42321-024-00181-0>

Satori, M. (2021). Effects of working memory on L2 linguistic knowledge and L2 listening comprehension. *Applied Psycholinguistics*, 42(5), 1313-1340. <https://doi.org/10.1017/S0142716421000345>

Schnelker, D. L. (2006). The student-as-bricoleur: Making sense of research paradigms. *Teaching and Teacher Education*, 22(1), 42-57. <https://doi.org/10.1016/j.tate.2005.07.001>

Schunk, D. H. (2012a). Attributions as motivators of self-regulated learning. In D, H. Schunk, & B. J. Zimmerman (Ed.), *Motivation and self-regulated learning* (pp. 245-266).

Routledge.

Schunk, D. H. (2012b). Social cognitive theory. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *Theories, constructs, and critical issues* (pp. 101-123). American Psychological Association.
<https://doi.org/10.1037/13273-005>

Serraj, S. (2015). Listening anxiety in Iranian EFL learners. *International Journal of Scientific and Research Publications*, 5(6), 543-550.

Shi, Y. Z., & Liu, Z. Q. (2006). Foreign language reading anxiety and its relationship to English achievement and gender. *Journal of PLA University of Foreign Languages*, 2, 59-64.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, 138(4), 1-28. doi: 10.1037/a0027473

Shotter, J. & Newson, J. (1982) An ecological approach to cognitive development: implicate orders, joint action, and intentionality. In G. Butterworth and P. Light (Eds.) *Social cognition: Studies in the development of understanding* (pp.32-51). Harvester.

Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, 9(1), 4-11. doi:10.12691/ajams-9-1-2.

Silver, R., Hu, G., & Iino, M. (2002). *English language education in China, Japan, and Singapore*. Graduate Programmes and Research Office, National Institute of Education, Nanyang Technological University

Simasangyaporn, N. (2016). *The effect of listening strategy instruction on Thai learners' self-efficacy, English listening comprehension and reported use of listening strategies* [Doctoral dissertation, University of Reading]. ProQuest One Academic.
<https://www.proquest.com/dissertations-theses/effect-listening-strategy-instruction-on-thai/docview/1896110853/se-2?accountid=13460>

Subon, F. (2016). Direct vocabulary instruction: The effects of contextualised word families on learners' vocabulary acquisition. *Procedia-Social and Behavioral Sciences*, 224, 284-291. <https://doi.org/10.1016/j.sbspro.2016.05.461>

Sun, Q., Zhang, L. J., & Carter, S. (2021). Investigating students' metacognitive experiences: Insights from the English as a foreign language Learners' writing metacognitive experiences questionnaire (EFLLWMEQ). *Frontiers in Psychology*, 12, 744842. <https://doi.org/10.3389/fpsyg.2021.744842>

Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, 24(3), 302-321. <https://doi.org/10.1017/S095834401200016X>

Teng, L. S., & Zhang, L. J. (2016). A questionnaire-based validation of multidimensional models of self-regulated learning strategies. *The Modern Language Journal*, 100(3), 674-701. <https://doi.org/10.1111/modl.12339>

Ting, K. Y., & Chao, M. S. (2013). The application of self-regulated strategies to blended learning. *English Language Teaching*, 6(7), 26-32. doi:10.5539/elt.v6n7p26

Tran, T. Q., & Duong, T. M. (2020). Insights into listening comprehension problems: A case study in Vietnam. *PASAA: Journal of Language Teaching and Learning in*

Thailand, 59, 77-100. doi: 10.58837/CHULA.PASAA.59.1.4

Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, 27(1), 78-102. <https://doi.org/10.1093/applin/ami046>

Ushioda, E. (2015). Context and complex dynamic systems theory. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 64-71). Multilingual Matters.

Valizadeh, M. R., & Alavinia, P. (2013). Listening comprehension performance viewed in the light of emotional intelligence and foreign language listening anxiety. *English Language Teaching*, 6(12), 11-26. doi:10.5539/elt.v6n12p11

Vandergrift, L. (2002). 'It was nice to see that our predictions were right': Developing metacognition in L2 listening comprehension. *Canadian Modern Language Review*, 58(4), 555-575. <https://doi.org/10.3138/cmlr.58.4.555>

Vandergrift, L., Goh, C. C., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language Learning*, 56(3), 431-462. <https://doi.org/10.1111/j.1467-9922.2006.00373.x>

Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences*, 39(2), 163-173. <https://doi.org/10.1080/01490400.2015.1127189>

Vogely, A. (1995). Perceived strategy use during performance on three authentic listening comprehension tasks. *The Modern Language Journal*, 79(1), 41-56. <https://doi.org/10.1111/j.1540-4781.1995.tb05414.x>

Wallace, M. P., & Lee, K. (2020). Examining second language listening, vocabulary, and executive functioning. *Frontiers in Psychology*, 11, 1122. <https://doi.org/10.3389/fpsyg.2020.01122>

Wang, C., Schwab, G., Fenn, P., & Chang, M. (2013). Self-efficacy and self-regulated learning strategies for English language learners: Comparison between Chinese and German college students. *Journal of Educational and Developmental Psychology*, 3(1), 173-191. <https://doi.org/10.5539/jedp.v3n1p173>

Wang, J. (2011). Culture differences and English teaching. *English Language Teaching*, 4(2), 223-230. <https://doi.org/10.5539/elt.v4n2p223>

Wang, Q. (2023). A study on the current state of college students' English listening anxiety. *Overseas English*, 13, 126-128.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 145-178). Routledge.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Routledge.

Winne, P. H., & Hadwin, A. F. (2012). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning* (pp. 309-326). Routledge.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, M. Zeidner, & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 531-566). Academic Press.

Wu, F., Chen, X., & Zheng, H. (2022). CET-4 listening test effect on listening learning based on machine learning. *Wireless Communications and Mobile Computing*, 2022, 5604032. <https://doi.org/10.1155/2022/5604032>

Wyatt, M. (2022). Self-efficacy. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge handbook of second language acquisition and individual differences* (pp. 207-219). Routledge.

Xu, F. (2011). Anxiety in EFL listening comprehension. *Theory & Practice in Language Studies*, 1(12), 1709-1717. <https://doi.org/10.4304/tpls.1.12.1709-1717>

Xu, J. (2017). The mediating effect of listening metacognitive awareness between test-taking motivation and listening test score: An expectancy-value theory approach. *Frontiers in Psychology*, 8, 2201. <https://doi.org/10.3389/fpsyg.2017.02201>

Xu, J., & Huang, Y. T. (2018). The mediating effect of listening metacognitive awareness between listening test anxiety and listening test performance. *The Asia-Pacific Education Researcher*, 27(4), 313-324. <https://doi.org/10.1007/s40299-018-0388-z>

Xu, J., & Luo, K. (2024). Immersing learners in English listening classroom: Does self-regulated learning instruction make a difference? *Applied Linguistics Review*, 15(1), 219-240. <https://doi.org/10.1515/applrev-2021-0171>

Xu, J., & Qiu, X. (2022). Second language psychological speaking and listening needs: Scale development, symbiosis, and demographic differences. *Studies in Second Language Learning and Teaching*, 12(3), 483-507. <https://doi.org/10.14746/ssllt.2022.12.3.7>

Xu, J., & Zheng, Y. (2024). Understanding the interconnectedness of L2 learners' writing assessment literacy and engagement: A longitudinal survey. *System*, 123, 103318. <https://doi.org/10.1016/j.system.2024.103318>

Xu, J., Fan, J., & Luo, K. (2021). Exploring L2 listening instruction, self-efficacy, and strategy use: A mediation analysis. *Frontiers in Psychology*, 12, 758757. <https://doi.org/10.3389/fpsyg.2021.758757>

Xu, T. T. (2024). Research on college English listening teaching strategies. *Knowledge Window (Teachers' Edition)*, 2, 12-14.

Yabukoshi, T. (2024). Validating a scale to measure self-regulated learning strategies for independent listening beyond the EFL classroom. *Research Methods in Applied Linguistics*, 3(1), 100090. <https://doi.org/10.1016/j.rmal.2023.100090>

Yaghoubi, S. T., & Seyyedi, F. (2017). The effect of explicit and implicit teaching vocabulary on Iranian EFL learners' vocabulary. *Studies in English Language and Education*, 4(1), 15-25. <https://doi.org/10.24815/siele.v4i1.7001>

Yeldham, M. (2019). Developing a framework for investigating L2 listeners' longitudinal development. *International Review of Applied Linguistics in Language Teaching*, 57(2), 235-263. <https://doi.org/10.1515/iral-2017-0004>

Yeung, S. S. S., Ng, M. L., Qiao, S., & Tsang, A. (2020). Effects of explicit L2 vocabulary instruction on developing kindergarten children's target and general vocabulary

and phonological awareness. *Reading and Writing*, 33, 671-689.
<https://doi.org/10.1007/s11145-019-09982-3>

Yu, S. (2015). Analysis of gender imbalance in higher education English programs: A case study of the School of Education at N College. *Overseas English*, 15, 242-244.

Yu, X., Janse, E., & Schoonen, R. (2021). The effect of learning context on L2 listening development: Knowledge and processing. *Studies in Second Language Acquisition*, 43(2), 329-354. <https://doi.org/10.1017/S0272263120000534>

Yurieva, O., Musiichuk, T., & Baisan, D. (2021). Informal English learning with online digital tools: Non-linguist students. *Advanced Education*, 17, 90-102.
<https://doi.org/10.20535/2410-8286.223896>

Zeng, Y., & Goh, C. C. (2018). A self-regulated learning approach to extensive listening and its impact on listening achievement and metacognitive awareness. *Studies in Second Language Learning and Teaching*, 8(2), 193-218.
<https://doi.org/10.14746/ssllt.2018.8.2.2>

Zhang, P. (2022). How does repetition affect vocabulary learning through listening to the teacher's explicit instruction? The moderating role of listening proficiency and preexisting vocabulary knowledge. *Language Teaching Research*, 0(0).
<https://doi.org/10.1177/13621688221140521>

Zhang, Y. (2009). An experimental study of the effects of listening on speaking for college students. *English Language Teaching*, 2(3), 194-204.

Zhang, Y., & Liu, G. (2024). Revisiting informal digital learning of English (IDLE): A structural equation modeling approach in a university EFL context. *Computer Assisted Language Learning*, 37(7), 1904-1936.
<https://doi.org/10.1080/09588221.2022.2134424>

Zhang, Y., & Xu, J. (2024). An investigation into self-efficacy, anxiety, and L2 listening comprehension: The mediation of metacognitive awareness. *International Journal of Listening*, 1-15. <https://doi.org/10.1080/10904018.2024.2386290>

Zhao, H., Chen, L., & Panda, S. (2013). Self-regulated learning ability of Chinese distance learners. *British Journal of Educational Technology*, 45(5), 941-958.
<https://doi.org/10.1111/bjet.12118>

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197-206.
<https://doi.org/10.1086/651257>

Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics*, 18(1), 49-68.
<https://doi.org/10.1093/applin/18.1.49>

Zhou, S., & Rose, H. (2022). English medium instruction in mainland China: National trends, and institutional developments. In N. Galloway, & J. McKinley (Ed.), *English medium instruction practices in higher education: International perspectives* (pp. 35-46). Bloomsbury Publishing.

Zhou, S., Hiver, P., & Al-Hoorie, A. H. (2021). Measuring L2 engagement: A review of issues and applications. In Phil, H., Ali, H. A., & Sarah, M (Ed.), *Student engagement in the language classroom* (pp. 75-98). Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781788923613-008>

Zhou, S., Xu, J., & Thomas, N. (2024). L2 listening in a digital era: Developing and validating the mobile-assisted self-regulated listening strategy questionnaire (MSRLS-Q). *System, 123*, 103310. <https://doi.org/10.1016/j.system.2024.103310>

Zhu, Z. (2021). A discussion of foreign language listening problems and their causes among intermediate EFL learners in Chinese universities. *Journal of Contemporary Educational Research, 5*(8), 47-59. <https://doi.org/10.26689/jcer.v5i8.2421>

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 1-19). Guilford Press.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). Academic Press.

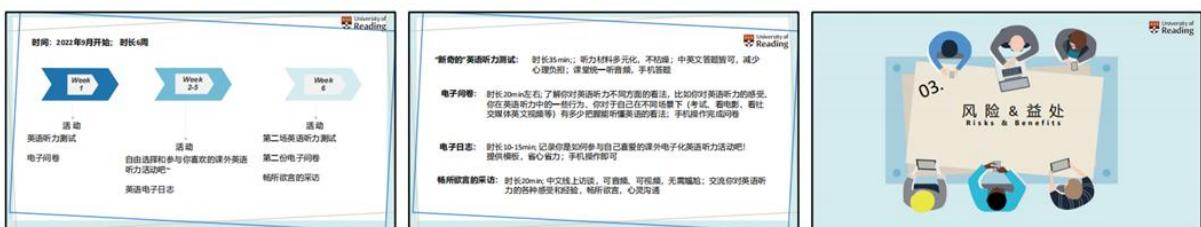
Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In DJ. Hacker, J. Dunlosky & AC. Graesser (Ed.), *Handbook of metacognition in education* (pp. 299-315). Routledge.

Öztürk, M., & Çakıroğlu, Ü. (2021). Flipped learning design in EFL classrooms: implementing self-regulated learning strategies to develop language skills. *Smart Learning Environments, 8*(1), 1-20. <https://doi.org/10.1186/s40561-021-00146-x>

APPENDICES

Appendix 1: PowerPoint for Participant Recruitment



Appendix 2: English Listening Questionnaire

Self-Regulated L2 Listening Questionnaire (SRLLQ)

(The items highlighted in yellow are the final items retained after the CFA analysis)

Please read each statement carefully, give your first reaction to each statement, and use the number below to mark an answer for every statement. Each number represents a different level of agreement, where 1 represents “strongly disagree” and 6 represents “strongly agree”.

1-strongly disagree 2 3 4 5 6-strongly agree

- 1) When I receive a listening task, I try to identify the reason or motivation for completing it.
- 2) When I receive a listening task, I try to figure out what the demands of the task are (for example, what kind of task it is, what its purpose is, what I have to do).
- 3) When I receive a listening task, I become aware of whether I can complete the task or not.
- 4) When I receive a listening task, I try to figure out what kind of listening material it is (for example, what is its genre/ topic/ cultural background?)
- 5) When I receive a listening task, I become aware of whether I feel anxious or not.
- 6) When I receive a listening task, I become aware of how important this task is for me.
- 7) When I receive a listening task, I become aware of what potential difficulties I may encounter.
- 8) Before listening, I predict what I will hear, such as topics, main ideas, words, etc.
- 9) Before listening, I plan where to focus my attention, such as on specific vocabulary, the main idea, tone, or specific parts of the listening material.
- 10) Before listening, I plan what strategies I can use in listening to help me understand the passage.

- 11) Before listening, I assess whether using certain listening strategies would effectively help me understand the passage.
- 12) Before listening, I plan how to make myself relaxed if I feel anxious.
- 13) Before listening, if I do not think I can successfully complete the task, I will plan how to convince myself that I can complete it.
- 14) When I am listening, I mainly focus on what I have planned to pay attention to, such as specific words, phrases, sentences, or gist.
- 15) When I am listening, I try to concentrate on understanding what I am listening to.
- 16) When I am listening, I assess whether parts I had difficulties with affect my overall understanding.
- 17) When I am listening, I use contextual clues to work out missing information or unfamiliar words.
- 18) When I am listening, I use common sense to work out missing information or unfamiliar words.
- 19) When I am listening, I use linguistic knowledge (e.g., vocabulary, grammar) to work out missing information or unfamiliar words.
- 20) When I am listening, I predict what I will hear next according to what I have already understood.
- 21) When I am listening, I use my knowledge of the topic to understand the passage.
- 22) When I am listening, I link different parts of the passage for better understanding.
- 23) When I am listening, I translate words, phrases, or sentences into Chinese in my mind to facilitate my understanding.
- 24) When I am listening, I pay attention to some key words, phrases or sentences that may help with my comprehension.
- 25) If I encounter listening difficulties during listening, I recall or imagine successful experiences to encourage myself.

- 26) If I encounter listening difficulties during listening, I give myself positive mental cues, believing that I can find a solution.
- 27) I use deep breathing to alleviate my listening anxiety.
- 28) When I encounter difficulties in listening, I comfort myself that I don't have to be nervous.
- 29) I compare what I have understood with what I predicted I would hear.
- 30) I check whether my understanding of a part of the passage makes sense in the light of what I have understood.
- 31) If I find myself struggling while completing a listening task, I immediately make changes to remedy the situation (e.g., changing listening strategies or reinterpreting the task requirements).
- 32) I check whether the actions I have taken to reduce my anxiety have been effective.
- 33) If I realise that the measures I have taken to reduce my anxiety are not effective, I immediately try a new approach.
- 34) I check whether the actions I have taken to encourage myself have been effective.
- 35) If I realise that the measures I have taken to encourage myself are ineffective, I will immediately adopt a new approach to motivate myself.
- 36) I evaluate whether my anxiety or lack of confidence has affected my listening comprehension.
- 37) I check whether my understanding of the passage is consistent with common sense.
- 38) I check whether my understanding of the passage is consistent with my linguistic knowledge.
- 39) After listening, I reflect on my overall performance during the listening task.
- 40) After listening, I reflect on my overall performance in managing my motivation (e.g., confidence) and affect (e.g., anxiety) during the listening task.

- 41) I look for reasons behind my success or failure in understanding the passage.
- 42) I attribute my success or failure in understanding to **internal** factors such as lack of English vocabulary, lack of English listening strategies or not putting enough effort into English listening.
- 43) I attribute my success or failure in understanding to **external** factors such as passage difficulty, fast speech input, heavy accent, noisy environment, poor sound quality, etc.
- 44) If necessary, I adjust my strategies after the task for future listening tasks.
- 45) Successful completion of one English listening task can reduce my anxiety when faced with other English listening tasks.
- 46) Successful completion of one English listening task can increase my confidence when faced with other English listening tasks.

L2 Listening Anxiety Scale (LLAS)

(The items highlighted in red are the final items retained after the CFA analysis)

The questions in this section are about how you feel about listening in English. Please read each statement carefully, give your first reaction to each statement, and use the number below to mark an answer for every statement. Each number represents a different level of agreement, where 1 represents “strongly disagree” and 6 represents “strongly agree”.

1-strongly disagree 2 3 4 5 6-strongly agree

- 1) I get upset when I'm not sure whether I understand what I'm hearing in English.
- 2) When I listen to English, I often understand the words but still can't quite understand what the speaker is saying.
- 3) When I'm listening to English, I get so confused I can't remember what I've heard.
- 4) I get anxious when I miss some information while listening to English.
- 5) I am nervous when I am listening to a passage in English when I'm not familiar with the topic.
- 6) I get upset whenever I hear unknown grammar while listening to English.
- 7) I get upset whenever I hear unknown words while listening to English.
- 8) I get anxious when I hear connected speech in the listening material.
- 9) When listening to English, I often worry about the speed of speech.
- 10) I feel anxious when I hear an unfamiliar English accent.
- 11) I don't feel nervous when listening to the passages in the typical English textbooks.
- 12) I feel nervous when listening to the passages in the typical English listening exams.
- 13) Listening to original English radio programmes (e.g., BBC, VOA) makes me nervous.
- 14) Watching English movies or TV series with original English soundtrack makes me feel nervous.
- 15) I don't feel nervous when listening to English songs

- 16) I don't feel nervous when I watch the videos on the social platforms posted by native English speakers.
- 17) I feel intimidated when I need to listen to English presentations (e.g., TED Talk), no matter they are short (e.g., less than 3 minutes) or long.

Listening Self-Efficacy Questionnaire (LSEQ)

(The items highlighted in red are the final items retained after the CFA analysis)

The questions in this section are about your belief in your capacity to perform English listening tasks in various settings. Please choose the number that best describes how sure you are that you could perform the listening skills below in different settings:



A. Think about a typical English listening passage you encounter in your textbook. How sure are you that you could

- 1) Understand the main ideas of the passage.
- 2) Understand details.
- 3) Work out the meaning of unknown or incomprehensible words.
- 4) Recognise opinions expressed in the passage.

B. Think about the last passage you listened to in the listening test. How sure are you that you could

- 5) Understand the main ideas of the passage.
- 6) Understand details.
- 7) Work out the meaning of unknown or incomprehensible words.
- 8) Recognise opinions expressed in the passage.

C. Think about a typical VOA/ BBC English radio programme. How sure are you that you could

- 9) Understand the main ideas of the passage.
- 10) Understand details.
- 11) Work out the meaning of unknown or incomprehensible words.
- 12) Recognise opinions expressed in the passage.

D. Think about a typical English TV series/ movie. How sure are you that you could

13) Understand the main ideas of the passage.

14) Understand details.

15) Work out the meaning of unknown or incomprehensible words.

16) Recognise opinions expressed in the passage.

E. Think about a typical video posted by English-speaking vloggers on social media platforms. How sure are you that you could

17) Understand the main ideas of the passage.

18) Understand details.

19) Work out the meaning of unknown or incomprehensible words.

20) Recognise opinions expressed in the passage.

F. Think about a typical presentation by English speakers (e.g., TED Talk).

How sure are you that you could

21) Understand the main ideas of the passage.

22) Understand details.

23) Work out the meaning of unknown or incomprehensible words.

24) Recognise opinions expressed in the passage.

G. Think about a typical English audio book, such as The Little Prince. How sure are you that you could

25) Understand the main ideas of the passage.

26) Understand details.

27) Work out the meaning of unknown or incomprehensible words.

28) Recognise opinions expressed in the passage.

H. Think about a typical English song. How sure are you that you could

29) Understand the main ideas of the song.

30) Understand details.

31) Work out the meaning of unknown or incomprehensible words.

32) Recognise opinions expressed in the song.

Appendix 3: IDLEL E-Log

This is the E-log that you can use to record your engagement in informal digital learning of English listening (IDLEL) activities. Please use English to complete the E-log.

1. Last four digits of your student ID*

2. This is the _____ week since I started using this E-log. [single choice]

- first
- second
- third
- fourth

3. How many IDLEL activities did you participate in this week? [single choice]

- 1
- 2
- 3
- 4
- 5
- Others (more than five) _____
- 0 (please indicate your reasons for not taking part in IDLEL activities here) _____

4. What IDLEL activity did you participate in? (Activity 1) [single choice]

- English movies
- English TV series
- English variety shows or talk-shows (e.g., Saturday Night Live; The Tonight Show Starring Jimmy Fallon)
- English videos posted by influencers on social platforms
- English listening materials in English learning apps

- English presentations (e.g., TED Talk)
- English radio programmes (e.g., BBC; VOA)
- English audio books (e.g., The Little Prince)
- English songs
- Listening materials for tests (e.g., TEM-4; IELTS) or textbooks
- Others (please indicate what the activity is) _____

5. What IDLEL activity did you participate in? (Activity 2) [single choice]

- English movies
- English TV series
- English variety shows or talk-shows (e.g., Saturday Night Live; The Tonight Show Starring Jimmy Fallon)
- English videos posted by influencers on social platforms
- English listening materials in English learning apps
- English presentations (e.g., TED Talk)
- English radio programmes (e.g., BBC; VOA)
- English audio books (e.g., The Little Prince)
- English songs
- Listening materials for tests (e.g., TEM-4; IELTS) or textbooks
- Others (please indicate what the activity is) _____

6. What IDLEL activity did you participate in? (Activity 3) [single choice]

- English movies
- English TV series
- English variety shows or talk-shows (e.g., Saturday Night Live; The Tonight Show Starring Jimmy Fallon)
- English videos posted by influencers on social platforms
- English listening materials in English learning apps
- English presentations (e.g., TED Talk)

- English radio programmes (e.g., BBC; VOA)
- English audio books (e.g., The Little Prince)
- English songs
- Listening materials for tests (e.g., TEM-4; IELTS) or textbooks
- Others (please indicate what the activity is) _____

7. What IDLEL activity did you participate in? (Activity 4) [single choice]

- English movies
- English TV series
- English variety shows or talk-shows (e.g., Saturday Night Live; The Tonight Show Starring Jimmy Fallon)
- English videos posted by influencers on social platforms
- English listening materials in English learning apps
- English presentations (e.g., TED Talk)
- English radio programmes (e.g., BBC; VOA)
- English audio books (e.g., The Little Prince)
- English songs
- Listening materials for tests (e.g., TEM-4; IELTS) or textbooks
- Others (please indicate what the activity is) _____

8. What IDLEL activity did you participate in? (Activity 5) [single choice]

- English movies
- English TV series
- English variety shows or talk-shows (e.g., Saturday Night Live; The Tonight Show Starring Jimmy Fallon)
- English videos posted by influencers on social platforms
- English listening materials in English learning apps
- English presentations (e.g., TED Talk)

- English radio programmes (e.g., BBC; VOA)
- English audio books (e.g., The Little Prince)
- English songs
- Listening materials for tests (e.g., TEM-4; IELTS) or textbooks
- Others (please indicate what the activity is) _____

9. In addition to the activities mentioned above, what other activities did you participate in?

10. I engaged in activity 1 for about _____ [single choice]

- <1 hour
- 1-2 hours
- 2-3 hours
- 3-4 hours
- Others _____

11. I engaged in activity 2 for about _____ [single choice]

- <1 hour
- 1-2 hours
- 2-3 hours
- 3-4 hours
- Others _____

12. I engaged in activity 3 for about _____ [single choice]

- <1 hour
- 1-2 hours
- 2-3 hours
- 3-4 hours

Others _____

13. I engaged in activity 4 for about _____ [single choice]

<1 hour

1-2 hours

2-3 hours

3-4 hours

Others _____

14. I engaged in activity 5 for about _____ [single choice]

<1 hour

1-2 hours

2-3 hours

3-4 hours

Others _____

15. Please record other activities and the time you spent on these activities here.

Example:

I engage in _____ for about _____ hour/hours.

In this part, please briefly describe how you made use of the IDLEL activities you mentioned above to enhance your English listening.

16. Reasons for taking part in activity 1. (Multiple choice acceptable)

Personal interest

It's quite popular recently

Classmates' recommendation

Teachers' recommendation

Others _____

17. Reasons for taking part in activity 2. (Multiple choice acceptable)

- Personal interest
- It's quite popular recently
- Classmates' recommendation
- Teachers' recommendation
- Others _____

18. Reasons for taking part in activity 3. (Multiple choice acceptable)

- Personal interest
- It's quite popular recently
- Classmates' recommendation
- Teachers' recommendation
- Others _____

19. Reasons for taking part in activity 4. (Multiple choice acceptable)

- Personal interest
- It's quite popular recently
- Classmates' recommendation
- Teachers' recommendation
- Others _____

20. Reasons for taking part in activity 5. (Multiple choice acceptable)

- Personal interest
- It's quite popular recently
- Classmates' recommendation
- Teachers' recommendation

Others _____

21. Please briefly describe your reasons for taking part in other activities.

Example, I took apart in activity _____ because _____

22. Any mental preparation I performed **before** I started being involved in activity 1 (e.g., any ideas, thoughts, or feelings about what you are going to listen to) _____

23. Any mental preparation I performed **before** I started being involved in activity 2 (e.g., any ideas, thoughts, or feelings about what you are going to listen to) _____

24. Any mental preparation I performed **before** I started being involved in activity 3 (e.g., any ideas, thoughts, or feelings about what you are going to listen to) _____

25. Any mental preparation I performed **before** I started being involved in activity 4 (e.g., any ideas, thoughts, or feelings about what you are going to listen to) _____

26. Any mental preparation I performed **before** I started being involved in activity 5 (e.g., any ideas, thoughts, or feelings about what you are going to listen to) _____

27. Any mental preparation I performed **before** I started being involved in other activities (e.g., any ideas, thoughts, or feelings about what you are going to listen to). Before _____,

I _____

28. To make it beneficial for improving my listening comprehension, what did I do during activity 1? (e.g., Did I encounter any listening comprehension difficulties? How did I solve these difficulties? Did I use any strategies to help with my listening? etc.) _____

29. To make it beneficial for improving my listening comprehension, what did I do during activity 2? (e.g., Did I encounter any listening comprehension difficulties? How did I solve these difficulties? Did I use any strategies to help with my listening? etc.) _____

30. To make it beneficial for improving my listening comprehension, what did I do during activity 3? (e.g., Did I encounter any listening comprehension difficulties? How did I solve these difficulties? Did I use any strategies to help with my listening? etc.) _____

31. To make it beneficial for improving my listening comprehension, what did I do during activity 4? (e.g., Did I encounter any listening comprehension difficulties? How did I solve these difficulties? Did I use any strategies to help with my listening? etc.) _____

32. To make it beneficial for improving my listening comprehension, what did I do during activity 5? (e.g., Did I encounter any listening comprehension difficulties? How did I solve these difficulties? Did I use any strategies to help with my listening? etc.) _____

33. To make it beneficial for improving my listening comprehension, what did I do during other activities? (e.g., Did I encounter any listening comprehension difficulties? How did I solve these difficulties? Did I use any strategies to help with my listening? etc.)

34. What did I do **after** activity 1 to help me improve English listening _____

35. What did I do after activity 2 to help me improve English listening _____

36. What did I do after activity 3 to help me improve English listening _____

37. What did I do after activity 4 to help me improve English listening _____

38. What did I do after activity 5 to help me improve English listening _____

39. What did I do **after** other activities to help me improve English listening?

I _____ after _____

40. My gains through activity 1 (multiple choice)

- Vocabulary expanded
- Got familiar with natural accent
- Got familiar with natural speech speed
- Learned about target culture
- Understood long and complex sentences
- Confidence in English listening increased
- Anxiety in English listening decreased
- Others _____

41. My gains through activity 2 (multiple choice)

- Vocabulary expanded
- Got familiar with natural accent
- Got familiar with natural speech speed
- Learned about target culture
- Understood long and complex sentences
- Confidence in English listening increased
- Anxiety in English listening decreased
- Others _____

42. My gains through activity 3 (multiple choice)

- Vocabulary expanded
- Got familiar with natural accent
- Got familiar with natural speech speed
- Learned about target culture

- Understood long and complex sentences
- Confidence in English listening increased
- Anxiety in English listening decreased
- Others _____

43. My gains through activity 4 (multiple choice)

- Vocabulary expanded
- Got familiar with natural accent
- Got familiar with natural speech speed
- Learned about target culture
- Understood long and complex sentences
- Confidence in English listening increased
- Anxiety in English listening decreased
- Others _____

44. My gains through activity 5 (multiple choice)

- Vocabulary expanded
- Got familiar with natural accent
- Got familiar with natural speech speed
- Learned about target culture
- Understood long and complex sentences
- Confidence in English listening increased
- Anxiety in English listening decreased
- Others _____

45. My gains through other activities (Please briefly describe your gains from other activities here). Example, My gains from activity _____ include _____

Well done! 🎉 Thanks for completing the E-log for this week!

Appendix 4: English Listening Tests

Listening Test 1

Listening Test Answer Sheet

Student ID (last four digits):

Task 1

In this section, you will hear four short passages. The first three passage will be played in two parts, with **ONE** minute pause in between for you to write down everything you understood from the first part. After the second part is played, you will have another **ONE** minute to write down your understanding of the second part. Every passage will only be played **ONCE**. Both **ENGLISH** and **CHINESE** are acceptable, so please write your understanding of the passage as much as possible.

Passage 1

Part 1

Part 2

Passage 2

Part 1

Part 2

Passage 3

Part 1

Part 2

Passage 4

Part 1

Part 2

Task 2

In this section, you will hear the four passages again. After each passage, please answer the listening comprehension questions below. There is only **ONE** right answer for each question. Every passage will only be played **ONCE**. Before the passage is played, you will have **30 Seconds** to read the questions. After each passage, you will have **One Minute** to answer the questions on it.

Passage 1

1) Luke read that one reason why we often forget dreams is that:

- A. Our memories cannot deal with too much information
- B. We might otherwise be confused about what is real
- C. We do not think they are important

2) What do Luke and Susie agree about dreams predicting the future?

- A. It may just be due to coincidence
- B. It only happens with certain types of event
- C. It happens more often than some people think

Passage 2

3) According to the news, what does “upcycled items” mean?

- A. Things modified from old things that are no longer in use

- B. Unwanted items
- C. Things that are broken-down

4) Why does the reporter say that the mall is clever?

- A. It is situated next to a recycling centre
- B. It is not only a mall but also a recycling centre
- C. The products it sells are very popular

Passage 3

5) When the speaker decided to learn German, what did she do?

- A. She read “Harry Potter” in German
- B. She read “Friends” in German
- C. She watched “Friends” in German

6) According to the speaker, what may be the secret of learning languages?

- A. Try to talk to native speakers instead of talking to yourself
- B. Keep listening to textbook materials even if you don’t like them
- C. Discover the fun of language learning by making it an enjoyable daily activity

Passage 4

7) Which of the following statements about “destiny” is TRUE?

- A. The woman believed that being a parent was everyone’s destiny
- B. The woman knew what the man’s destiny was but did not tell him before she died
- C. The woman thought the man could make and discover his own destiny

8) What does the mother mean by comparing life to a box of chocolates?

- A. Life is as sweet as chocolates

- B. Life is full of uncertainty
- C. Life is full of predictable events

Listening Test 2

Listening Test Answer Sheet

Student ID (last four digits):

Task 1

In this section, you will hear four short passages. The first three passage will be played in two parts, with **ONE** minute pause in between for you to write down everything you understood from the first part. After the second part is played, you will have another **ONE** minute to write down your understanding of the second part. Every passage will only be played **ONCE**. Both **ENGLISH** and **CHINESE** are acceptable, so please write your understanding of the passage as much as possible.

Passage 1

Part 1

Part 2

Passage 2

Part 1

Part 2

Passage 3

Part 1

Part 2

Passage 4

Part 1

Part 2

Task 2

In this section, you will hear the four passages again. After each passage, please answer the listening comprehension questions below. There is only **ONE** right answer for each question. Every passage will only be played **ONCE**. Before the passage is played, you will have **30 Seconds** to read the questions. After each passage, you will have **One Minute** to answer the questions on it.

Passage 1

1) What does Adam suggest that the restaurants could do to reduce obesity?

- A. offer fewer food options
- B. have more low-calorie foods
- C. organise menus in a particular way

2) What do Adam and Rosie think about the levels of exercise in England?

- A. The amount recommended is much too low
- B. Most people overestimate how much they do
- C. Women now exercise more than they used to do

Passage 2

3) What does “leftover bread” mean?

- A. Outdated bread
- B. Daily unsold bread

○C. Bad-tasting bread

4) The main reason for turning bread into beer may be that

○A. it is beneficial to solving food waste problems

○B. it is a tradition that needs be followed

○C. this product can make huge profits

Passage 3

5) What does the speaker mean by saying “live the language”?

○A. Making language learning interesting

○B. Using that language as much as possible

○C. Living in countries where that language is spoken

6) Why does “making mistakes” help us learn language?

○A. Because it can make us less vulnerable

○B. Because we expected to do that when we were children

○C. Because it provides us with the freedom and space to progress

Passage 4

7) At first, what was the father’s attitude towards the boy’s dream of becoming a professional basketball player?

○A. He thought the boy should spend time practising playing basketball

○B. He thought the boy had talent in playing basketball

○C. He did not think the boy could be a professional basketball player

8) What’s the main point of the man’s words in the end?

○A. The boy should take advice from others

○B. The boy should stick to achieving his dream

○C. The man's suggestions are more important than others'

Listening Test 3

Listening Test Answer Sheet

Student ID (last four digits):

Task 1

In this section, you will hear four short passages. The first three passage will be played in two parts, with **ONE** minute pause in between for you to write down everything you understood from the first part. After the second part is played, you will have another **ONE** minute to write down your understanding of the second part. Every passage will only be played **ONCE**. Both **ENGLISH** and **CHINESE** are acceptable, so please write your understanding of the passage as much as possible.

Passage 1

Part 1

Part 2

Passage 2

Part 1

Part 2

Passage 3

Part 1

Part 2

Passage 4

Part 1

Part 2

Task 2

In this section, you will hear the four passages again. After each passage, please answer the listening comprehension questions below. There is only **ONE** right answer for each question. Every passage will only be played **ONCE**. Before the passage is played, you will have **30 Seconds** to read the questions. After each passage, you will have **One Minute** to answer the questions on it.

Passage 1

5) When the speaker decided to learn German, what did she do?

- A. She read “Harry Potter” in German
- B. She watched “Friends” in German
- C. She read “Friends” in German

6) According to the speaker, what may be the secret of learning languages?

- A. Keep listening to textbook materials even if you don’t like them
- B. Discover the fun of language learning by making it an enjoyable daily activity
- C. Try to talk to native speakers instead of talking to yourself

Passage 2

7) Which of the following statements about “destiny” is TRUE?

- A. The woman believed that being a parent was everyone's destiny
- B. The woman thought the man could make and discover his own destiny
- C. The woman knew what the man's destiny was but did not tell him before she died

8) What does the mother mean by comparing life to a box of chocolates?

- A. Life is full of predictable events
- B. Life is as sweet as chocolates
- C. Life is full of uncertainty

Passage 3

3) According to the news, what does "upcycled items" mean?

- A. Unwanted items
- B. Things modified from old things that are no longer in use
- C. Things that are broken-down

4) Why does the reporter say that the mall is clever?

- A. It is not only a mall but also a recycling centre
- B. It is situated next to a recycling centre
- C. The products it sells are very popular

Passage 4

1) Luke read that one reason why we often forget dreams is that:

- A. We might otherwise be confused about what is real
- B. Our memories cannot deal with too much information
- C. We do not think they are important

2) What do Luke and Susie agree about dreams predicting the future?

- A. It happens more often than some people think
- B. It only happens with certain types of event
- C. It may just be due to coincidence

Appendix 5: Main Ideas for the Free-Recall Task

Test 1

Total main ideas: 20

Passage 1

1. It's a sort of protection for us to forget most of our dreams.
2. Dreams predicting the future are purely coincidental.
3. We remember dreams when real-world events trigger their memory.
4. Segmented sleeping involves waking up in the middle of the night for activities before returning to sleep.
5. The speaker stopped practicing segmented sleeping after a week as it didn't suit their lifestyle.

Passage 2

1. Everything for sale in the mall is second-hand.
2. Upcycled items are unwanted items broken-down and reinvented as something new.
3. The clever thing about this mall is its location: next to the city's recycling centre.
4. The mall's shopkeepers select items from the recycled goods that they want to sell or use as materials for upcycling.
5. One of the most popular shops in the mall specialises in selling items made from old leather jackets.

Passage 3

1. Persistent exposure to favourite sitcoms and books enabled the presenters' comprehension improvement in foreign language learning.
2. Transforming language learning into an enjoyable and engaging daily activity.
3. If you prefer not to write on paper, you can use an app instead.
4. Make language learning more enjoyable by choosing engaging online content over boring textbook material.
5. Use self-talk to practice speaking if you're introverted and not yet comfortable speaking with native speakers.

Passage 4

1. Death is something we're all destined to do.
2. You make your own destiny by doing the best with what God gave you.
3. Life is unpredictable.
4. Momma had a unique way of making things easy for me to understand.
5. Mom died from cancer.

Test 2
Total main ideas: 20

Passage 1

1. Placing low-calorie items at the beginning and end of menus may encourage healthier choices.
2. Some food labels have misleading calorie counts, possibly intentionally.
3. Get people to exercise more.
4. Many claim to meet exercise recommendations, but actual measurements show few do.
5. Obesity is growing.

Passage 2

1. The leftover bread at this shop is made into beer to prevent food waste.
2. Leftover buns accumulate significantly over time.
3. Making beer out of bread is an old tradition which dates back to Ancient Egypt.
4. This beer is a niche product that cannot compete with large breweries.
5. Making bread into beer is about raising awareness of the problem of food waste.

Passage 3

1. There are three golden rules of language learning.
2. The first rule is immersing yourself in the language in every possible way.
3. The second rule is making as many mistakes as you want.
4. Making mistakes are essential for learning and improvement.
5. The third rule is making language learning enjoyable to stay motivated.

Passage 4

1. The dad and the boy went to play basketball.
2. The boy wants to be a professional basketball player.
3. The dad thinks that he was below average and his son will be like him.
4. Don't ever let somebody tell you that you can't do something.
5. Protect your dreams and actively work towards achieving them.

Appendix 6: The Codebook For the Thematic Analysis Purpose

Codebook for the Task-Level Strategy

Task Perception	Goal Setting	Listening Strategy Planning	Strategy Employment
Task demand perception	When receiving a listening task, participants try to figure out what the demands of the task are (for example, what kind of task it is, what its purpose is, what they have to do). When participants receive a listening task, they try to figure out the nature of the listening material, e.g. its genre, theme, cultural context, difficulty, etc.	Setting goals for task completion Before listening, participants set goals for task completion.	Selective attention In listening, participants mainly focus on specific words, phrases, sentences, gist, accent, cultural aspects, etc.
Listening material perception	 Setting goals for comprehension levels Setting goals for selected attention Setting goals for skill enhancement	Before listening, participants set goals for the level of comprehension of the listening material. Before listening begins, participants set goals to plan what to focus on, e.g., focusing on specific vocabulary words, main ideas, detailed information, or specific parts of the listening material (e.g., dialogues, beginnings, endings, etc.). Before listening, participants set goals for the skills they want to improve or the knowledge they want to gain through the listening task	Focus on ongoing contents In listening, participants focus on understanding what they are listening to.
Previous knowledge recall	When participants receive a listening task, they try to recall related knowledge	Listening strategy planning Before listening, participants plan which listening strategies to use or not to use to help them understand the passage or achieve their goals	Contextual-clue assistance In listening, participants use contextual clues to work out missing information or unfamiliar words.
	 Content prediction	 Before listening, participants predict what they are going to listen to, such as possible vocabulary and arguments. These predictions may become standards or targets for participants to monitor their comprehension later.	Common-sense assistance In listening, participants use common sense to work out missing information or unfamiliar words.
		 Linguistic-knowledge assistance Topic-knowledge assistance Cultural background assistance Understanding-based assistance Mental Translation	Subtitles/ Scripts Dictionary/ Translation websites/ apps Pause/ Speed Adjustment Imitation Question-clue assistance
			In listening, participants use linguistic knowledge (e.g., vocabulary, grammar) to work out missing information or unfamiliar words or the meaning of complex sentences. In listening, participants use their knowledge of a topic (e.g. sports, food culture) to understand the passage. In listening, participants use their knowledge of the target culture to understand the passage. In listening, participants use what they have already understood as assistance, to predict or what they will hear next or to inference what they have not understood In listening, participants translate words, phrases, or sentences into Chinese in their mind to facilitate their understanding.
			Participants may re-play the listening materials for many times Participants use or close the subtitles/ scripts to help them better understand the listening materials or achieve their goals Look up unfamiliar words participants may pause or adjust the speed of the listening material to gain more time to understand it or to take notes During listening, participants may familiarise themselves with the pronunciation or accent of unfamiliar words by mimicking or following along. Participants use the stem or options of a question to infer or understand the content of the original text

		Visualisation	In listening, participants visualise words, phrases, and sentences they hear in their mind to aid their comprehension.	Skip	Participants will skip over words they don't understand at first and try to understand their meaning when the listening is over or on a second listen
--	--	---------------	--	------	---

	Listening Monitoring	Reflection	Attribution	Adjustment	
Comprehension monitoring	Participants compare what they have understood with a variety of sources to monitor their listening comprehension. More detailed strategies under this category can be seen below.	Task performance reflection	Participants reflect on whether the goals set before listening are met or not or on listening knowledge	Internal-factor attribution Participants attribute their success or failure in understanding to internal factors such as lack of English vocabulary, lack of English listening strategies or not putting enough effort into English listening.	Plan for future study Participants set goals/ plans after the task for future, maybe not limited to listening.
		Strategy-use reflection	Participants reflect on the effectiveness of the strategies used in listening.	External-factor attribution Participants attribute their success or failure in understanding to external factors such as passage difficulty, fast speech input, heavy accent, noisy environment, poor sound quality, etc.	Help & Seeking seeking Participants will seek help or feedback from others for difficulties that they cannot solve on their own
Difficulty-based monitoring	Participants assess whether parts they had difficulties with affect their overall understanding.	Knowledge accumulation	Memorise unfamiliar or 'useful' words, sentence structures, expressions, cultural phenomena or even other skills in listening/ after listening	Sharing	Participants will share the knowledge/ ideas they gained from the listening activities
		Task-gained knowledge review	Participants review or expand the knowledge gained from the task or re-do the task		

	Motivation & Affect Awareness	Goal clarification	Motivation and Affect-level Strategy Employment	Motivation and Affect Reflection & Adjustment
Self-efficacy awareness	When receiving a listening task, participants become aware of whether they can complete the task or not.	Before listening, participants determine their goal orientations (e.g., mastery-approach/avoidance orientations; performance-approach/avoidance orientations). In other words, participants determine their reasons for engaging in listening activities.	Anxiety-relieving physical strategies	Before listening, participants try to alleviate their anxiety by using some physical behaviours, such as taking deep breath, talking to others or themselves or jumping
	When receiving a listening task, participants become aware of whether they feel anxious or not.		Gather mind and focus	
Anxiety awareness	When receiving a listening task, participants become aware of whether they feel anxious or not.	Before listening, participants encourage themselves to be confident by imagining good results, such as getting good grades in this listening task	Imaging good results	Participants adjust their motivational & affective states in the face of future tasks based on the completion of one listening task
	When receiving a listening task, participants become aware of how important this task is for them.		Recall previous experiences	
Task-value awareness	When receiving a listening task, participants become aware of how important this task is for them.	Before listening, even though participants may be aware that the task will be difficult or unsure if they will be able to complete the task, they give themselves the courage to push themselves to start the task	Self-push/encourage/ comfort	Before listening, participants may choose/ create a quiet/ clean environment to help them focus on listening or relieve their anxiety
	When receiving a listening task, participants become aware of what potential difficulties they may encounter.		Mental preparation for potential difficulties	
Potential difficulty awareness	When receiving a listening task, participants become aware of what potential difficulties they may encounter.	Before listening begins, participants can gain mental preparation by envisioning the potential difficulties that may arise during the listening process, thus easing their anxiety	Selecting/ Creating an environment conducive to concentration	

Appendix 7: Ethical Approval & Consent Forms

Ethical Approval Form

Please tick one:

Staff:

PhD:

EdD:

Name of applicant(s): Huining Yang

Title of project: The development of learners' listening comprehension and motivational variables within informal digital learning of English listening (IDLEL) context: Examining the role of frequency, diversity of use, and self-regulation

Name of supervisor (s) (for student projects): Prof. Suzanne Graham; Dr. Pengchong Anthony Zhang

Please complete the form below.

Have you prepared an Information Sheet for participants and/or their parents/carers that	YES	NO	N.A.
a) explains the purpose(s) of the project	√		
b) explains how they have been selected as potential participants	√		
c) gives a full, fair, and clear account of what will be asked of them and how the information that they provide will be used	√		
d) makes clear that engagement in the project is voluntary	√		
e) explains the arrangements to allow participants to withdraw at any stage if they wish	√		
f) explains the arrangements to ensure the confidentiality of any material collected during the project, including secure arrangements for its storage, retention, and disposal	√		
g) explains the arrangements for publishing the research results and, if confidentiality might be affected, for obtaining written consent for this	√		
h) explains the arrangements for providing participants with the research results if they wish to have them	√		
i) gives the name and designation of the member of staff with responsibility for the project together with contact details, including email. If any of the project investigators are students at the IoE, then this information must be included, and their name provided	√		
j) explains, where applicable, the arrangements for expenses and other payments to be made to the participants	√		
k) includes a standard statement indicating the process of ethical review at the University undergone by the project, as follows: "This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct."	√		
l) includes a standard statement regarding insurance: "The University has the appropriate insurances in place. Full details are available on request."	√		

Please answer the following questions:	YES	NO	N.A.
1) Will you provide participants involved in your research with all the information necessary to ensure that they are fully informed and not in any way deceived or misled as to the purpose(s) and nature of the research? (Please use the subheadings used in the example information sheets on blackboard to ensure this).	√		
2) Will you seek written or other formal consent from all participants, if they are able to provide it, in addition to 1)?	√		
3) Is there any risk that participants may experience physical or psychological distress in taking part in your research?		√	
4) Staff Only – Have you taken the online training modules in data protection and information security (which can be found here: http://www.reading.ac.uk/internal/humanresources/PeopleDevelopment/newstaff/humres-MandatoryOnlineCourses.aspx			√
For all student projects, please tick N.A. and complete the Data Protection Declaration form (which is included in this document) and submit it with this application to the ethics committee.			
5) Have you read the Health and Safety booklet (available on Blackboard) and completed a Risk Assessment Form (included below with this ethics application)?	√		
6) Does your research comply with the University's Code of Good Practice in Research?	√		
7) If your research is taking place in a school, have you prepared an information sheet and consent form to gain the permission in writing of the head teacher or other relevant supervisory professional?	√		
8) Has the data collector obtained satisfactory DBS clearance?	√		
9) If your research involves working with children under the age of 16 (or those whose special educational needs mean they are unable to give informed consent), have you prepared an information sheet and consent form for parents/carers to seek permission in writing, or to give parents/carers the opportunity to decline consent?			√
10) If your research involves processing sensitive personal data ¹ , or if it involves audio/video recordings, have you obtained the explicit consent of participants/parents?	√		
11) If you are using a data processor to subcontract any part of your research, have you got a written contract with that contractor which (a) specifies that the contractor is required to act only on your instructions, and (b) provides for appropriate technical and organisational security measures to protect the data?			√
12a) Does your research involve data collection outside the UK?	√		

¹ Sensitive personal data consists of information relating to the racial or ethnic origin of a data subject, their political opinions, religious beliefs, trade union membership, sexual life, physical or mental health or condition, or criminal offences or record.

12b) If the answer to question 12a is “yes”, does your research comply with the legal and ethical requirements for doing research in that country?	√		
13a) Does your research involve collecting data in a language other than English?	√		
13b) If the answer to question 13a is “yes”, please confirm that information sheets, consent forms, and research instruments, where appropriate, have been directly translated from the English versions submitted with this application.	√		
14a. Does the proposed research involve children under the age of 5?		√	
14b. If the answer to question 14a is “yes”: My Head of School (or authorised Head of Department) has given details of the proposed research to the University’s insurance officer, and the research will not proceed until I have confirmation that insurance cover is in place.			√

If you have answered YES to Question 3, please complete Section B below

- Complete either **Section A** or **Section B** below with details of your research project.
- Complete a **Risk Assessment**.
- Sign the form in **Section C**.
- For all student projects, complete a **Data Protection Declaration form**.
- Append at the end of this form all relevant documents: information sheets, consent forms, and ALL research instruments which may include tests, and questionnaires, and for staff, evidence that you have completed information security training (e.g., screen shot/copy of certificate).
- Email the completed form, as a **SINGLE** document, to the Institute’s Ethics Committee for consideration.

Any missing information will result in the form being returned to you.

Section A: My research goes beyond the “accepted custom and practice of teaching” but I consider that this project has **no** significant ethical implications. (Please tick the box.)



Please state the total number of participants that will be involved in the project and give a breakdown of how many there are in each category e.g., teachers, parents, pupils etc.

Main study: One hundred and fifty second-year undergraduates

Pilot study: Thirty-three first-year undergraduates

Give a succinct description of the aims and the methods (participants, instruments, and procedures) of the project in up to 500 words noting:

1. Title of project

The development of learners’ listening comprehension and motivational variables within informal digital learning of English listening (IDLEL) context: Examining the role of frequency, diversity of use, and self-regulation

2. Purpose of project and its academic rationale

The worldwide spread of the epidemic in recent years has significantly increased the demand for technology in education. Technology-assisted language learning has quite a long research tradition but with less focus on informal digital learning of English listening (IDLEL). This mixed-method research aims to fill the gap by investigating the role of frequency (i.e., the time that learners spend on IDLE activities), diversity (i.e., the range of IDLE activities that learners engage in), and self-regulation in learners' development of listening comprehension and motivational variables (i.e., anxiety and self-efficacy) through Chinese undergraduates' IDLEL practices.

3. Brief description of methods and measurements

Both quantitative and qualitative methods will be employed in this research, with two listening comprehension tests, a questionnaire, and English e-logs being the instruments. Specifically, two listening comprehension tests of comparable levels of difficulty (i.e., similar speech speed, duration, vocabulary difficulty, accent, and topic) will be administered as a pre-test and a post-test to explore participants' listening proficiency before and after the study respectively. The tests will be developed for this study, including free-call tasks (to investigate participants' comprehension of the gist and opinions of the passage), fill-in-the-blank tasks, and open-ended short answer tasks (to explore participants' comprehension of detailed information of the passage). Each test will contain four audio clips of around 2 minutes from the IELTS Listening test, TED TALK, BBC radio programmes and English language films respectively. Design like this may help provide a comprehensive picture of participants' listening proficiency in both formal and informal settings.

As for participants' listening self-efficacy, listening anxiety and self-regulated listening ability, they will be investigated in different sections of one questionnaire. The questionnaire will be administrated via a questionnaire app (i.e., wenjuan.com) to participants both before and after the research to collect quantitative data. Whenjuan.com is a professional online questionnaire platform, the results of which can only be reviewed by the data collector. The security and privacy of the data can be guaranteed in the absence of cyber security incidents such as hacking.

English e-logs will be distributed to participants via the Word software to collect both qualitative and quantitative data. Participants need to complete their e-logs and send their e-logs by email to the researcher. Only the researcher can receive and review e-logs, thus the security and privacy of the data on e-logs can be guaranteed as long as there are no cyber security incidents, such as hacking.

4. Participants: Recruitment methods, number, age, gender, exclusion/inclusion criteria

For the main study, one hundred and fifty second-year undergraduates majoring in English from

a Chinese university will be recruited to participate in the study. One week before the start of the research, students will be informed of the purpose, length, process of the study, the principles of privacy protection, benefits (i.e., students who take part in the study will receive feedback from the researcher on English writing regarding their e-logs; participants will obtain learning materials developed by the researcher on self-regulated L2 listening and L2 listening strategies at the end of the study), and potential risks. All students are free to decide whether to take part in the research or not and they will become participants of the study as long as they agree to get involved in the research. Participants will be required to complete listening tests (N=150), and questionnaires (N=150), E-logs (N=150) over a six-week period.

For the pilot study, thirty-three first-year English undergraduates enrolled in the first semester of the 2021-2022 Media English Listening and Speaking course from a Chinese university will be recruited to participate in the study. One week before the start of the research, students will be informed of the purpose, length, process of the study, the principles of privacy protection, benefits (i.e., students who take part in the study as participants will obtain learning materials on self-regulated L2 listening and L2 listening strategies at the end of the study; participants who complete their English e-logs will receive feedback from the researcher on their English writing as rewards). Participants will be required to complete listening tests (N=33), questionnaires (N=33), and E-logs (N=2) over a three-week period. Unlike the main study, the questionnaire for the pilot study will end with feedback from students on the clarity of the questionnaire questions so as to make potential improvements to the instruments of the main study.

5. Consent and participant information arrangements, debriefing (attach forms where necessary)
Please see information and consent forms

6. A clear and concise statement of the ethical considerations raised by the project and how you intend to deal with them.

Firstly, regarding informed consent issue, participants will be informed of the purpose, processes, risks, and benefits of the research before they decide to take part in the research or not. Secondly, participants have the right to withdraw from the research at any point and for any reason. Thirdly, as for the confidentiality issue, apart from basic information about the participants, such as gender, age and sex, other privacy issues will not be collected. In addition, the above-mentioned information will be anonymised during the data analysis, writing and publication process. Fourthly, this research is basically free from any type of harm (e.g., physical, social, or psychological). Finally, it is for sure that this research will be free of any type of research misconduct or plagiarism, and all the results will be represented appropriately and accurately.

7. Estimated start date and duration of project

Start date: Pilot study, 17th May 2022; Duration, three weeks.

Start date: Main study: 5th September 2022; Duration: six weeks

<p>Section B: I consider that this project may have ethical implications that should be brought before the Institute's Ethics Committee. <input type="checkbox"/></p>	
<p>Please state the total number of participants that will be involved in the project and give a breakdown of how many there are in each category e.g., teachers, parents, pupils etc.</p>	
<p>Give a succinct description of the aims and the methods (participants, instruments, and procedures) of the project in up to 500 words.</p> <ol style="list-style-type: none"> 1. Title of project 2. Purpose of project and its academic rationale 3. Brief description of methods and measurements 4. Participants: Recruitment methods, number, age, gender, exclusion/inclusion criteria 5. Consent and participant information arrangements, debriefing (attach forms where necessary) 6. A clear and concise statement of the ethical considerations raised by the project and how you intend to deal with them. 7. Estimated start date and duration of project 	

RISK ASSESSMENT

Brief outline of Work/activity:	Before the study, participants need to take an English listening comprehension test and a questionnaire with three sections on English listening self-efficacy, listening anxiety and self-regulated L2 listening as the pre-test. During the research, participants need to engage in various informal digital learning of English listening (IDLEL) activities. Meanwhile, they are required to record their engagement experiences in e-logs every week and submit the e-logs to the researcher once a week (i.e., main study for 4 weeks; pilot study for 1 week). After the study, participants need to take the listening comprehension test and the questionnaire again as the post-test.
Where will data be collected?	Pilot study: A Chinese university located in Hunan Province, China Main study: A Chinese university located in Hunan Province, China (different from the one in the pilot study)
Significant hazards:	Firstly, potential hazards may exist in terms of the data privacy. Since the listening comprehension tests will be in the form of written tests. In the case of the pilot study, as the study will be conducted remotely, the teacher will need to collect the papers and send them to the researcher with photographs. This process may lead to the disclosure of information

	on the papers. In addition, e-logs need to utilize electronic devices and networks. If the devices are damaged, or if network security is attacked, the content of the logs might be breached. Secondly, the main study is expected to take place in September. The university where the study will take place is located in the south of China, where the climate is hot, thus the researchers as well as the participants may feel physically uncomfortable. Thirdly, the researchers need to fly back to China from the UK to conduct the main study, and there is a possible risk of infecting the COVID-19 along the way. Fourthly, several universities in China have recently experienced outbreaks of COVID-19 infections. It's hard to ensure that the target university will not have similar situations during the period of pilot study. In addition, the researcher may be at risk of contracting the virus through contact with students and teachers while conducting the main research.
--	---

Who might be exposed to hazards?	The researcher; participants
----------------------------------	------------------------------

Existing control measures:	Firstly, the researcher will take a close look at the data protection measures of the software that will be used in the study and regularly check the computer equipment for viruses to protect the electronic equipment. For pilot study, the researcher will ask the teacher to take photographs of the test papers and send them to the researcher for data preservation as soon as they are collected and will ask the teacher to seal the papers pending the researcher's processing. Secondly, the researcher will take care to protect myself against viruses and ensure my health during the research process as well as during the overseas travel. Meanwhile, the researcher will be keeping a close eye on the pandemic outbreak in China. If there is an outbreak in the target university, the researcher will adjust either the format of the listening tests or the time of the study to suit the situation.
----------------------------	---

Are risks adequately controlled:	Yes
----------------------------------	-----

If NO, list additional controls and actions required:	Additional controls	Action by:

Section C: SIGNATURE OF APPLICANT

Note: a signature is required. Typed names are not acceptable.

I have declared all relevant information regarding my proposed project and confirm that ethical

good practice will be followed within the project.

Signed: Print Name: Huining Yang Date: 6/13/25

STATEMENT OF ETHICAL APPROVAL FOR PROPOSALS SUBMITTED TO THE
INSTITUTE ETHICS COMMITTEE

This project has been considered using agreed Institute procedures and is now approved.

Signed: Print Name: Date:

(IoE Research Ethics Committee representative) *

* A decision to allow a project to proceed is not an expert assessment of its content or of the possible risks involved in the investigation, nor does it detract in any way from the ultimate responsibility which students/investigators must themselves have for these matters. Approval is granted on the basis of the information declared by the applicant.

Data Protection Declaration for Ethical Approval (PhD/EdD projects)

This document can be used to provide assurances to your ethics committee where confirmation of data protection training and awareness is required for ethical approval.

By signing this declaration, I confirm that:

- I have read and understood the requirements for data protection within the *Data Protection for Researchers* document located here:
<https://www.reading.ac.uk/imps/-/media/49b402bbe9a74ae59dd8f4f080652123.ashx>
- I have asked for advice on any elements that I am *unclear on* prior to submitting my ethics approval request, either from my supervisor, or the data protection team at: imps@reading.ac.uk
- I understand that I am responsible for the secure handling, and protection of, my research data.
- I know who to contact in the event of an information security incident, a data protection complaint or a request made under data subject access rights.

Researcher to complete

Project / Study Title:

NAME	STUDENT ID NUMBER	DATE
Huining Yang		6/13/25

Supervisor signature

Note for supervisors: Please verify that your student has completed the above actions

NAME	STAFF ID NUMBER	DATE
Suzanne Graham		6/13/25

Submit your completed signed copy along with the other documents pertaining to the ethics application.

Copies to be retained by ethics committee.

VERSION	KEEPER	REVIEWED	APPROVED BY	APPROVAL DATE
1.0	IMPS	Annually	IMPS	

Information Sheet and Consent Form (Students)

Research Project: An investigation of second language learners' listening within an informal learning context

Project Team Members: Huining Yang; Supervisors: Professor Suzanne Graham and Dr. Pengchong Zhang

Dear Student

We would like to invite you to take part in a research study we are undertaking.

What is the study?

The worldwide spread of the epidemic in recent years has significantly increased the demand for technology in education. Technology-assisted language learning has quite a long research tradition but with less focus on informal digital learning of English listening (IDDEL). This research aims to investigate second language learners' listening within an informal learning context.

Why have I been chosen to take part?

The focus of this research is second-year undergraduates majoring in English in China and you meet those criteria. In addition, your faculty has expressed an interest in taking part in this kind of research. Therefore, I'd be glad to invite you to take part in this research.

Do I have to take part?

It is entirely up to you whether you give your consent to participate. You may also withdraw your consent to engagement at any time during the project, without any repercussions to you, by contacting Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

What will happen if I take part?

The study is expected to be conducted from early September 2022 and will last for a total of six weeks. In week 1, participants will be asked to complete an English listening test and an online questionnaire about how you feel different aspects of English listening (in Chinese) outside your normal lesson time. During the following four weeks (week 2-week 5), participants can freely choose their IDDEL activities and will need to briefly describe the way they engage in IDDEL activities in their e-logs (in English) every week. Participants will be asked to send their e-logs to the researcher by email at the start of the next week (e.g., the e-logs for week 2 need to be sent to the researcher on Monday of week 3). In week 6, all participants will be asked to take a second English listening test and complete the questionnaire as before. In week 19, you will be asked to take a third English listening test.

What are the risks and benefits of taking part?

The information given by participants in the study will remain confidential and will only be seen by the research team listed at the start of this letter. Neither you nor the university will be identifiable in any published report resulting from the study. Information about individuals will not be shared with the University. As for potential benefits, students who take part in the study will receive feedback from the researcher on their English writing regarding e-logs and obtain learning materials for developing their English listening at the end of the study.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. You will be assigned a pseudonym and will be referred to by the pseudonym in all records. The records will be sorted securely in password-protected files on a non-shared PC and all paper documentation will be kept in a locked filing cabinet and only the researcher will have access to the records. The data collected in this study will be rigorously analysed to answer a series of research questions.

In line with the University's policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. All anonymised research data will be retained indefinitely whereas any identifying information such as consent forms will be disposed of securely after the research findings have been written up. The results of the study will be presented at national and international conferences, and in written reports and articles. We can send you electronic copies of these publications if you wish.

Who has reviewed the study?

This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

What happens if I change my mind?

You can change your mind at any time without any repercussions. If you change your mind after data collection has ended, we will discard your data.

What happens if something goes wrong?

In the unlikely case of concern or complaint, you can contact Professor Suzanne Graham, Tel: +44 (0) 118 378 2684, E-mail: s.j.graham@reading.ac.uk

Where can I get more information?

For more detailed information, please feel free to contact Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

If you are happy to take part, please complete and return to Huining Yang the attached consent form.

Yours faithfully
Huining Yang

data protection for information sheets

The organisation responsible for protection of your personal information is the University of Reading (the Data Controller). Queries regarding data protection and your rights should be directed to the University Data Protection Officer at imps@reading.ac.uk, or in writing to: Information Management & Policy Services, University of Reading, Whiteknights, P O Box 217, Reading, RG6 6AH.

The University of Reading collects, analyses, uses, shares, and retains personal data for the purposes of research in the public interest. Under data protection law we are required to inform you that this use of the personal data we may hold about you is on the lawful basis of being a public task in the public interest and where it is necessary for scientific or historical research purposes. If you withdraw from a research study, which processes your personal data, dependant on the stage of withdrawal, we may still rely on this lawful basis to continue using your data if your withdrawal would be of significant detriment to the research study aims. We will always have in place appropriate safeguards to protect your personal data.

If we have included any additional requests for use of your data, for example adding you to a registration list for the purposes of inviting you to take part in future studies, this will be done only with your consent where you have provided it to us and should you wish to be removed from the register at a later date, you should contact Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

You have certain rights under data protection law which are:

- Withdraw your consent, for example if you opted in to be added to a participant register
- Access your personal data or ask for a copy
- Rectify inaccuracies in personal data that we hold about you
- Be forgotten, that is your details to be removed from systems that we use to process your personal data
- Restrict uses of your data
- Object to uses of your data, for example retention after you have withdrawn from a study

Some restrictions apply to the above rights where data is collected and used for research purposes.

You can find out more about your rights on the website of the Information Commissioners Office (ICO) at <https://ico.org.uk>

You also have a right to complain the ICO if you are unhappy with how your data has been

handled. Please contact the University Data Protection Officer in the first instance.

Research Project:

Please complete and return this form to: Huining Yang

1. I have read the information sheet about the project and received a copy of it.
2. I understand what the purpose of the study is and what you want me to do. All my questions have been answered.
3. I agree to take part in the research.
4. I agree to take two listening tests in the research.
5. I agree to complete questionnaires in the research.
6. I agree to complete e-logs and send them in time to the researcher in the research.

Name:

Signed:

Date:

Information Sheet and Consent Form (Teachers)

Research Project: An investigation of second language learners' listening within an informal learning context

Project Team Members: Huining Yang; Supervisors: Professor Suzanne Graham and Dr. Pengchong Zhang

Dear Teacher

We would like to invite your students to take part in a research study we are undertaking.

What is the study?

The worldwide spread of the epidemic in recent years has significantly increased the demand for technology in education. Technology-assisted language learning has quite a long research tradition but with less focus on informal digital learning of English listening (IDDEL). This research aims to investigate second language learners' listening within an informal learning context.

Why has my students been chosen to take part?

The focus of this research is second-year undergraduates majoring in English in China and your students meet those criteria. In addition, your faculty has expressed an interest in taking part in this kind of research. Therefore, I'd be glad to invite your students to take part in this research.

Does my students have to take part?

It is entirely up to you whether your students take part in this study. If you are happy for your students to take part, please return the consent form to Huining Yang. You may also withdraw your consent to their engagement at any time during the project, without any repercussions by contacting Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

What will happen if my students take part?

The study is expected to be conducted from early September 2022 and will last for a total of six weeks. In week 1, participants will be asked to complete an English listening test and an online questionnaire about how your students feel different aspects of English listening (in Chinese) outside your normal lesson time. During the following four weeks (week 2-week 5), participants can freely choose their IDDEL activities and will need to briefly describe the way they engage in IDDEL activities in their e-logs (in English) every week. Participants will be asked to send their e-logs to the researcher by email at the start of the next week (e.g., the e-logs for week 2 need to be sent to the researcher on Monday of week 3). In week 6, all participants will be asked to take a second English listening test and complete the questionnaire as before. In week 19, participants will be asked to take a third English listening test.

What are the risks and benefits of taking part?

The information given by your students in the study will remain confidential and will only be seen by the research team listed at the start of this letter. Neither your students nor the university will be identifiable in any published report resulting from the study. Information about individuals will not be shared with the University. As for potential benefits, students who take part in the study will receive feedback from the researcher on their English writing regarding e-logs and obtain learning materials for developing their English listening at the end of the study.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The data collected in this study will be rigorously analysed to answer a series of research questions.

In line with the University's policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. All anonymised research data will be retained indefinitely whereas any identifying information such as consent forms will be disposed of securely after the research findings have been written up. The results of the study will be presented at national and international conferences, and in written reports and articles. We can send you electronic copies of these publications if you wish.

Who has reviewed the study?

This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

What happens if I change my mind?

You can change your mind at any time without any repercussions. If you change your mind after data collection has ended, we will discard the data collected from your students.

What happens if something goes wrong?

In the unlikely case of concern or complaint, you can contact Professor Suzanne Graham, Tel: +44 (0) 118 378 2684, E-mail: s.j.graham@reading.ac.uk

Where can I get more information?

For more detailed information, please feel free to contact Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

If you are happy for your students to take part, please complete and return to Huining Yang the attached consent form.

Yours faithfully
Huining Yang

data protection for information sheets

The organisation responsible for protection of your personal information is the University of Reading (the Data Controller). Queries regarding data protection and your rights should be directed to the University Data Protection Officer at imps@reading.ac.uk, or in writing to: Information Management & Policy Services, University of Reading, Whiteknights, P O Box 217, Reading, RG6 6AH.

The University of Reading collects, analyses, uses, shares, and retains personal data for the purposes of research in the public interest. Under data protection law we are required to inform you that this use of the personal data we may hold about you is on the lawful basis of being a public task in the public interest and where it is necessary for scientific or historical research purposes. If you withdraw from a research study, which processes your personal data, dependant on the stage of withdrawal, we may still rely on this lawful basis to continue using your data if your withdrawal would be of significant detriment to the research study aims. We will always have in place appropriate safeguards to protect your personal data.

If we have included any additional requests for use of your data, for example adding you to a registration list for the purposes of inviting you to take part in future studies, this will be done only with your consent where you have provided it to us and should you wish to be removed from the register at a later date, you should contact For more detailed information, please feel free to contact Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

You have certain rights under data protection law which are:

- Withdraw your consent, for example if you opted in to be added to a participant register
- Access your personal data or ask for a copy
- Rectify inaccuracies in personal data that we hold about you
- Be forgotten, that is your details to be removed from systems that we use to process your personal data
- Restrict uses of your data
- Object to uses of your data, for example retention after you have withdrawn from a study

Some restrictions apply to the above rights where data is collected and used for research purposes.

You can find out more about your rights on the website of the Information Commissioners Office (ICO) at <https://ico.org.uk>

You also have a right to complain the ICO if you are unhappy with how your data has been handled. Please contact the University Data Protection Officer in the first instance.

Research Project:

Please complete and return this form to: Huining Yang

1. I have read the information sheet about the project and received a copy of it.
2. I understand what the purpose of the study is and what is required of my students and me. All my questions have been answered.
3. I agree to my students take part in the research.
4. I agree to my students take two listening tests in the research.
5. I agree to my students complete questionnaires in the research.
6. I agree to my students complete e-logs and send them in time to the researcher in the research.

Name:

Signed:

Date:

Information Sheet and Consent Form (Head of the faculty)

Research Project: An investigation of second language learners' listening within an informal learning context

Project Team Members: Huining Yang; Supervisors: Professor Suzanne Graham and Dr. Pengchong Zhang

Dear Head of the faculty

We would like to invite your students to take part in a research study we are undertaking.

What is the study?

The worldwide spread of the epidemic in recent years has significantly increased the demand for technology in education. Technology-assisted language learning has quite a long research tradition but with less focus on informal digital learning of English listening (IDDEL). This research aims to investigate second language learners' listening within an informal learning context.

Why has my students been chosen to take part?

The focus of this research is second-year undergraduates majoring in English in China and you meet those criteria. In addition, your faculty has expressed an interest in taking part in this kind of research. Therefore, I'd be glad to invite your students to take part in this research.

Does my students have to take part?

It is entirely up to you whether your students take part in this study. If you are happy for your students to take part, please return the consent form to Huining Yang. You may also withdraw your consent to their engagement at any time during the project, without any repercussions by contacting Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

What will happen if my students take part?

The study is expected to be conducted from early September 2022 and will last for a total of six weeks. In week 1, participants will be asked to complete an English listening test and an online questionnaire about how your students feel different aspects of English listening (in Chinese) outside your normal lesson time. During the following four weeks (week 2-week 5), participants can freely choose their IDDEL activities and will need to briefly describe the way they engage in IDDEL activities in their e-logs (in English) every week. Participants will be asked to send their e-logs to the researcher by email at the start of the next week (e.g., the e-logs for week 2 need to be sent to the researcher on Monday of week 3). In week 6, all participants will be asked to take a second English listening test and complete the questionnaire as before. In week 19, participants will be asked to take a third English listening test.

What are the risks and benefits of taking part?

The information given by your students in the study will remain confidential and will only be seen by the research team listed at the start of this letter. Neither your students nor the university will be identifiable in any published report resulting from the study. Information about individuals will not be shared with the University. As for potential benefits, students who take part in the study will receive feedback from the researcher on their English writing regarding e-logs and obtain learning materials for developing their English listening at the end of the study.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The data collected in this study will be rigorously analysed to answer a series of research questions.

In line with the University's policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. All anonymised research data will be retained indefinitely whereas any identifying information such as consent forms will be disposed of securely after the research findings have been written up. The results of the study will be presented at national and international conferences, and in written reports and articles. We can send you electronic copies of these publications if you wish.

Who has reviewed the study?

This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

What happens if I change my mind?

You can change your mind at any time without any repercussions. If you change your mind after data collection has ended, we will discard the data collected from your students.

What happens if something goes wrong?

In the unlikely case of concern or complaint, you can contact Professor Suzanne Graham, Tel: +44 (0) 118 378 2684, E-mail: s.j.graham@reading.ac.uk

Where can I get more information?

For more detailed information, please feel free to contact Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

If you are happy for your students to take part, please complete and return to Huining Yang the attached consent form.

Yours faithfully
Huining Yang

data protection for information sheets

The organisation responsible for protection of your personal information is the University of Reading (the Data Controller). Queries regarding data protection and your rights should be directed to the University Data Protection Officer at imps@reading.ac.uk, or in writing to: Information Management & Policy Services, University of Reading, Whiteknights, P O Box 217, Reading, RG6 6AH.

The University of Reading collects, analyses, uses, shares, and retains personal data for the purposes of research in the public interest. Under data protection law we are required to inform you that this use of the personal data we may hold about you is on the lawful basis of being a public task in the public interest and where it is necessary for scientific or historical research purposes. If you withdraw from a research study, which processes your personal data, dependant on the stage of withdrawal, we may still rely on this lawful basis to continue using your data if your withdrawal would be of significant detriment to the research study aims. We will always have in place appropriate safeguards to protect your personal data.

If we have included any additional requests for use of your data, for example adding you to a registration list for the purposes of inviting you to take part in future studies, this will be done only with your consent where you have provided it to us and should you wish to be removed from the register at a later date, you should contact For more detailed information, please feel free to contact Huining Yang, Tel: 18163979527, E-mail: ct823443@student.reading.ac.uk

You have certain rights under data protection law which are:

- Withdraw your consent, for example if you opted in to be added to a participant register
- Access your personal data or ask for a copy
- Rectify inaccuracies in personal data that we hold about you
- Be forgotten, that is your details to be removed from systems that we use to process your personal data
- Restrict uses of your data
- Object to uses of your data, for example retention after you have withdrawn from a study

Some restrictions apply to the above rights where data is collected and used for research purposes.

You can find out more about your rights on the website of the Information Commissioners Office (ICO) at <https://ico.org.uk>

You also have a right to complain the ICO if you are unhappy with how your data has been handled. Please contact the University Data Protection Officer in the first instance.

Research Project:

Please complete and return this form to: Huining Yang

1. I have read the information sheet about the project and received a copy of it.
2. I understand what the purpose of the study is and what is required of my students and me. All my questions have been answered.
3. I agree to my students take part in the research.
4. I agree to my students take two listening tests in the research.
5. I agree to my students complete questionnaires in the research.
6. I agree to my students complete e-logs and send them in time to the researcher in the research.

Name:

Signed:

Date: