



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Language Education
English Language Teaching Program

THE IMPACT OF METACOGNITIVE INSTRUCTION ON EFL LEARNERS'
LISTENING COMPREHENSION SKILL

Kübra YETİŞ

Master's Thesis

Ankara, (2021)



With leadership, research, innovation, high quality education and change,

To the leading edge... Toward being the best...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Department of Foreign Language Education

English Language Teaching Program

THE IMPACT OF METACOGNITIVE INSTRUCTION ON EFL LEARNERS'
LISTENING COMPREHENSION SKILL

ÜSTBİLİŞSEL EĞİTİMİN YABANCI DİL OLARAK İNGİLİZCE ÖĞRENENLERİN
DİNLEME ANLAMA BECERİSİ ÜZERİNE ETKİSİ

Kübra YETİŞ

Master's Thesis

Ankara, (2021)

Acceptance and Approval

To the Graduate School of Educational Sciences,

This thesis, prepared by **Kübra YETİŞ** and entitled “**The Impact of Metacognitive Instruction on EFL Learners’ Listening Comprehension Skill,**” has been approved as a thesis for the Degree of **Master** in the **Program of English Language Teaching** in the **Department of Foreign Language Education** by the members of the Examining Committee.

Chair	Prof. Dr. İsmail Hakkı MİRİCİ	Signature
Member (Supervisor)	Assist. Prof. Dr. İsmail Fırat ALTAY	Signature
Member	Assoc. Prof. Dr. Ayhan KAHRAMAN	Signature

This is to certify that this thesis/dissertation has been approved by the aforementioned examining committee members on .../.../.... in accordance with the relevant articles of the Rules and Regulations of Hacettepe University Graduate School of Educational Sciences, and was accepted as a **Master’s Thesis** in the **Program of English Language Teaching** by the Board of Directors of the Graduate School of Educational Sciences from/...../.....

Prof. Dr. Selahattin GELBAL
Director of Graduate School of Educational Sciences

Abstract

The study explored the impact of an 8-week process-based metacognitive instruction with pedagogical procedures on listening comprehension performance and metacognitive awareness of 37 EFL listeners in Turkey regarding different listening proficiency levels and preparatory school backgrounds. A mixed-methods research design with a quasi-experimental study, including a pretest-posttest and an experimental-control group, was adopted in this study. The quantitative data were collected through the Metacognitive Awareness Listening Questionnaire (MALQ) to measure metacognitive awareness of listening before, during, and after metacognitive instruction, and an IELTS listening test as pretest and posttest to measure listening comprehension performance. The qualitative data was collected through stimulated recall protocols from the experimental group during and after the instruction to determine any changes in their metacognitive awareness over metacognitive instruction. Results showed that the experimental group showed a statistically significant increase in their listening performance on the final test to a medium extent. However, their metacognitive awareness of listening showed a non-significant increase at the end of the study, and this was also analyzed regarding the subdimensions of the MALQ. Besides, the experimental group had a significantly higher level of overall metacognitive awareness than the control group to a medium extent. Furthermore, more- and less-successful L2 listeners and preparatory school backgrounds made considerable differences in the study. Also, the quantitative findings on metacognitive awareness of listening were supported by some qualitative findings. Implications and suggestions for future research on learning and teaching L2 listening with the process-based metacognitive instruction were lastly discussed.

Keywords: EFL, L2 listening, listening comprehension skill, metacognitive instruction, metacognitive awareness of listening, process-based approach.

Öz

Bu çalışma; 8 haftalık, süreç odaklı eğitsel süreçleri benimsemiş üstbilişsel eğitimin Türkiye’de İngilizce’yi yabancı dil olarak öğrenen 37 öğrencinin dinleme anlama başarısını ve dinleme becerisi üstbilişsel farkındalıklarına etkisini farklı dinleme yeterlilik seviyesine ve hazırlık okulu deneyimine göre incelemiştir. Karışık desenli araştırma tasarımı ve öntest, sontest, deney ve kontrol grubu içeren yarı-deneyssel çalışma yöntemi benimsenmiştir. Nicel veri, ikinci dil dinleyicilerinin dinleme becerisi üstbilişsel farkındalıklarını ölçmek için Vandergrift ve arkadaşları (2006) tarafından tasarlanan Dinleme Becerisi Üstbilişsel Farkındalık Ölçeği (MALQ) aracılığıyla çalışmanın başında, ortasında ve sonunda, ve dinleme anlama performanslarını ölçmek için IELTS dinleme öntesti ve sontesti aracılığıyla toplanırken üstbilişsel eğitim süreci boyunca katılımcıların ölçeğe verdikleri cevaplardaki değişiklikleri keşfetmek üzere çalışmanın ortasında ve sonunda deney grubu ile gerçekleştirilen uyarılmış-geri çağırma protokolleri ile nitel veri toplanmıştır. Nicel veri bulgularına göre, deney grubu sontestte dinleme performanslarında istatistiksel olarak anlamlı orta etki düzeyinde fark gösterirken, MALQ altboyutları açısından da analiz edilmiş olan, genel dinleme becerisi üstbilişsel farkındalıklarındaki çalışma sonunda gözlenen artış orta düzeyde olup anlamlı bulunmamıştır. Bunun yanında, çok ve az başarılı dinleyiciler ile hazırlık okulu deneyiminin çalışmanın sonuçlarını etkilediği saptanmıştır. Aynı zamanda, dinleme becerisi üstbilişsel farkındalık üzerine elde edilen nicel veri sonuçları bazıları nitel veri sonuçları desteklenmiştir. İngilizce’yi yabancı dil olarak öğrenen dinleyiciler için süreç odaklı eğitsel süreçleri benimsemiş üstbilişsel eğitimin mevcut çalışma üzerindeki çıkarımları ve gelecekte bu konuda yapılacak çalışmalara dair tavsiyeler de tartışılmıştır.

Anahtar sözcükler: Yabancı dil olarak İngilizce, dinleme anlama becerisi, ikinci dil dinleme becerisi, üstbilişsel eğitim, dinleme becerisi üstbilişsel farkındalık, süreç odaklı yaklaşım.

Acknowledgements

This three-year journey could not be gone through without the support of some people. I would like to thank all of them.

First of all, I owe my gratitude to my supervisor Assist. Prof. Dr. İsmail Fırat ALTAY for his encouragement and help during the process of writing my thesis.

Also, I would like to thank Prof. Dr. İsmail Hakkı MİRİCİ and Assoc. Prof. Dr. Ayhan KAHRAMAN as my thesis committee members for their valuable comments and inspiring opinions made for the thesis.

Besides, I owe my gratitude to my mother and father, and my sisters for their patience and understanding during this process. I would like to thank my friends, my colleagues for their continuous support and reminding me of my unfinished thesis waiting for me during the difficult days with Covid19, “Yes, Kübra, it is not so much left, keep your social distance, just sit down and write it, come on!” Also, I would like to thank my professors and students at Düzce University for their help and support while conducting my thesis.

Last but not least, I would love to owe my gratitude to my dear teacher, the late Assoc. Prof. Dr. Hüseyin Öz, as my pre-supervisor and the first person in the academic world I trusted in after my undergraduate education, for guiding me in deciding the topic of my thesis and showing me the way that I have been going through after him. I always remember him as a guiding light in my academic life.

After all, I would like to thank Küb to write this thesis and present it one day, filled up with all feelings of gratitude, confidence, and mindfulness.

Table of Contents

Abstract	ii
Öz	iii
Acknowledgements	iv
List of Tables	vii
List of Figures	x
Symbols and Abbreviations	xi
Chapter 1 Introduction	1
Statement of the Problem	1
Aim and Significance of the Study	6
Research Questions	8
Assumptions	8
Limitations	9
Definitions	9
Chapter 2 Literature Review	11
Metacognition	11
Listening	13
Metacognition and Listening	17
Related Studies	21
Chapter 3 Methodology	35
Research Design	35
Setting and Participants	40
Data Collection	42
Instruments	47
Data Analysis	53
Reliability and Validity of the Study	55
Chapter 4 Findings	59

Quantitative findings.....	59
Qualitative Findings.....	120
Summary of Qualitative Findings	147
Chapter 5 Conclusion, Discussion and Suggestions	149
Conclusion and Discussion on Listening Comprehension Performance of L2 Listeners over Metacognitive Instruction	149
Conclusion and Discussion on Metacognitive Awareness of Listening of L2 Listeners over Metacognitive Instruction	156
Overall Conclusion of the Study	187
Implications of the Study	191
Suggestions for Future Research	196
References	197
APPENDIX-A: Consent Form	207
APPENDIX-B: Weekly Listening Texts.....	208
APPENDIX-C: Sample Page of Listening Diary	209
APPENDIX-D: The IELTS Listening Test	211
APPENDIX-E: Metacognitive Awareness Listening Questionnaire (MALQ)	217
APPENDIX-F: Sample Questions of Stimulated Recall Protocols	220
APPENDIX-G: Ethics Committee Approval	221
APPENDIX H: Declaration of Ethical Conduct.....	222
APPENDIX-I: Thesis/Dissertation Originality Report	223
APPENDIX-J: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı	224

List of Tables

Table 1 <i>Three Types of Metacognitive Knowledge about L2 Listening and Examples (Vandergrift & Goh, 2012, p. 87-88)</i>	18
Table 2 <i>Metacognitive Processes for Listening (Vandergrift, 2004, p.11; Vandergrift and Goh, 2012, p. 109).</i>	21
Table 3 <i>Research Design of the Study</i>	38
Table 4 <i>Gender and Participants</i>	42
Table 5 <i>Level of Listening Proficiency and Participants</i>	42
Table 6 <i>Preparatory School Background and Participants</i>	42
Table 7 <i>Definition of the Factors in the MALQ (Vandergrift et al., 2006, p. 450-451)</i>	49
Table 8 <i>The Reliability Analysis of the MALQ (Vandergrift et al., 2006, p.446)</i>	50
Table 9 <i>Data Collection Instruments</i>	53
Table 10 <i>Descriptive Statistics of the Pretest and Posttest Scores by the Groups</i>	59
Table 11 <i>Wilcoxon Signed Ranks Tests on Pretest and Posttest within the Groups</i>	59
Table 12 <i>Mann-Whitney U Test on Pretest and Posttest between the Groups</i>	60
Table 13 <i>Descriptive Statistics of the Pretest and Posttest by the Groups regarding Level of Listening Proficiency</i>	61
Table 14 <i>Wilcoxon Signed Ranks Tests on Pretest and Posttest within the Groups regarding Level of Listening Proficiency</i>	61
Table 15 <i>Mann-Whitney U Test on Pretest and Posttest within the Groups regarding Level of Listening Proficiency</i>	63
Table 16 <i>Mann-Whitney U Test on Pretest and Posttest between the Groups regarding Level of Listening Proficiency</i>	65
Table 17 <i>Descriptive Statistics of the Pretest and Posttest by the Groups regarding Preparatory School Backgrounds</i>	66
Table 18 <i>Wilcoxon Signed Ranks Tests on Pretest and Posttest within the Groups regarding Preparatory School Background</i>	66
Table 19 <i>Mann-Whitney U Test on Pretest and Posttest within the Groups regarding Preparatory School Background</i>	67
Table 20 <i>Mann-Whitney U Test on Pretest and Posttest between the Groups regarding Preparatory School Background</i>	69

Table 21 <i>Descriptive Statistics of the MALQ for the Groups</i>	70
Table 22 <i>Friedman Test on the MALQ within the Groups</i>	71
Table 23 <i>Wilcoxon Signed Ranks Tests on Person Knowledge by the Experimental Group</i>	73
Table 24 <i>Wilcoxon Signed Ranks Tests on the Overall MALQ for the Control Group</i>	74
Table 25 <i>Wilcoxon Signed Ranks Tests on Problem Solving for the Control Group</i>	75
Table 26 <i>Descriptive Statistics of the MALQ between the Groups</i>	76
Table 27 <i>Mann-Whitney U Test on the MALQ between the Groups</i>	77
Table 28 <i>Descriptive Statistics of the MALQ for the Experimental Group regarding Level of Listening Proficiency</i>	81
Table 29 <i>Friedman Test on the MALQ within the Experimental Group regarding Level of Listening Proficiency</i>	82
Table 30 <i>Descriptive Statistics of the MALQ for the Control Group regarding Level of Listening Proficiency</i>	83
Table 31 <i>Friedman Test on the MALQ within the Control Group regarding Level of Listening Proficiency</i>	85
Table 32 <i>Wilcoxon Signed Ranks Tests on Overall MALQ for Less-Successful Listeners in the Control Group</i>	86
Table 33 <i>Wilcoxon Signed Ranks Tests on Problem Solving for Less-Successful Listeners in the Control Group</i>	87
Table 34 <i>Mann-Whitney U Test on the MALQ for the Experimental Group regarding Level of Listening Proficiency</i>	89
Table 35 <i>Mann-Whitney U Test on the MALQ for the Control Group regarding Level of Listening Proficiency</i>	92
Table 36 <i>Mann-Whitney U Test on the MALQ Scores of Less-Successful L2 listeners between the Groups</i>	94
Table 37 <i>Mann-Whitney U Test on the MALQ Scores of More-Successful L2 Listeners between the Groups</i>	97
Table 38 <i>Descriptive Statistics of the MALQ for the Experimental Group regarding Preparatory School Background</i>	100
Table 39 <i>Friedman Test on the MALQ within the Experimental Group regarding Preparatory School Background</i>	101

Table 40 <i>Descriptive Statistics of the MALQ for the Control Group regarding Preparatory School Background</i>	103
Table 41 <i>Friedman Test on the MALQ within the Control Group regarding Preparatory School Background</i>	104
Table 42 <i>Wilcoxon Signed Ranks Tests on Mental Translation for the Control Group regarding Different Preparatory School Background</i>	106
Table 43 <i>Wilcoxon Signed Ranks Tests of Overall MALQ for the Control Group regarding Preparatory School Background</i>	107
Table 44 <i>Mann-Whitney U Test on the MALQ of the Experimental Group regarding Preparatory School Background</i>	108
Table 45 <i>Mann-Whitney U Test on the MALQ for the Control Group regarding Preparatory School Background</i>	112
Table 46 <i>Mann-Whitney U Test on the MALQ between the Groups for L2 listeners who studied at the Preparatory School Before</i>	115
Table 47 <i>Mann-Whitney U Test on the MALQ between the Groups for L2 listeners who did not study at the Preparatory School Before</i>	117
Table 48 <i>Descriptive Information of the Participants of Stimulated Recall Protocols</i>	120
Table 49 <i>Summary of Qualitative Findings</i>	148

List of Figures

<i>Figure 1.</i> A metacognitive framework for L2 listening teaching (Vandergrift & Goh, 2012, p. 85)	18
<i>Figure 2.</i> Positive changes from first stimulated recall protocol.....	125
<i>Figure 3.</i> Negative changes from the first stimulated recall protocol	133
<i>Figure 4.</i> Positive changes from the second stimulated recall protocol	141
<i>Figure 5.</i> Negative changes from the second stimulated recall protocol	146
<i>Figure 6.</i> Suggestions on L2 listening comprehension success	147



Symbols and Abbreviations

DA: Directed Attention

EFL: English as a Foreign Language

ESL: English as a Second Language

ELT: English Language Teaching

IELTS: International English Language Testing System

L1: First Language

L2: Second Language

MALQ: Metacognitive Awareness Listening Questionnaire

MT: Mental Translation

SPSS: Statistical Package for Social Sciences

PE: Planning and Evaluation

PK: Person Knowledge

PS: Problem Solving

Chapter 1

Introduction

The research intends to explore the effect of metacognitive instruction on EFL learners' listening comprehension skills. The current research problem statement with the support of its background is given in this chapter. It is followed by the presentation of the aim and significance of the study. After that, the research questions are listed. Finally, the study's assumptions, limitations, and definitions used in the present study are presented.

Statement of the Problem

During a listening course, it is familiar to meet this setting (Goh, 2010): A teacher at a primary school or an instructor at a university starts their listening course by introducing a listening text by asking learners to tell themselves what they had knowledge of the text topic. The learners then talk about their ideas in class and discuss possible unknown vocabularies they may meet. Later, the teacher/the instructor asks them to carefully examine the listening task instructions to find out what is asked on the listening text. After that, they listen to the listening text so as to complete the activity. The teacher/the instructor then makes them listen to it again to check their answers or complete the missing parts. After all, they answer the listening task's questions altogether, or the teacher/ the instructor tells the learners the correct answers. Therefore, they see their correct and incorrect answers about the listening text, and the listening course is dismissed. Is this scene typical in the listening classes you have met so far, the reader of this study?

The emphasis above is on the extent to which the listeners' answers are correct; what is done above is to solely test their listening comprehension without teaching how to facilitate their listening comprehension, assuming that those know how to listen effectively (Goh, 2010). In this kind of listening course, the listening product is the focus; the process of listening, which leads to that product, takes place very little or is not even mentioned (Field, 2019).

Also, when analyzed L2 listening coursebooks or materials, it is apparent that pre-listening activities serve the aim of activating prior knowledge of L2 listeners on the text while post-listening activities focus on what is understood in the text and on

completing grammar, pronunciation, or vocabulary activities based on what they understood from the text. Here the matter is again the product of listening, not engaging in the process of listening. Under these circumstances, it is a general expectation that L2 listeners show more remarkable development in their L2 listening performance over their language learning process (Goh, 2010).

It is also evident that listening has an important place in daily life; listening is used more than the other three language skills (Mendelsohn, 1994). It is also the most used second language learning skill (Scarcella & Oxford, 1992). Rost (2001) points out that listening has been accepted as essential for language learning. It was used to present grammatical forms via sample dialogues in the early times of ELT; despite its importance, listening has been neglected and underestimated until the late 1970s in which it started to be taught for communication in the ELT field (Field, 2008). Nunan (2002, p. 238) considered listening as “the Cinderella skill” among the others because it has been disregarded by “its elder sister-speaking” and seen as a means to speak and write in the language, which was defined as knowing the language by that time.

Listening has also been regarded as a passive activity; nonetheless, it consists of a process which is active and complex that listeners decode speech sounds, comprehend vocabulary and grammatical forms, and interpret stress, intonation, and all available contextual knowledge while bringing and tapping their background knowledge into the process of listening, which needs listeners to involve a good deal of mental process (Vandergrift, 1999). Altuwaresh (2016) claims that this complex listening process makes the listening skill an ignored and overlooked language skill by teachers and researchers.

Given the complex nature of listening, L2 listeners’ listening comprehension could be affected by some factors such as background knowledge about text topic, vocabulary, grammatical and contextual knowledge, speech rate, accent, input type, language ability, psychological situation, time available, environmental factors, attention and concentration issues, and therefore misunderstanding and misinterpretation of the listening text because of those factors (Underwood, 1989; Goh, 1999, 2000; Graham, 2006).

About this, Goh (2000) investigated listening comprehension problems L2 learners experienced based on a cognitive perspective of Anderson (1995), including perception (segmenting speech sounds and phonemes), parsing (recognizing of words), and utilization (interpreting the intended message through applying knowledge sources in the long-term memory) processes via listening dairies, retrospective, and small group interviews. L2 learners experienced problems in the perception phase such as not identifying the known words, ignoring the rest when considering the meaning, not chunking the parts of speaking, losing track of the beginning part of the text, focusing on listening too hard. They had problems in the parsing phase, such as immediately forgetting what they just heard, failing to compose a reflection of the word just heard in mind, and not comprehending the next part of the input due to previous problems. Furthermore, they had trouble in the utilization phase, for example, comprehending the words but not the message meant and not ensuring the importance of the input's critical parts. Goh (2000) examined the results in terms of different levels of listening ability. According to the results, both higher and lower-skilled listeners had a problem understanding the words they already knew and immediately forgot what they just heard and comprehended, which were two of the most frequent listening problems they had. High-skilled listeners also experienced problems understanding the message, despite understanding the words, because of the lack of contextual or background information. Besides, less-skilled listeners had difficulty understanding the rest due to focusing excessively on certain parts of the text they just listened to because of the lack of vocabulary knowledge. However, the higher ones did not experience this problem because they kept going on listening using a strategy called directed attention, which was one of the metacognitive strategies emphasizing staying on the task. She finally suggested two teaching strategies for improving L2 listeners' listening comprehension performance: direct strategy including cognitive, metacognitive, and socio-affective tactics along with perception activities on sounds, words selected, pronunciation or intonation, and indirect strategy with metacognitive awareness-raising activities.

Considering all this, how do listeners manage and control the process of listening effectively despite its complex cognitive nature and the difficulties listeners experience during this complex listening process mentioned above? Vandergrift and

Goh (2012) answer this question as listeners with a high-proficiency level of listening are capable of managing these processes through using “metacognitive knowledge” consisting of “person, task, and strategy knowledge,” in other words, it means the awareness and knowledge of the cognitive processes regarding themselves as listeners, listening tasks, listening strategies used, and the potential to see, regulate, control these processes (Goh, 2008). The theoretical concept of “metacognition” that indicates “thinking about thinking” (Flavell, 1979) appears here and forms this basis of understanding. The study of Vandergrift et al. (2006) explains that the reason for almost thirteen percent of the difference in the listeners’ listening comprehension performance might be metacognition. Therefore, the listeners who are metacognitively aware of the listening process apply their metacognitive knowledge to plan, monitor, solve problems in listening, and evaluate the listening process to comprehend what they listen to wholly. Accordingly, how do listeners use metacognitive knowledge to manage the listening process?

According to Goh (2000, 2010), because L2 listeners have a necessary to notice the listening process and learn how to listen, they need to develop better metacognitive awareness of listening through metacognitive awareness-raising activities. To do this, Vandergrift and Tafaghodtari (2010) proposed a process-based metacognitive instruction in listening, including pedagogical procedures increasing their metacognitive awareness of the process, regulating and appraising their comprehension and overall learning process, and therefore helping them become more autonomous and self-regulated learners (Wenden, 1988).

Vandergrift and Goh (2012) list this metacognitive view to listening’ aims, which are to improve learners who: a) comprehend the difficulties of L2 listening; b) have an idea about their development of learning on an individual and collaborative basis with others; c) have a habit which is to plan and control their own listening process; d) use listening strategies in an appropriate way; e) have growing motivation and self-efficacy, and developing listening to process the spoken language and involve in oral interaction effectively. It can be summarized that through this metacognitive approach to listening, L2 listeners become autonomous, who are conscient and responsible for their learning process, and use several strategies to overcome the challenges they experience and meet their needs in listening in different contexts.

The researchers also made a group of listening strategies using different classifications of listening strategies listed by different researchers (O'Malley & Chamot, 1990; Vandergrift, 1997, 2003a; Goh, 1998, 2002; Young, 1997 as cited in Vandergrift & Goh, 2012, p. 90). The listening strategies are a) to enable to process and interpret the oral input; b) oversee how the information is processed; c) take action to direct and control the cognitive processes; d) guide listeners' feelings; e) engage others or learning sources to help in comprehension and learning.

Accordingly, how does this metacognitive approach to listening and listeners' metacognitive awareness affect the conclusion of listening comprehension? Listeners who have enough metacognitive knowledge can make themselves ready, check, and assess their comprehension of what they listen to; it differentiates them from listening to the listening tasks randomly or incidentally. How L2 learners perceive their learning straight-forwardly influences their learning process and product (Vandergrift, 2004). Their knowledge of a listening task's difficulties can guide them to choose, evaluate, change, or ignore goals, tasks, and strategies (Goh, 2010). The answer might be that it positively influences their attitudes towards listening and learning to listen.

As it is proved that metacognition is a predictor of learning and key to success (Vandergrift, 1999), it is crucial to make L2 learners become aware of the process and increase their metacognitive knowledge to facilitate listening. Several studies have highlighted that metacognitive awareness and listening comprehension skill of L2 listeners could increase via metacognitive instruction in language learning classroom (e.g., Vandergrift, 2002, 2003b; Vandergrift & Tafaghodtari, 2010, Goh & Taib, 2006; Liu & Goh, 2006; Cross, 2009; Rahimi & Katal, 2012a; Katrancı & Yangın, 2012, Mareschal, 2007; Zeng, 2007; Goh & Hu, 2013; Altuweish, 2016; Graham & Macaro, 2008; Maftoon & Alamdari, 2020).

The present study has been motivated by this pedagogical process-based metacognitive approach to listening supported by Vandergrift and Goh's earlier studies (Vandergrift, 1998, 1999, 2000, 2004; Goh, 1997, 2000, 2005, 2008) based on Flavell (1979) and Wenden (1998)'s approach to metacognition. The current research has also regarded the study of Vandergrift and Tafaghodtari (2010) as the seminal study. Regarding the literature of the field of English language teaching, the research on metacognitive instruction into L2 listening is limited by some crucial

studies (e.g., Vandergrift & Tafaghodtari 2010; Mareschal, 2007; Cross, 2009; Graham & Macaro, 2008; Goh & Taib, 2006; Bozorgian, 2014, 2015; Rahimi & Katal, 2012a, 2012b, 2013; Zeng, 2012; Goh & Hu, 2013; Altuwairesh, 2016, Bozorgian & Alamdari, 2017; Maftoon & Alamdari, 2020). The studies show that teaching how to listen through a process-based metacognitive instruction facilitates listening in EFL and ESL contexts. Considering the literature in Turkey, few studies have focused on that area. Çevikbaş (2016) and Coşkun (2010) are two significant studies in the EFL context in Turkey concerning this approach. In Çevikbaş's study (2016), metacognitive instruction was given in L1 to the preparatory school students, while in Coşkun's research (2010), the duration of metacognitive instruction to the preparatory school students was only five weeks. However, no reports questioning the level of metacognitive awareness and listening comprehension performance of L2 listeners majoring in the ELT department after receiving an 8-week process-based metacognitive instruction in listening was found in the literature. Regarding those situations, the current study will shed light on the literature of the ELT field.

Aim and Significance of the Study

The paper aims to investigate the effect of an 8-week process-based metacognitive instruction with pedagogical procedures on EFL listeners' metacognitive awareness of listening as well as listening comprehension performance in Turkey. First of all, it explores to what extent metacognitive instruction affects EFL listeners' listening comprehension performance. Secondly, it intends to demonstrate to what extent metacognitive instruction impacts their metacognitive awareness of listening. In addition, it attempts to question whether their listening performance and metacognitive awareness show any differences over metacognitive instruction between L2 listeners with different levels of listening proficiency and preparatory school backgrounds.

Concerning its significance, the present study could make several critical contributions to the ELT field. First of all, it has been the first research in which 8-week process-based metacognitive listening instruction has been given the ELT students in Turkey. Thus, it has offered more reliable and deeper insights into the L2 listening area in Turkey and different contexts. In this way, metacognitive

instruction to listening could be questioned with particular points of view, contributing to L2 listening teaching and learning area.

Moreover, it could give pedagogical implications for language teachers, policymakers, curriculum designers, teacher educators, and researchers about teaching how to listen through a process-based metacognitive instruction in order to solve the problems of listeners who do not know how to listen but know what to listen for, which is very common and creates the problem of this study. It could also shed light on them to show L2 listeners how to become more-autonomous and self-regulated learners by becoming metacognitively aware of the listening process and using their metacognitive knowledge. Besides, it could show what kind of comprehension problems EFL listeners meet in a different context and how to deal with those problems in a classroom environment by teachers, learners within pedagogical procedures.

Besides, it presents significant methodological implications for the field. It differentiates from the seminal study of Vandergrift and Tafaghodtari (2010) in two aspects: in the study, at the end of the lessons, a transcript was given to L2 listeners, which helped them to engage in the bottom-up processes of listening, recognizing words and their pronunciations, but the instruction length was shorter, lasting eight weeks. Except for Vandergrift and Tafaghodtari (2010)'s seminal study, studies on this subject based on a mixed-method research design are quite limited. Additionally, many earlier related studies have only been quantitative studies (Bozorgian, 2014; Rahimi & Katal, 2012a, 2012b, 2013; Goh & Hu, 2013, Alamdari & Maftoon, 2020; Bozorgian & Alamdari, 2017) with different grade levels, age, and gender groups, proficiency levels, the length of the instruction. In comparison, much earlier studies (Chamot & Küpper, 1989; Bacon, 1992; Thomas & Rubin, 1996; Murphy, 1985; Vandergrift, 2002; O'Malley, Chamot, & Küpper, 1989) have been based on qualitative research as well, but they adopted different methodology on strategy instruction concerning different cognitive and/or metacognitive strategies from the seminal study. Above all, however, in this study, the quantitative data on listening comprehension performance and metacognitive awareness could be supported by the qualitative data on the perceptions of L2 listeners on metacognitive awareness of listening so as to elaborate and validate the findings with each other.

About the rationales and the point of view of the study, this research is not only significant but also necessary in the EFL context in Turkey.

Research Questions

The paper addresses the research questions below concerning the gap in the research literature:

1. To what extent does metacognitive instruction affect the listening comprehension performance of L2 listeners?
 - 1.1. Are there any statistically significant differences in L2 listeners' listening comprehension performance over metacognitive instruction regarding the level of listening proficiency?
 - 1.2. Are there any statistically significant differences in the listening comprehension performance of L2 listeners regarding preparatory school background?
2. To what extent does metacognitive instruction affect the metacognitive awareness of listening of L2 listeners?
 - 2.1. Are there any statistically significant differences in L2 listeners' metacognitive awareness of listening over metacognitive instruction regarding the level of listening proficiency?
 - 2.2. Are there any statistically significant differences in L2 listeners' metacognitive awareness of listening over metacognitive instruction regarding preparatory school background?
 - 2.3. What are the perceptions of L2 listeners on possible changes in metacognitive awareness of listening over metacognitive instruction?

Assumptions

The present study, a quasi-experimental study based on a mixed-method research design, has some assumptions regarding data collection and its instruments. Firstly, the data collection instruments are assumed to be suitable for collecting data on L2 listeners' metacognitive awareness of listening and listening comprehension performance. Secondly, all of the research participants are

accepted to give sincere and honest responses to the instruments. Also, the necessary permission has been gotten from the owners for the instruments used; therefore, it is accepted that there were no problems to use them. Besides, the participants who were the first-grade ELT students at a Turkish state university are assumed to be at least in the B2 level of general English proficiency. Finally, the participants are assumed to represent the majority of the study's target population.

Limitations

As for the research limitations, the participants and setting of the study were the major limitations. The participants were limited to 37 first-grade ELT students at a Turkish state university. The research could be administered with more participants in different contexts. Also, because the data was not normally distributed, the participants' assignment into the treatment and control groups could be another limitation. Thirdly, the instruction length is only eight weeks, which is one of the limitations of this study. More extended instruction periods could be achieved in other contexts. Lastly, the researcher as an instructor gave the 8-week metacognitive instruction to both groups with the aim of preventing the possible undesirable effects because of different instructors with different implementation on the study; however, the classes and the instructor were not observed by another person during the instruction; that could be considered as a limitation of the study.

Definitions

Metacognition. It refers to “one’s knowledge concerning one’s own cognitive processes and ... active monitoring and consequent regulation and orchestration of these processes in relation to the cognitive objects or data on which they bear, usually in the service of some concrete goal or objective” (Flavell, 1976, p. 232).

Metacognitive Instruction. It is identified as “pedagogical procedures that enable learners to increase awareness of the listening process by developing richer metacognitive knowledge about themselves as listeners, the nature and demands of listening, and strategies for listening as well as to plan, monitor, and evaluate their comprehension efforts and the progress of their overall listening development” (Vandergrift & Goh, 2012, p. 97).

Listening. It is defined as “the active and dynamic process of attending, perceiving, interpreting, remembering, and responding to the expressed (verbal and nonverbal), needs, concerns, and information offered by other human beings” (Purdy, 1997, p. 8).

Listening Comprehension. It refers to “an active and conscious process in which the listener constructs meaning by using cues from contextual information and from existing knowledge, while relying upon multiple strategic resources to fulfill the task requirements” (O’Malley, Chamot & Küpper, 1989, p. 434).



Chapter 2

Literature Review

The literature review on metacognition and listening with definitions of listening, listening processes, listening difficulties, and metacognitive approach to listening, metacognitive instruction is presented in this part. Later, earlier studies related to the current study and these fundamental concepts are given.

Metacognition

Flavell (1976) is the leading researcher who identified the concept of metacognition as the capability to think about thinking, monitor, and regulate what is in mind while thinking. According to Flavell (1979, p. 906), monitoring and regulating the cognitive processes in the mind takes place in the interaction with four phenomena of “metacognitive knowledge,” “metacognitive experiences,” “goals (tasks),” and “actions (strategies).” An attempt to thinking’s aims is called goals; the steps taken to accomplish those aims are regarded as actions or strategies; one’s knowledge of cognitive processes and beliefs on factors affecting those processes are defined as metacognitive knowledge, including three knowledge types about cognition: “person, task, and strategy knowledge”; aware mental and emotional practices during thinking are called as metacognitive experience (Flavell, 1979, p. 906).

As mentioned above, among the three types of metacognitive knowledge, firstly, person knowledge is knowledge about a person himself as a “cognitive processor” and factors affecting their thinking (Flavell, 1979). Also, it is of the one’s consciousness of their mental processes (Oz, 2005) with three categories: intraindividual differences (beliefs about abilities of oneself), interindividual differences (beliefs about abilities of others), and universal of cognition. Secondly, task knowledge is knowledge about a task’s features, purposes, demands, information about a task’s procedures, and difficulty a person has. Thirdly, strategy knowledge is knowledge about what strategies help effectively accomplish a goal and complete a task successfully (Flavell, 1979).

According to Paris and Winograd (1990), this approach of Flavell (1976, 1979) reflects two important characteristics of metacognition: “self-appraisal” and “self-regulation” of thinking. Self-appraisal consists of individual reflections related

to the knowledge, capabilities, judgments about the individual's abilities, tasks, or strategies, and it answers the questions about what knowledge they have, the way they think, when and why they use their knowledge or strategies (Paris & Winograd, 1990). Self-regulation is about how metacognition organizes and controls the cognitive aspects, which is about preparing, checking, and appraising the processes.

Metacognition and Language Learning. According to Flavell (1979, p. 906), metacognition has an essential role in cognitive tasks relating to languages such as speaking, listening and reading comprehension, writing, language acquisition, memory, several kinds of self-management, and self-instruction.

The concept of metacognition was first used in the L2 learning area by Wenden (1987, 1991, & 1998). Wenden (1998) explained metacognitive knowledge and its components in L2 learning and distinguished between metacognitive knowledge and strategies. According to the researcher, person knowledge was meant what a learner had knowledge of their particular features helping or harming their learning process such as age, language aptitude, motivation, self-efficacy beliefs about their capability of learning and of achieving a specific goal in language learning; task knowledge was defined as what knowledge a learner had of the listening task's goal, nature, and demands; strategic knowledge was identified as knowledge a learner had of what strategies were useful, when and how to use in the process of learning. While metacognitive knowledge was what learners were aware of their learning, metacognitive strategies were actions learners take to plan, check, regulate, evaluate, and guide their learning (Wenden, 1998).

Besides, Wenden (1998) briefly indicates that metacognitive knowledge helps reach learning outcomes, enables recall, comprehension of listening texts, achieves learning, and develops the rate of learning progress. Furthermore, Wenden (1987) points out the role of metacognition in learning and expresses that it helps develop learner autonomy and shows the differences in each learner's cognitive processes. Also, Wenden (1991) suggests a different aspect of good language learners and asserts that those who are metacognitively aware of themselves are self-controlled and responsible for their learning process. It is summarized that they are autonomous learners and aware of their strengths and weaknesses; they use strategies when in trouble during learning, they are a self-

regulated learner; they regulate their learning process. All Wenden's works have pioneered other researchers to study metacognition in learning and teaching language skills, especially listening (Vandergrift & Goh, 2012).

Listening

Definition of Listening. Several definitions of listening have been made for different situations by different researchers. According to Underwood (1989), listening is the task of attending and attempting to understand what a person hears. Mendelson (1994) explains listening as a skill to get meaning from the speaking of native speakers. Purdy (1997) offers another definition for listening as an active and dynamic process, including attention, perception, interpretation, and response to speakers' existing reaction.

Rost (2002, p.1) states that its definition has changed in various ways in time: in the 1900s, it was defined regarding a reliable recording process of human brain signals; around 1920s and 1930s, through more knowledge reached on the brain, listening was identified as a process, not consciously held, but managed by uncertain cognitive systems; in the 1940s, due to the improvement in telecommunications, it was defined as a better way of transmitting or composing the messages; in the 1950, with the computational sciences' advances, it was called as segmenting and labeling the input to store and retrieve it efficiently; in the 1960s, it was known as hearing and feeling both speakers and listeners' feelings due to the appearance of transpersonal psychology; in the 1970s, with the development of globalism and anthropology, its definition was as bringing about well-accepted cultural experience; in the 1980s, it was identified as skills of and actions made by a listener to be active; in the 1990s, it was described as "input processing" due to developments in computer technology; in the 2000s, with the occurrence of digital networking, listening included the concept of a network where people actively connected with each other and organized various activities in a fast and efficient way. Rost (2002) summarizes that this entire change has shown the shift in our thoughts about what we could do via listening. Moreover, Rost (2002) defines listening within four aspects: a) receptive as receiving what the speaker expresses while paying attention to, hearing, comprehending, and keeping in mind aural symbols; b) constructive as constructing and finding meaning the speaker means

which is interesting for you; c) collaborative as negotiating what the speaker means with the speaker interpersonally and giving response; d) transformative as generating meaning through involving with, understanding, showing empathy with the speaker.

Listening Comprehension. A complex process where listeners actively differentiate between sounds, comprehend words and syntactic forms, clarify intonation and stress, and keep in mind what is understood from those above and explain it using exiting contextual background or related prior knowledge on the input is a definition of Vandergrift (1999) for listening comprehension.

According to Buck (2001), it is a skill of understanding the broader meanings of the oral input automatically and of real-time comprehending the linguistic knowledge in the text very clearly and inferencing the intended message in an exact way.

Listening comprehension is a procedure that listeners consciously use cues available in the context and from their background and create meaning based on several strategic sources to accomplish the task (O'Malley, Chamot, & Küpper, 1989). Similarly, a multifaceted process that consists of getting vocabularies and expressions in speech, clarifying their meanings through applying syntactic knowledge, creating an understanding through vocabularies from background knowledge, giving responses properly and through oral communication, and putting the information into memory for recalling and using it next time is an identification of Goh (2014) for it.

Studies on listening mentioned above have focused on similar points: listening comprehension is a many-sided, conscious, interactive, and meaning-making procedure in a sociocultural context where a listener tries to attend to sounds, understand vocabularies, and interpret the intended message by using available contextual knowledge and their background knowledge.

Significance of Listening in Language Learning. Listening has a crucial role in daily communication: it composes “40-50%; speaking, 25-30%; reading, 11-16%; and writing, about 9%” according to Mendelsohn (1994, p. 9). It is considered as a language skill that is practiced most in language classrooms (Nunan, 1998; Scarcella & Oxford, 1992; Rost, 2001). Its significance is emphasized by Rost

(2002) as a) for it enables the input in the language classrooms, and learning comes out when the input is understood at the right point, it is crucial; b) speaking and listening are the means of interaction and listeners need to interact with speakers to understand the message; c) oral language is also a difficulty for the learner to understand natives of the language; d) listening comprehension tasks help teachers make learners aware of the forms of grammar and vocabulary, and communication models in the language. Vandergrift (1997) asserts that listening incorporates the language's rules within itself and enables learners to speak, write, and read in the language.

Process of Listening. Considering the definitions above, it is obvious that it consists of a conscious and multifaceted process. During the listening comprehension process, listeners go through two kinds of process: bottom-up processing, taking the sounds into meaningful forms where listeners decode the meaningful units from sounds to phonemes, phonemes to words, words to sentences while constructing the meaning; and top-down processing, applying the contextual and world knowledge to making inferences on the listening text (Mendelsohn, 1994; Vandergrift, 2004; Oxford, 1993; Rost, 2002).

While listeners hear sounds, keep them in their working memory to connect them, and try to understand what they just heard, listeners use their prior knowledge and contextual cues to construct the meaning intended in what they heard. The top-down processing could be described as "listener as an active model builder" while the bottom-up processing could be defined as "listener as a tape-recorder." (Anderson & Lynch, 1988 as cited in Seferoğlu & Uzakgören, 2004). The purpose of listening, learners' characteristics, and the context of listening text affect the extent to which the process is basically used (Vandergrift & Goh, 2012). For example, a listener who needs to scan the text for specific information is involved in mostly bottom-up processing, while a listener who attempts to understand the main point of the spoken language deals with more top-down processing.

Besides, Anderson (1995) offered a cognitive model explaining the process of listening in three stages to provide a better insight into making meaning while top-down and bottom-up processing during listening. Considering this model, listeners recognize speech sounds, hesitations, and intonations in the perceptual stage, keeping them in the working memory. Listeners give their attention to the text and

ignore the rest of the sounds around them; they become aware of pauses, the stress in the stream of sound related to a specific language; and classify them regarding the language categories. As a result, listeners decode the spoken language. This process is the first stage of segmenting the words and includes bottom-up processing. In Anderson (1995)'s this model, in the parsing stage, which also involves bottom-up processing, listeners take the decoded phonetic representations of the input in the working memory into parts and bring out possible words from the long-term memory, using some points such as phonotactic features, word choices; thus, listeners create a representational idea to keep the meaningful representations of these potential words in their working memory. According to this model, listeners use non-linguistic knowledge to comprehend the message in the utilization stage. Listeners develop the parsed knowledge in detail and check their clarification by correspondence with their background knowledge and the representations of the intended message in memory by using contextual, pragmatic knowledge, and listeners' schemata. The utilization stage includes top-down processing. In that stage, listeners link the expressive patterns with knowledge types in their long-term memory to deduce the pointed implication (Anderson, 1995).

Listening Difficulties. As it is understood, listening is a relatively complex process. Goh (2000) studied the difficulties of listening comprehension related to the three-stage listening process of Anderson (1995) and found out ten problems including that: in the perceptual stage, listeners had trouble in recognizing words, ignoring parts of speech following, parting the sound stream, not catching the beginning of a phrase or a sentence, and in concentrating on the text; in the parsing stage, listeners easily ignored what they just perceived, could not constitute a cognitive depiction with vocabularies they listened to, and comprehend the rest of the text due to what listeners missed before; in the utilization stage, listeners did not comprehend the intended message; they comprehended only the words, and become disoriented because of available inconsistency in the message.

Goh (1999) also investigated the factors influencing learners' listening comprehension and found out that vocabularies in the text, background knowledge about the theme of the text, rate of speech, accent of speakers, type of input affected their listening. Graham (2006) added poor grammatical knowledge, limited

vocabularies, and misconstruction of the text's meaning due to the difficulties learners have during listening.

Underwood (1989) explained a list of obstacles preventing listeners from listening efficiently: first, listeners could not regulate the speed of speech; second, listeners did not always have a chance to make words repeated; third, they had limited vocabulary knowledge; fourth, they could not recognize the points such as transitions, repeating a part; fifth, they had a lack of contextual knowledge; sixth, they could have trouble in concentrating in another language such as distraction for not even a second from the text; last, listeners could tend to want to comprehend all words they hear, and the possibility of missing any of them could make them worried.

Metacognition and Listening

In the study of Vandergrift et al. (2006), it was emphasized that metacognition could justify almost thirteen percent of the differences in listening achievement. Thus, listeners who metacognitively know the listening process prepare themselves for listening, control their comprehending and solve their understanding problems, and assess their understanding process in listening. How do listeners use their metacognitive knowledge to manage the process of listening? It is essential to understand Vandergrift and Goh's (2012) metacognitive approach to listening.

Metacognitive Approach to Listening. In the first part of Chapter 2, metacognition is explained with Flavell's (1979) and Wenden's (1991) description. They point out that metacognition includes metacognitive knowledge, classified as person, task, strategy knowledge, metacognitive experiences, tasks, and strategies. Vandergrift and Goh (2012) created a structure for teaching and learning second language listening with these four concepts (see Figure 1).

In light of this metacognitive framework, metacognitive experience is regarded as a conscious feeling a person has at a specific moment (Flavell, 1979). For example, when listeners are aware that they have no idea of a word meaning while listening, they try to remember a strategy, which they did before to solve a similar problem. It is beneficial to listeners when they produce and apply strategies and understand the task, themselves, or the world better. In Figure 1, two arrows

show that metacognitive experience affects the metacognitive knowledge and the preference of strategy use.

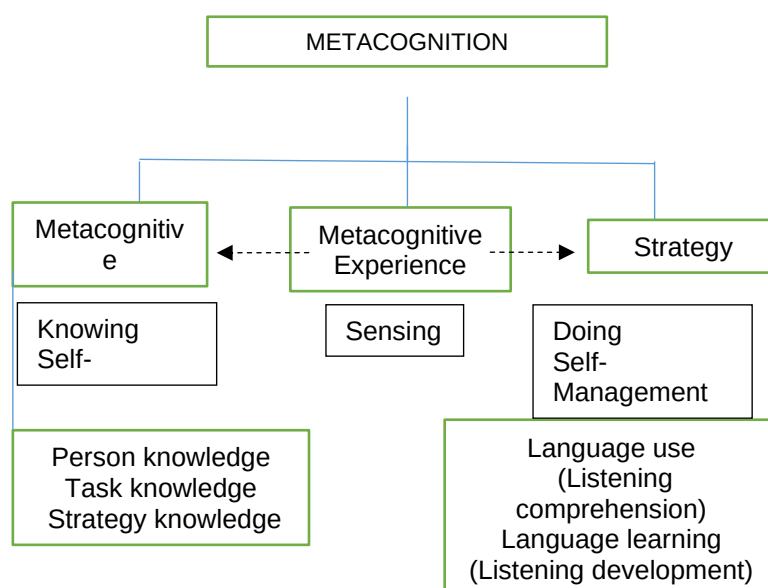


Figure 1. A metacognitive framework for L2 listening teaching (Vandergrift & Goh, 2012, p. 85)

In Table 1, the three types of metacognitive knowledge are explained regarding second language listening, and some examples related to them are summarized by Vandergrift's (2002, 2003b) and Goh's (2002) studies together with Wenden (1991) and Flavell (1979).

Table 1

Three Types of Metacognitive Knowledge about L2 Listening and Examples
(Vandergrift & Goh, 2012, p. 87-88)

Types	Examples for L2 listening
"Person Knowledge": Knowledge of the mental and affective variables that affect listening comprehension performance of listeners.	"Self-concept and self-efficacy about listening" • I am a nervous listener. If I try to do my best, I can develop my listening Certain listening difficulties, their reasons, and potential answers. • I experience some troubles understanding the vocabularies if they are said too fast.

“Task Knowledge”: Knowledge of the listening task’s aim and features, difficulties, and knowledge of when conscious attention is needed.	Cognitive, emotional, and interactive listening processes • You need to focus on the task if you are not good at that kind of activity. Abilities to carry out a listening task • I know you can understand the topic if you listen to it carefully Factors that influence listening • His accent is too fast, and it makes listening difficult for me. How to develop listening out of the class • I think I should listen to some podcasts for different accents.
“Strategy Knowledge”: Knowledge of fruitful listening strategies and of the way to apply them appropriately at the right time	Strategies to ease understanding and control learning • When you miss what you listen to, make a guess. Strategies for certain listening tasks • To understand the numbers, you need to listen to them carefully Strategies which do not work effectively • Do not concentrate too much on the text; it makes you stressed.

Strategy use is of applying appropriate strategies to accomplish tasks and goals when needed. It is of taking specific actions to provide learners with learning more efficiently, more effectively, and more autonomously (Vandergrift & Goh, 2012). Learners who know when and how to use strategic knowledge effectively tend to use strategies (Zhang & Goh, 2006). Furthermore, learning strategies directly or indirectly facilitates language learning (O’Malley & Chamot, 1990; Oxford, 1993; Vandergrift, 1997; Goh, 2002).

Metacognitive Instruction. A process-based methodology including pedagogical procedures that help listeners develop their metacognitive awareness of listening by increasing their metacognitive knowledge in terms of themselves as a listener, the task they involve in, and the strategies they use; at the same time, they learn to organize and prepare themselves by planning, check and assess their comprehension and their development on second language listening skill is called as metacognitive instruction (Vandergrift, 2004). It is useful since it is learner-oriented and enables listeners to take part in the evaluation and regulation process of their own learning by activities scaffolded by teachers, and it also helps teachers and learners oversee the complex and unseen listening processes while focusing on the listening task in order to notice learners’ learning styles and skills as an

individual (Goh, 2010). According to Veenman et al. (2006), effective metacognitive processes integrate metacognitive intervention into classroom listening activities, including listeners in considering and learning the way of listening.

Metacognitive Process. L2 listening metacognitive instruction includes pedagogical approaches to growing listeners' awareness of listening. Following processes are used to make listeners more successful through metacognitive instruction (Vandergrift, 2004; Goh, 2010):

Planning. It is the stage in which listeners make themselves ready for what they listen to and what they are supposed to do during listening. Listeners try to activate their background knowledge on the listening text, examine the text's genre, and predict what they hear, the text's main idea, and details based on the contextual cues.

Monitoring. It is the second stage in which listeners monitor their listening comprehension based on their predictions. Listeners check what they understand, confirm what they understood, and determine what needs to be done for comprehending the text. Also, assessing their progress during listening and deciding if their approach works or not are related to it.

Problem Solving. It is the third stage where listeners check and fit their methods to understand the text and apply appropriate strategies to overcome their difficulties. Listeners change their approach accordingly, use appropriate strategies, check and infer about the parts they do not get or ask for an explanation on that part.

Evaluation. It is the final stage of the metacognitive processes in which listeners evaluate if their approach is useful or not during listening. They make reflections about the difficulties they confronted, what they did wrong, and their success in listening comprehension, and confirm their achievement on their choices about strategies or problem-solving methods.

According to Vandergrift (2004), these metacognitive processes operate interactively with each other, not in a linear way. It depends on the needs that listeners have during the processes of listening.

Metacognitive Pedagogical Sequence. Metacognitive instruction aims to enable learners to control listening comprehension and eventually succeed in L2

listening. Table 2 shows the instructional phases and metacognitive procedures of listening instruction (Vandergrift, 2004, p.11; Vandergrift and Goh, 2012, p. 109).

Table 2

Metacognitive Processes for Listening (Vandergrift, 2004, p.11; Vandergrift and Goh, 2012, p. 109).

Stages of Pedagogical Listening Instruction	Metacognitive Processes (Metacognitive Strategies)
1. Planning/predicting stage (Pre-listening) After listeners find out the listening text's topic and genre, they make predictions about what kind of information and potential vocabularies they could hear.	1.Planning & Directed attention
2. First verifications stage (First listening)	
a. Listeners verify initial hypotheses, correct as required, and note additional information understood.	2a. Selective Attention, Monitoring, and Evaluation
b. Listeners compare what they have noted with their pair, adjust as required, decide what parts needs still need more attention	2b. Monitoring, Evaluation, and Planning & Selective Attention
3. Second verification stage (Second listening)	
a. Listeners verify the parts that need resolution after the first listening, correct their notes and understanding and take additional notes that are just understood.	3a. Selective Attention, Monitoring, Evaluation, and Problem Solving
b. A whole-class discussion on the main and related parts of the text and reflections on how listeners achieved their understanding of certain parts occur.	3b. Monitoring, Evaluation, and Problem Solving
4. Final verification stage (Third listening) Students listen to the text for the parts they could not understand earlier during the discussion. It can be fulfilled by following a transcript of the text.	4.Selective Attention, Monitoring, and Problem Solving
5. Reflection and Goal-setting stage Considering the earlier discussion of the strategies they used to make up for the parts not understood, listeners write their goals for future listening tasks.	5. Evaluation and Planning

Related Studies

For more than two decades, several researchers have focused on studying on metacognitive approach to listening instruction. Several important studies related to metacognition, metacognitive knowledge and strategies, metacognitive instruction regarding second language listening skills in distinct settings with listening proficiency levels are briefly mentioned.

Murphy (1985) was one of the first researchers who studied the listening strategies used by ESL listeners with different proficiency levels, but he did not aim to find out which strategy they used most. It was demonstrated that more-successful learners preferred to apply different strategies more, such as inferring, predicting actively, and that they dealt with top-down processes rather than bottom-up processes. In contrast, less-successful listeners preferred to focus on the text and related textual processes so much that it took a long time to elaborate on what they hear. Murphy did not list any categories as metacognitive or cognitive strategies but suggested that listening proficiency made a difference regarding using listening strategies. As a result, Murphy (1985) suggested that it was necessary to guide ESL listeners to experience those different strategies while focusing on the listening process by peer-supported and teacher-supported techniques rather than only answering the questions following listening to the text.

Among the first researchers emphasizing listening comprehension using metacognitive strategies, especially monitoring, Henner Stanchina (1987) showed that successful listeners used predicting, evaluating, and problem-solving strategies during listening, such as applying their world knowledge to make guesses related to the task and interpret the meaning through what they know and understand; assessing and reviewing their interpretations when needed. According to the researcher, successful listeners tended to listen and consistently monitor themselves by using their world knowledge while listening.

In another study on using listening strategies concerning different listening proficiency levels through think-aloud procedures, that listening proficiency level created a difference in the use of the strategies among listeners was emphasized by Chamot and Küpper (1989), as Murphy (1985) did. The quantitative results showed few differences in monitoring and inferencing but greater differences in selective attention and evaluation between more-successful and less-successful high-school listeners with an intermediate level. At the same time, the qualitative findings claimed that more-successful listeners used those strategies persistently and purposefully. They concluded that more-skilled college students applied much more metacognitive strategies such as monitoring understanding and identifying the problems and mostly integrated them with inferencing and elaboration, which are among cognitive strategies.

One other study, including a think-aloud methodology, by O'Malley, Chamot, and Küpper (1989), was based on high-school ESL listeners' strategy use with an intermediate level. They asserted that more-effective listeners tended to prepare themselves for what to focus on during listening, stay focused, keep going on listening, and immediately recover concentration when lost. Besides, they preferred to interact with the text through much more top-down or utilization processes such as using their world or personal knowledge, asking themselves questions about the text they listened to or using contextual information to help their comprehension. Less-effective listeners tended to give up listening when they met unknown words or phrases, interacted with the text using the bottom-up processes like word recognition, and did not much integrate what they knew about the text with what they understood. According to the quantitative findings, more-effective listeners preferred to use monitoring, using their world knowledge, and inferencing more than their counterparts.

The other study based on a think-aloud methodology about using listening strategies of university-level effective and less-effective L2 listeners was examined by Bacon (1992). The researcher deduced that listening comprehension success was related to applying and being open to trying different strategies, self-monitoring, staying motivated and focused while listening, and effectively using prior knowledge. According to this study, the proficiency level did not cause any differences in monitoring, but effective listeners assess their comprehension level more realistically.

The effect of metacognitive and cognitive strategy instruction on Russian college listeners' listening performance was investigated by Thompson and Rubin (1996). They concluded that over two-year strategy instruction, the listeners who received the training showed significant development in their listening performance and developed a regulative approach to listening using metacognitive strategies.

Vandergrift and Goh are among the leading researchers who researched metacognitive awareness and strategy use in second language listening for two decades.

Vandergrift (1997) studied listening strategies with high-school French students depending on the proficiency level through a think-aloud methodology.

According to the findings, the frequency of metacognitive strategy use increased as the listening proficiency level increased. Primarily, monitoring, problem identification, and selective attention differed in using between effective and less-effective listeners. While the latter used cognitive strategies like translation, repetition, transferring, they failed to control their use of metacognitive strategies to attend to the meaning. He suggested that in earlier years of language learning, listening should be taught within a framework in which the strategies were taught, employed, and discussed.

Goh (1997) examined beliefs and knowledge on listening of L2 listeners writing their beliefs on diaries and referred to the importance of increasing L2 listeners' metacognitive awareness by using diaries as a tool in the classroom because it could guide L2 listeners to think over their listening performance and the ways to improve it. According to the study, they reported the metacognitive knowledge types on their diaries. As for person knowledge, they identified cognitive processes and problems during listening and difficulties in listening comprehension and development; for task knowledge, they defined issues influencing listening comprehension, the types of input improving comprehension, second language nature; for strategic knowledge, they wrote strategies helping comprehension and recall, improving listening comprehension skill, and not working every time. Also, she suggested that more discussion on the process of listening within a metacognitive perspective in the classroom based on the curriculum of the course together and on their beliefs and strategy use about the listening task would make L2 listeners aware of the process of listening and of their own learning, and therefore become more autonomous learners. Besides, she recommended that metacognitive awareness of listening could be fostered by teaching L2 listeners how to plan during pre-listening activities, control their understanding during listening, and assess themselves considering their comprehension and strategy use with post-listening activities.

Vandergrift (1998) was also the one who examined listening strategies to help listening comprehension of L2 listeners with different levels of proficiency through a think-aloud methodology as a case study to find out how they constructed the meaning. He claimed that "listening is a selective process," what the listener selected for processing was about using metacognitive strategies and essential for

better comprehension. According to the results, monitoring the listening process was the primary metacognitive strategies leading to other metacognitive strategies such as guessing and attending selectively together with cognitive strategies such as using prior knowledge and making an inference; in other words, successful listeners monitored their comprehension by applying successful prediction strategies following successful elaboration and inferencing practices. Less-successful listeners tended to depend more on using their background knowledge to interpret the unknown rather than on the cues such as contextual or extralinguistic cues directing themselves the meaning to compose a framework helping comprehension on their mind. Also, they spent their time mostly on bottom-up strategies as well. However, instead of lower-level processes, successful listeners deal with the meaning by using top-down processes and metacognitive strategies.

Goh (1998) determined the sorts of cognitive and metacognitive strategies and tactics practiced by L2 listeners through retrospective protocols and diaries and found out that higher proficient L2 listeners applied more than the less-proficient L2 listeners, that cognitive strategies and the tactics were used much more than metacognitive strategies by both groups. Higher-proficient listeners used “selective attention, directed attention, comprehension monitoring, real-time assessment of input, comprehension evaluation,” while the less-proficient ones used these first three metacognitive strategies (p.13). The former also used “contextualization, fixation, reconstruction, inferencing, elaboration”; the latter mainly used the last two cognitive strategies. As for the tactics, which were specific actions, higher-proficient listeners used much more metacognitive tactics for dealing with the difficulties they met, such as continuing listening despite them, while the lower ones became worried when they met any unknown or difficult words and missed the rest of the test. Goh (1998) also concluded that higher-proficient listeners used top-down processes as well as bottom-up processes as in fixation and reconstruction. Finally, she suggested that through direct listening strategy training or awareness-raising activities, L2 listeners were guided to apply those strategies and tactics flexibly, developing their listening comprehension and having better metacognitive knowledge on learning how to listen.

Vandergrift (1999) highlighted how listeners could apply strategies so as to develop their learning process. He presented an existing research basis for raising

the listeners' metacognitive awareness about the listening procedures and metacognitive strategy use, including planning, monitoring, and evaluating through listening comprehension checklists.

Goh (1999) conducted a study on factors affecting the listening comprehension of Chinese L2 listeners and their metacognitive awareness of those factors. In the study, she examined one of the types of task knowledge she studied in her earlier study (1997) mentioned above, which was the factors affecting listening comprehension. She concluded that twenty factors related to five main factors: "listening task, listening text, listeners, environment, and speakers." The top five factors concluded were speech rate, input type, accent, vocabulary, and background knowledge. According to high-skilled listeners, listening was an interactive process in which the meaning was constructed together with listeners and speakers. In contrast, listening was mainly regarded as a text-oriented activity and textual factors such as vocabulary, rate of speech, and input type rather than environment and speaker, which were considered influencing listening comprehension for lower-skilled listeners. Listeners' point of interest for listening, their psychological and physical situation including anxiety, tiredness, their knowledge about the topic of the text, their ability to stay focused, visual-supported listening activities, the speaking ability of the speaker were among other factors affecting listening comprehension Goh (1999) concluded.

Vandergrift (2002) investigated developing metacognitive knowledge of L2 listening to studying with beginning-level French students by using listening comprehension activities with reflective tasks. They completed a survey to improve the listening and reflective exercises' qualities. Their responses were analyzed in terms of metacognitive knowledge and strategies of listening. It was concluded that reflective exercises caused to raise their metacognitive awareness of listening. They also benefited more strategic knowledge by using planning, monitoring and evaluation strategies, and task knowledge on factors influencing their listening comprehension but less person knowledge on themselves as an L2 listener. Reflective exercises helped students enhance their metacognitive knowledge of listening and their achievement on L2 listening comprehension as well as be motivated and have self-efficacy beliefs for their second language learning and

listening skills regarding the strategies' use, and therefore get them to be more autonomous L2 learners.

Vandergrift (2003a) studied the listening comprehension strategy treatment for two years to seven-grade more-proficient and less-proficient French listeners. He revealed that metacognitive strategies, including planning and monitoring, except for evaluation, were reported due to language proficiency through a think-aloud procedure. Also, more successful listeners practiced mostly metacognitive strategies, particularly comprehension monitoring, more questioning elaboration, which was related to being flexible and open to different potentials while interpreting the meaning, and less translation; those metacognitive strategies controlled the listening procedures and activated appropriate cognitive strategy use such as elaboration mostly. Less-successful listeners mostly used translation as a bottom-up process, which caused them not to keep in mind what they heard and create their mental representations because of the limited memory capacity, and therefore, not to understand it. Even though both used equally elaboration and inferencing strategies, more-successful listeners used them at a deeper level while the lower ones applied them superficially depending on the context at the discourse level. It was also noted that the former constructed the meaning within a constant metacognitive cycle where they interacted with the input and monitored it using their prior or contextual knowledge through elaboration, and made inferences about the meaning depending on that. He suggested that guiding listeners to use their metacognitive knowledge within a pedagogical cycle by planning, monitoring, and evaluating their listening process could enhance their listening comprehension performance and become self-regulated learners.

An 8-week metacognitive instruction to ten younger students studying at a primary school was examined by Goh and Taib (2006). Listening comprehension and reflective exercises with teacher-led discussions on their metacognitive awareness of listening took place in the study. When the instruction ended, the listeners showed an increasing level of metacognitive awareness of listening, feeling confident while engaging in listening comprehension tasks, and using their strategic knowledge while dealing with challenges while listening to as many adult L2 listeners. Also, the learners took advantage of a process-based metacognitive instruction on listening, especially the less successful listeners at most: they learned

to listen by dealing with the mental processes of listening actively rather than only answering the comprehension questions. However, because their knowledge of strategies was limited, it was suggested that explicit strategy training could be better for younger language learners.

Vandergrift (2005) investigated the interaction of metacognition, listening proficiency, and three types of motivation studying with French learners. He found a supportive relation between metacognitive strategy use and metacognitive knowledge with motivation for learning to listen. It was concluded that the proficiency level was not positively related to amotivation, but its relationship with intrinsic and extrinsic motivation was not found as higher as expected. Tafaghodtari and Vandergrift (2008) studied the contribution of L1 listening, L2 proficiency, metacognitive knowledge, and motivation to L2 listening comprehension investigating with Persian EFL college students. It was found that the four factors may not present a significant explanation for the difference in L2 listening comprehension solely, but the interaction of these factors could provide a better L2 listening performance.

Zeng's (2007) experimental research on Chinese university-level L2 listeners' listening comprehension performance over a seven-week process-based listening instruction including learner diaries and discussion with listening strategies concluded that the treatment group indicated statistically significant growth in listening performance comparing with the control group.

The effect of a process-based listening instruction with a metacognitive approach on 60 French learners' listening performance and metacognitive awareness was questioned by Vandergrift (2007). He concluded that the treatment group showed a greater level of metacognitive awareness and listening performance significantly.

Mareschal (2007) examined a self-regulatory approach and a process-based listening treatment to adult French listeners with different proficiency levels for nine weeks within a pedagogical cycle, including reporting their responses and discussing them with each other. According to results, both high-skilled and low-skilled listeners increased their metacognitive awareness and strategy use, and their

self-confidence and interests in listening were also developed, but the latter were the ones who most benefited from the approach.

The impact of strategy treatment on the learners' listening performance and self-efficacy in two intervention groups: one high-scaffolded, one low-scaffolded, and one control group was investigated by Graham and Macaro (2008). The students in the treatment groups outperformed the comparison group. It was concluded that the listening instruction developed the proficiency of listening and self-efficacy of the students. It was claimed that it was unlikely to teach individual strategies such as isolating the strategy of inferencing; combining cognitive strategies with metacognitive ones, including monitoring and evaluating and engaging listeners in a specific task with scaffolding, strategy instruction was useful for L2 listeners.

Cross (2009) intended to determine the impacts of explicit strategy treatment on EFL listeners. It was a classroom-based quasi-experimental study with a treatment-control group that included advanced Japanese listeners. The treatment group received the explicit strategy treatment including the presentation, practice, and post-reflection activities using BBC news video texts on listening strategy use including self-management, selective-attention, planning, and self-evaluation together with seven cognitive and three social-affective strategies for 12 hours. The results showed that strategy instruction facilitated learners' listening performance; both groups showed more remarkable development in their listening performance, but no significant differences were indicated in the final performances of both groups.

Cross (2010) studied on increasing advanced-level adult Japanese EFL listeners' metacognitive awareness of listening through a pedagogical cycle emphasizing the sociocultural theory with peer-peer interaction and learner diaries, and through creating metacognitive awareness between the pairs and their discussion on the use of planning, monitoring and evaluating strategies over five lessons. According to the results, the use of dialogue could contribute to raising L2 listeners' metacognitive awareness.

Coskun (2010) studied how metacognitive listening strategy instruction influenced 20 prep-school students' listening performance, with A1-level of English

proficiency. The training group received a five-week metacognitive strategy instruction in which the CALLA method including the five steps of preparation, presentation, practice, evaluation, expansion, which O'Malley and Chamot (1990) proposed, was applied to explicitly teach planning, monitoring, evaluation, and problem identification strategies to use for future listening (Vandergrift, 1997), using their embedded coursebook. The MALQ was used as a tool to discuss using the strategies in each task to enhance their awareness of listening. According to the findings, it was concluded that the treatment group demonstrated a better result in their listening performance; embedding it into the listening course curriculum for training better L2 listeners were needed.

In Vandergrift and Tafaghodtari's (2010) seminal research, the listening comprehension and metacognitive awareness of EFL listeners were investigated through applying metacognitive instruction based on a process-based methodology to teaching L2 listening. The treatment group, including more-successful and less-successful listeners, received metacognitive instruction, with the whole guided practice of planning, monitoring, evaluating, and problem-solving strategies through the guide of the instructor, rather than one-by-one explicit strategy training. The control group with more-successful and less-successful listeners listened to the identical texts during 12 weeks with the same teacher without any attention to the metacognitive processes. The instrument to determine the learners' metacognitive awareness was the MALQ applied three times: before, during, and after the training. The result showed that less-successful listeners who received the instruction demonstrated more significant development in their listening comprehension performance. The treatment group had a higher level of listening comprehension performance on the final test, and more-successful listeners who received the training demonstrated better results in the posttest. However, the difference between more-successful L2 listeners comparing to the groups was slightly different. The greater development of less-successful listeners could be why the process-based listening practices with the instructor's guide and more-successful listeners helped them be aware of the process and therefore show a better listening performance. The guided practice also contributed to the ongoing cycle of metacognitive and cognitive processes to be automatized by the listeners. Besides, during the study, the listeners' metacognitive awareness from both groups increased

through instruction, especially for less-successful L2 listeners; however, the empirical differences were seen only in “Problem Solving and Mental Translation” (p. 483). While the growth in Problem Solving strategy use could be attributed to implicit strategy training over the process, the increase in Mental Translation could be paradoxical because decreasing the use of translation over the process was expected. According to the results of the stimulated-recall protocols, the increase in translation could be attributed to the growth in their lexical knowledge and the ability to identify the vocabularies’ meaning or the increase in the use of Problem Solving to inference more on what they did not understand. All in all, metacognitive training based on a procedural approach to L2 listening enhanced L2 listeners’ metacognitive awareness of listening and listening success, especially the lower-skilled listeners.

Rahimi and Katal (2012a) studied metacognitive awareness of listening among university and high-school EFL learners in Iran. They applied to the MALQ to determine if the level of education made a variance in their metacognitive awareness of listening with the subdimensions of the MALQ. According to the results, both groups had a higher level of metacognitive awareness of listening. However, high-school students showed a greater metacognitive awareness of listening than university-level students, which could be because of their level of motivation, self-efficacy, or listening ability (Vandergrift, 2003, 2005). Also, it showed that Problem Solving was used strategy among all students at most and that guessing the unknown vocabularies by using their prior experience and world knowledge and assessing their inferences, comparing what they know about the topic, and revising them when needed were among those strategies related to Problem Solving. On the other hand, it was seen that the level of Person Knowledge was the lowest one for both groups, but high-school students were higher ranked in Person knowledge and Mental Translation strategy use. Rahimi and Katal (2012b) searched the metacognitive awareness of EFL learners and their readiness to use podcasts in their English learning process. Their awareness was found at a moderate level and that the more metacognitive awareness of listening they had, the more they felt ready for the technology of podcast-use to develop their listening ability. It was also found that Person Knowledge and Problem Solving were linked to podcasts-use in English language teaching and learning.

Katrancı and Yangın (2012) searched the listening performance and attitudes towards L2 listening of 65 fifth-grade learners of Turkish by applying 12-week metacognitive strategy instruction. According to the results, metacognitive instruction positively impacted their attitudes towards listening and success in listening.

In another study, 62 Iranian L2 listeners' listening comprehension achievement was searched by applying the process-based metacognitive training developed by Vandergrift and Tafaghodtari (2010). The treatment group showed a better listening achievement in the last test. Similarly, by applying the same approach, Rahimi and Katal (2013) investigated an experimental study on metacognitive instruction to upper-intermediate EFL listeners, searching their speaking and listening ability. According to the results, the treatment group showed a higher level of listening and speaking proficiency, but no statistically significant changes were found between the groups regarding listening performance.

Bozorgian (2012) explored the effect of strategy-based metacognitive instruction on 28 high-intermediate adult male Iranian EFL listeners' listening comprehension. The listeners took place in four listening courses in which advance organization, self-management, selective and directed attention were taught with the IELTS listening texts. According to the results, less-effective listeners outperformed more-effective listeners in the final IELTS test. It also demonstrated that the instruction promoted the listeners' metacognitive awareness and listening success.

In Goh and Hu's (2013) study, ESL learners' listening performance and metacognitive awareness of listening and their relationship were questioned. They conducted the MALQ and the IELTS listening proficiency test. According to the findings, their listening performance indicated 22% of the difference that was accounted for by their metacognitive awareness of listening. Also, it was found that there was a considerable association between metacognitive strategy use, mainly Directed Attention and Problem Solving, and listening success.

Bozorgian (2014) questioned whether metacognitive instruction influenced high-intermediate EFL listeners' listening achievement after an 8-week metacognitive pedagogical training in listening to 30 male students for eight weeks.

According to the findings obtained from the MALQ and the IELTS test, the listeners' listening performance increased over metacognitive instruction. However, no statistically significant differences were determined regarding their overall metacognitive awareness of listening. However, Problem Solving and Planning and Evaluation strategies caused a statistically significant difference in use, but Person Knowledge, Directed Attention, and Mental Translation strategies did not.

Altuwairesh (2016) examined the metacognitive listening strategy use by 82 female Saudi students in L2 listening. According to the results of the MALQ, the strategies of Problem Solving and Directed Attention were mostly practiced while the strategies of Mental Translation and Person Knowledge were least applied. In terms of Problem Solving, the participants mostly and significantly used the strategies of guessing the unknown words with the known ones, deducing the unknown words by using the main idea of the text, and using prior knowledge or experience to help to understand. In terms of Directed Attention, they mostly used the concentration strategies harder in the moments of trouble in understanding, secondly on trying to keep listening when the concentration is lost, and thirdly on sustaining concentration. In contrast, they did not commonly use the strategy of giving up when they lost track of listening.

In her thesis, Çevikbaş (2016) aimed to investigate the impact of listening activities designed with metacognitive strategies on listening to English preparatory school students in Turkey. The treatment group received metacognitive instruction in their L1 consciously for seven weeks while the control group did not. An achievement test, structured learning diaries, and reflective exercises were applied. According to the results, the training group had a nonsignificant growing listening performance. It was also found that metacognitive instruction increased the students' motivation and positively affected their attitudes toward listening.

Bozorgian and Alamdari (2017) studied the impact of metacognitive instruction with dialogic interaction on advanced Iranian L2 listeners through multimedia tools. The data were gathered via the MALQ and multimedia listening tests from one experimental group after 10-session metacognitive instruction through dialogic intervention and from the other experimental group after 10-session of metacognitive instruction with one-hour twice a week, and from the control group. It was concluded that both groups that received one of both instructions showed a

development in their overall metacognitive awareness of listening with multimedia listening achievement. However, no statistically significant increases were determined in the use of Directed Attention, Person Knowledge, and Mental Translation.

Enrich and Henderson (2019) conducted a Rasch analysis of the MALQ to measure its validation and analyze its psychometric properties. The MALQ was conducted on 299 adult Korean L2 listeners and found that the MALQ with the other four subscales had reliable and useful psychometric properties except for Person Knowledge.

The other research on explicit metacognitive strategy instruction on EFL listeners in Iran was carried out by Maftoon and Alamdari (2020). After a 10-week metacognitive instruction to an experimental group with 30 intermediate EFL listeners, with a control group with 30 intermediate EFL listeners without any guided instruction, the results demonstrated that metacognitive instruction led to remarkable growth in the listening success and overall metacognitive awareness of the EFL listeners. Their metacognitive awareness on the subdimensions of the MALQ except for Problem Solving was significantly affected by metacognitive strategy instruction.

To sum up, it was indicated that metacognitive instruction positively affected listeners' metacognitive awareness of listening and listening achievement; even though some were not statistically significant, most of them supported that metacognitive instruction facilitated L2 listening learning. Through metacognitive instruction, more-successful listeners and especially less-successful listeners showed a more remarkable development in their metacognitive awareness of listening and listening performance. While Problem Solving, in some studies, was the most developed aspect of metacognitive awareness after metacognitive instruction while Person Knowledge was one of the least used strategies of metacognitive awareness in some studies. Also, metacognitive instruction made the listeners have more positive attitudes to L2 listening, have increased their self-efficacy beliefs, and enhanced their motivation. In light of the literature review and related studies, the study was conducted, analyzed, concluded, and discussed in the next chapters.

Chapter 3

Methodology

The methodological framework of the report in this part starts with listing the research design, data collection methods chosen to answer the research questions, and then the study context and the participants who contributed the data to be collected. After that, the data collection procedures and instruments used during this process are listed. Later, the data analysis and the study's reliability and validity are clarified.

Research Design

Before explaining the research design, it is crucial to bring the research questions into mind. Primarily, the study questions a) to what extent metacognitive instruction affects L2 listeners' listening comprehension performance; b) to what extent metacognitive instruction affects L2 listeners' metacognitive awareness of listening.

As regards the objectives of the research, the overall research strategy is based on an experimental research design, a quantitative data collection methodology, which enables to test of any possible cause-and-effect relationships between two or more factors while systematically controlling other possible extraneous factors to minimize their effects on possible research outcomes and keep them constant, therefore strengthening the internal validity of the research (Phakiti, 2014; Johnson & Christensen, 2012). In this design, one group of participants, called an experimental group, receives a particular treatment, while another as the control group does not receive the treatment. The experimental group's progress is measured to check if the treatment factor causes any change or not, compared with the control group (Dörnyei, 2007).

Among experimental research designs, a quasi-experimental research methodology involving a treatment-control group with a pretest-posttest design was adopted for the current paper to search to what extent metacognitive instruction influences listening comprehension skills. Participants were purposefully chosen among first-year ELT students of a state university called purposeful or judgmental sampling among the non-probability sampling forms. The purposeful sampling is applicable in the quantitative and qualitative methods and aims to identify

participants' features according to the interest of the study and choose among them who have those features (Johnson & Christensen, 2012). The rationale for choosing this sampling is that it composes a research group suitable for the study's aims and tries to minimize any possible extraneous factors (Marshall, 1996). The reason the first-year students were preferred for this study was that they were supposed to be at the same level, B2 level, of English proficiency to start the ELT department. After the purposeful sampling, the study participants were determined on a voluntary basis among the first-year ELT students at a Turkish state university, called the convenience sampling among the non-probability sampling forms (Dörnyei, 2007). Details on the participants are presented in the stage of Setting and Participants.

In order to find out to what extent metacognitive instruction influenced L2 listeners' listening comprehension performance, a quantitative research methodology was adopted. Quantitative research consists of data collection providing numerical data that is later statistically analyzed; it defines the statistical relationship between variables. Therefore, it provides objective and reliable results that could be generalized to other contexts (Dörnyei, 2007). As mentioned above, the quasi-experimental research design with an experimental-control group and a pretest-posttest design was adopted for this study: the IELTS listening test as a pretest and a posttest was applied to collect the numerical data on L2 listeners' listening comprehension performance and analyze it statistically later.

In order to determine to what extent metacognitive instruction impacted L2 listeners' metacognitive awareness, a mixed-methods study design was adopted. It has been considered as the third research methodology (Johnson & Onwuegbuzie, 2004) in addition to traditional research methods that were quantitative and qualitative methods about which a paradigm war in the 70s-80s was broken out between researchers who were either on the side of quantitative or qualitative research method and who later started to accept and integrate both methods (Dörnyei, 2007). As a result of this, a mixed-methods research design appeared.

Before presenting the mixed-methods research design, it is necessary to clarify the qualitative research methodology. Dörnyei (2007) explains that qualitative research, unlike quantitative research, includes non-numerical data; during the process, the data is emerged out, and there could be some constructed hypotheses about the theoretical framework of the research as well as no preconditioned

hypotheses to be tested; it occurs in a natural setting and is interested in subjective experiences, ideas, and emotions of participants; the sample size is generally small, and the researcher interprets the data. Nevertheless, the small sample size might cause not to generalize findings to broader contexts because of its “idiosyncratic” nature; the subjective interpretation of the data may affect the validity and the reliability of the study as well (p. 41). Regarding the advantages and disadvantages of both methods, implementing a mixed methods study becomes more vital for the current research.

As for the mixed methods design for the present study, the reasons why mixed-methods research was preferred could be identified with four concepts by Greene, Caracelli, and Graham (1989): triangulation, which aims to corroborate and validate the findings obtained from different research methods; complementarity, which aims clarify and elaborate the findings from the one with from the other methodology; initiation, which aims to discover contradictory or different aspects using the results of one of the methods; development, which aims to benefit from the results of the one to give information to another method; and expansion, which aims to increase the scope of a study through distinct methods and components. The mixed-methods design was adopted to clarify, compare, and validate the findings of both methodologies with each other for the current research.

Considering the mixed methods design types, Creswell and Clark (2011) note that four principles for deciding an appropriate mixed-methods design are crucial: “the level of interaction” and “the relative priority” between the study’s quantitative and qualitative parts, “the timing” and “the procedures for mixing” of both parts (p. 64). The embedded design, in which the researcher mixes the procedures of both methodologies, and embeds both within a traditional method, was found suitable for the study. Among its types, the embedded experimental study, in which the qualitative data is gathered and analyzed “before, during, and/or after” the process of the treatment in an experimental study, was considered as convenient for the current study and the reasons of mixing methods (Creswell & Clark, 2011, p. 92).

When analyzed in terms of the four principles of choosing the mixed methods design, in the present embedded mixed-methods experimental study, there is a direct interaction between the study’s quantitative and qualitative parts, which both are combined in different ways at different times before the last interpretation. There

exists a quantitative priority in which the qualitative data is used in a supportive role; a multiple timing phase including sequential and/or concurrent collection and analysis of secondary data before/during/ after the treatment process; a procedure for mixing occurs at the level of design when the qualitative data is embedded within a broader experimental design (Creswell & Clark, 2011).

Regarding all information on the mixed methods design mentioned above, for the second main question, during the 8-week experimental process of metacognitive instruction, the Metacognitive Awareness Listening Questionnaire (MALQ) to identify the metacognitive awareness of listening of the participants was applied before, during, and after the instruction. That is the quantitative part of the study. During that process, to broaden the understanding of the quantitative data and validate its results with a secondary research method, discover contradictory or different aspects using the findings from it, a qualitative method, stimulated recall protocols, one of the introspection methods, were administered after conducting the midpoint MALQ and the last MALQ questionnaire for the third sub-question of the second main question. In that method, participants try to recall and express what happens in their minds after completing a task to reach unseen mental processes during the task (Dörnyei, 2007). In that method, some kinds of stimulus are used as a reminder to achieve successful recall, for example; showing a video from the interview, listening to an audio text from the interview, or giving the participants their written materials to see the changes they made (Gass & Mackey, 2000). The MALQs responded by the experimental group in the current research were used as a reminder to see the changes they made on their MALQ responses.

As clarified, while the quantitative data through the IELTS listening test and the MALQ were collected before, during, and after 8-week metacognitive instruction, the qualitative data via stimulated recall protocols on the MALQ were collected in the middle and the endpoint of metacognitive instruction. Then, the data from both were analyzed. In Table 3, the overall research design is summarized.

Table 3

Research Design of the Study

RQ	Research Question	Instrument	Sample	N	Data Analysis	Analysis Method
----	-------------------	------------	--------	---	---------------	-----------------

1 Main RQ	To what extent does metacognitive instruction affect the listening comprehension performance of L2 listeners?	Listening Proficiency Test (IELTS)	First-year ELT students at a state university	37	Quantitative	Wilcoxon signed ranks test Mann-Whitney U test
1.1.	Are there any statistically significant differences in L2 listeners' listening comprehension performance over metacognitive instruction regarding the level of listening proficiency?	Listening Proficiency Test (IELTS)	First-year ELT students at a state university	37	Quantitative	Wilcoxon signed ranks test Mann-Whitney U test
1.2.	Are there any statistically significant differences in the listening comprehension performance of L2 listeners regarding preparatory school background?	Listening Proficiency Test (IELTS)	First-year ELT students at a state university	37	Quantitative	Wilcoxon signed ranks test Mann-Whitney U test
2. Main RQ	To what extent does metacognitive instruction affect the metacognitive awareness of listening of L2 listeners?	MALQ	First-year ELT students at a state university	37	Quantitative	Friedman test Wilcoxon signed ranks test Mann-Whitney U test
2.1.	Are there any statistically significant differences in L2 listeners' metacognitive awareness of listening over metacognitive instruction regarding the level of listening proficiency?	MALQ	First-year ELT students at a state university	37	Quantitative	Friedman test Wilcoxon signed ranks test Mann-Whitney U test
2.2.	Are there any statistically significant differences in L2 listeners' metacognitive awareness of listening over metacognitive instruction regarding	MALQ	First-year ELT students at a state university	37	Quantitative	Friedman test Wilcoxon signed ranks test Mann-Whitney U test

	preparatory school background?					
2.3.	What are the perceptions of L2 listeners on possible changes in metacognitive awareness of listening over metacognitive instruction?	Stimulated Recall Protocols on the MALQ	First-year students at the ELT department of a state university (from the experimental group)	4	Qualitatively	Thematic Content Analysis

Setting and Participants

This research was administered at the Foreign Languages Education Department, Education Faculty of a state university in Turkey. The Foreign Languages Department offers a Bachelor's degree in the field of English Language Teaching (ELT), which started to provide undergraduate education in the 2015-2016 academic year. The program aims to improve students' language proficiency, educate them with current language teaching methods, and become well-educated and reformer English teachers. Before studying at the ELT department, students have to take a compulsory proficiency exam at the beginning of the semester to either place in the preparatory school or directly transfer to the department. If students do not have a B2 level of English proficiency, they have to study at the preparatory school until they achieve that level at the end of the term. If they possess a B2 level of English proficiency, they transfer to the faculty to major in the ELT program.

As mentioned above, the participants were purposefully chosen among the first-year students of ELT students who started to study in the ELT department in the fall term of 2019-2020 academic year, which is called purposeful or judgmental sampling among the non-probability sampling forms. The reason for it was that they were supposed to be at the same level, B2 level of English proficiency, to start to major in the ELT department. However, as mentioned above, first-graders may be the ones who pass the proficiency exam and directly transfer to the faculty and the ones who fail the proficiency exam, study at the preparatory school until they get a B2 level of English proficiency, and after passing the proficiency exam, transfer to study at the faculty next semester. Considering this criterion, the study's sampling group included the first graders having different preparatory school backgrounds.

After the sampling group was determined, the first graders including almost 60 students at the ELT department were informed about the research, its objectives, and process, and invited them to participate the study as a volunteer. In that phase, the convenience sampling based on the first-year ELT students' desire to volunteer was applied. At the beginning of the study, a consent form was asked from the volunteers to show their acceptance of participating in the study as an evident (Appendix A), and their background information was collected from the volunteers about their gender, age, preparatory school background. The information gathered was examined, and it was found out that the total number of volunteers was 48, including 15 male and 33 female students, and their ages were ranged between 19 and 27 at the beginning of the study. Thirty-two students were the ones who had never studied at the prep-school, while 16 students were the ones who had a preparatory school background. Considering all these differences, they were grouped as the participants who studied at the preparatory school and who did not study at the preparatory school before.

In order to minimize any confounding effects that may occur as a result of the discrepancy in preparatory school backgrounds among the participants, the volunteers were allocated into the experimental and control groups equally but randomly regarding preparatory school background; therefore, the homogeneity between the groups aimed to be maximized. Despite this, the homogeneity between the groups could be affected by other factors such as L2 listening background, their prior exposure to English; that is why the study was accepted as a quasi-experimental study, and they were regarded as the study limitation.

During the study, 11 out of 48 participants, including seven female and four male participants, missed at least one of the data collection sessions and were extracted from the study. Considering the gender distribution at the ELT department, where females were much more than males regarding the number, and considering the participants who were removed from the study, the gender distribution became unequal between the groups. However, because gender was not analyzed as an independent variable in the study, this issue has been assumed not to affect the study results, but it has been evaluated as a limitation in any case.

Besides, in order to find out whether metacognitive instruction affects L2 listeners regarding the level of listening proficiency, the participants were also

grouped as less- and more-successful listeners, considering their pretest performance, which aimed to measure their level of listening proficiency. The ones whose score was above the mean were identified as successful listeners, and the ones whose score was below the mean were identified as less-successful listeners in each group.

As a result, the data collected from both groups were analyzed regarding the level of listening proficiency and preparatory school background. Demographic information of the participants who took place in each data collection session and accepted to be evaluated during the data analysis is presented in Tables 4, 5, and 6 below.

Table 4
Gender and Participants

Groups	Gender	N	%
Experimental Group	Female	11	61,1
	Male	7	38,9
	Total	18	100,0
Control Group	Female	15	78,9
	Male	4	21,1
	Total	19	100,0

Table 5
Level of Listening Proficiency and Participants

Groups	Level of Listening Proficiency	N	%
Experimental Group	Less successful listeners	8	44.4
	More successful listeners	10	55.6
	Total	18	100.0
Control Group	Less successful listeners	10	52.6
	More successful listeners	9	47.4
	Total	19	100.0

Table 6
Preparatory School Background and Participants

Groups	Preparatory School Experience	N	%
Experimental Group	Yes	6	33.3
	No	12	66.7
	Total	18	100.0
Control Group	Yes	9	47.4
	No	10	52.6
	Total	19	100.0

Data Collection

Pilot Study. Before the real implementation of the study, to prevent any possible problems that may occur during the study, pilot studies were fulfilled.

The pilot study of the IELTS listening test. Firstly, the first-year ELT students were informed and invited to voluntarily take place in the pilot studies. Nine first-grade students were volunteered for the pilot studies. A consent form was collected from those volunteers who joined only the pilot studies and did not participate in the research's real implementation.

Firstly, the pilot study of the IELTS listening test was conducted on nine volunteer students. For this pilot study, it aimed to check its reliability and validity statistically in the context of the study, although it is evident that the IELTS test is an objective high-stake English proficiency test all around the world and used as a reliable and valid tool in several studies (Bozorgian, 2014; the details on the IELTS test are explained under the title of Data Instruments). On a predetermined day, the researcher conducted the IELTS listening test as a pilot study by following the test's instructions after all audio files and equipment were checked and prepared. No problems were observed before, during, or after applying it. Therefore, this could enhance the reliability of the study.

To assess the pilot test results, the answer key provided on the website of the IELTS for this test was used. Using this objective and clear answer key, the researcher scored the tests and checked the given scores. Thanks to the objective answer key, no faults were observed as expected. The Cronbach's alpha internal consistency reliability coefficients were calculated to determine the IELTS listening test's internal validity for the pilot study and found as .85, which is quite preferable (Pallant, 2011).

The pilot study of the MALQ. The MALQ was conducted on the same nine volunteer students for the pilot research. The MALQ is a reliable and valid questionnaire used in several studies (e.g., Bozorgian, 2014; Goh & Taib, 2006; Cross, 2009). The details on the MALQ are explained under the title of Data Instruments). The MALQ is originally in English and suitable for the intermediate level of English proficiency listeners, which means that the participants of the study are supposed to understand it. However, to prevent any misunderstanding while conducting it, the researcher translated it, then the translated form was checked by

another expert, and the Turkish meaning of each statement was stated under each statement in English. To ensure the understanding of the MALQ and check its reliability and validity statistically in the study context before the real implementation, the pilot study of the MALQ was applied. According to its findings, the value of Cronbach's Alpha for the MALQ itself was .69, which is almost .70 and acceptable (Pallant, 2011). Also, for the subscales of the MALQ, they were calculated for Planning and Evaluation as $\alpha = .68$; for Problem Solving as $\alpha = .69$; which are still acceptable, while for Mental Translation as $\alpha = .74$; for Directed Attention as $\alpha = .65$; for Person Knowledge as $\alpha = .85$, which are acceptable as well (Pallant, 2011).

The pilot study of the stimulated recall protocol. With a volunteer participant from the experimental group, who left the study after this pilot study, the stimulated recall protocol was tested in L1. It intended to develop the instructions or the questions to be followed in the study's real implementation if necessary. The pilot study lasted for 12-13 minutes. During the process, no problems were observed in terms of the clarity of the questions, management of time, understanding, or protocol setting. However, two more questions were added, which were "What did you realize on your listening performance when choosing this option?" and "What could be done to be a successful listener?") (see the chapter of Data Instruments for the questions in English and see Appendix F for both Turkish and English versions). Therefore, this pilot study made the research process clearer and enhanced the study's reliability and validity.

Data collection procedure. After the pilot studies, for the current study, the data collection procedures took place as follows:

The data were collected at the Department of ELT of a Turkish state university during the fall term of the 2019-2020 academic year. At first, the study, its aim, and methods were explained to all the ELT department's first-year students, and they were invited to participate in the study. Then, 48 first-year students were volunteered for the study, and a consent form which let the students know the permission received for implementing the current research at the ELT department from the Ethical Committee of Hacettepe University (for the permission from the Ethical Committee, see Appendix G) was collected from the participants who were eager to take place in the research (for the consent form, see Appendix A).

At first, the participants' background information on gender, age, and preparatory school background was collected through a question-answer session by the researcher immediately after the consent forms were gathered. Later, the participants were allotted to the experimental and control groups equally regarding different preparatory school backgrounds. However, because 11 participants missed at least one of the data collection sessions, 37 out of 48 participants were accepted to be analyzed in the study.

While assigning the participants into the groups, they found out that they were divided into two groups and took place in an investigation on learning to listen in English, but they were not supposed to know which group they were assigned to. A specific date and a specific place for the 8-week process were then determined together with both groups. After all this, the process of data collection was started.

Firstly, the IELTS test was administered as a pretest and a posttest to assess the participants' listening comprehension performance to see any difference in their listening performance over metacognitive instruction.

Besides, the MALQ was applied after the pretest, at the middle of the study after the 4th lesson, and after the posttest at the end of the metacognitive instruction to determine the metacognitive awareness of listening of the participants.

Moreover, four participants from the treatment group were chosen based voluntarily for two stimulated-recall sessions. In the first session in the middle of the study, the researcher shared the participants' first and second MALQ responses and talked about significant contrast points in their responses with them. For the second stimulated-recall session, the researcher shared the second and final MALQ responses and discussed the potential reasons for the available and future contradictions. All sessions were tape-recorded, transcribed word-by-word, put into code, and analyzed via the *Atlas.ti* software program.

As metacognitive instruction procedures for the groups, the following processes were followed:

Instruction procedure for the experimental group. Metacognitive instruction lasted for eight weeks in the 2019-2020 academic year's fall term, and the examiner as an instructor gave metacognitive instruction to the group at a predetermined date once a week. That instruction took one hour once a week. The

L2 listeners listened to a dissimilar authentic listening text per week, and the listening texts lasted 4-6 minutes, which were suitable for the intermediate level of English proficiency. (For the information on weekly listening texts, see Appendix B). The procedure was the same each week and consisted of the steps below:

- Every student was asked to keep a notebook for this study. During each lesson, they wrote the lesson's date and the listening text topic on their notebooks. Four parts, My Predictions, First Listening, and Second Listening and Reflection section, would be on each page (For a sample page, see Appendix C).

- At first, the instructor introduced the listening text's topic and activated the background of the students by providing the title of the listening text. Therefore, the researcher guided students to brainstorm and make predictions about what they could hear and any related words, phrases, or topics and made them write these ideas (in English or Turkish) in the "My Prediction" part. This stage was practiced as a whole class activity for the first four weeks of the study; later, it went on as a pair activity for two weeks and continued being practiced with students on their own until last week.

- After the part of brainstorming, students listened to the text as a first listening. During that time, they were desired to tick beside the guessed ideas if any of them was heard. They are also asked to take notes of any additional information they could hear and understand on the part of First Listening.

- After that, students worked with their peers and compared their guesses, and shared what they comprehended until that moment with their peers. The instructor guided them to discuss contradictory ideas, think of any other possible ideas, and decide the text points that needed to be listened to carefully while the second listening.

- Then, as a second listening, they listened to the same text, trying to understand the sections that needed careful listening and writing down new understood information in "Second Listening." After students completed noting down the information, they would be encouraged to discuss as a whole class to assure their understanding of the text and tell others how they achieved to comprehend it.

- Later, as a third listening, they listened to the same text to confirm their comprehending of what they may have missed before. At that point, a transcription of the text was presented to students to realize the speech form with its written form and its syntactic structures, and the development of word recognition and auditory discrimination skills.

- Finally, students evaluated themselves and wrote down their thoughts, answering the questions of “What I succeed,” “What I had troubled in,” and “What I will do for next/future listening differently” on the part of “Reflection.”

Instruction procedure for the control group. The researcher, as the instructor, managed the lessons at a predetermined date and place. They took one hour once a week for eight weeks. The L2 listeners listened to a dissimilar authentic text each week, but the same texts for both groups. The listening texts took 4-6 minutes, which were suitable for the intermediate level of English proficiency (For the information on weekly listening texts, see Appendix B). The procedure was the same each week; however, it was different from the experimental groups’ because they did not receive the metacognitive instruction during the process.

It consisted of the steps below:

- Every student was asked to keep a notebook. However, they took their notes under different titles from the experimental group, and it included one column for all three listening practices (for a sample page, see Appendix C).

- During each of the three listens, the control group did not practice guessing activities and did not discuss and monitor their understanding of the text with their peers.

- Before the third listening, the instructor handed out the transcription of the text to the class and enabled them to discuss with each other in an attempt to verify their comprehension of the text. There were no places for talking about strategy use and involving in any reflection activity for the listening process.

Instruments

The data were gathered via an IELTS listening proficiency test, the MALQ, and stimulated-recall protocols for the study.

Listening proficiency test. The International English Language Testing System (IELTS) was applied as a pretest and a posttest to assess the L2 listeners' listening performance before and after the study. The IELTS test is one of the standard high-stakes English tests for different purposes such as study, work, and immigration worldwide, and assesses English proficiency in listening, reading, writing, and speaking (What is IELTS?, 2019). It has two formats, General and Academic Training: the IELTS Academic test is for candidates who would like to study in higher education or professional institution in an environment where English is spoken; the General IELTS test is for people who want to go to English speaking countries for work, education, migration (What is IELTS?, 2019). For both, listening and speaking tests are the same, while reading and writing tests are different for both formats.

In the current paper, the IELTS listening test was applied. It assesses participants' listening proficiency and the extent to which they understand the general ideas and specific points, comprehend speakers' thoughts, manners, aims, and follow the idea development. Because the IELTS is one of the reliable and valid, high-stakes English tests worldwide (Test statistics, 2020), it has been used in several studies as a reliable and valid research tool (e.g., Goh & Hu, 2013; Bozorgian, 2012, 2014; Rahimi & Katal, 2012;).

For this research, a sample IELTS listening practice test which is free of charge and open to access for everyone without registration in the website of British Council charity, who is a co-owner of IELTS and to whom the researcher mailed to get permission to use it for the study, was conducted. The participants listened to four recordings: the first one was a conversation between two people taken place in a daily context; the second was a monologue in a daily context; the third was a conversation among a maximum of four people in an educational context; the fourth was a monologue related to an academic topic (Listening practice tests, 2019). They answered 40 questions, and it took almost 40 minutes (for the IELTS listening test, see Appendix D). However, before the IELTS's real application as pretest and posttest, its pilot study was conducted, and its internal validity for the pilot study was found .85, which showed good reliability (Pallant, 2011).

Metacognitive awareness listening questionnaire. The second data collection tool was the Metacognitive Awareness Listening Questionnaire (MALQ)

developed by Vandergrift et al. (2006) in order to measure L2 listeners' perceived metacognitive strategy use and their metacognitive awareness. It includes five subgroups with a complete 21 items regarding metacognitive awareness in L2 listening. It consists of five metacognitive strategies as its subdimensions, including Planning and Evaluation, Problem Solving, Directed Attention, Mental Translation, and Person Knowledge. A definition of these factors and the items matched to these factors in the MALQ are available in Table 7. It is a 6-point Likert scale lined from 1 "strongly disagree" to 6 "strongly agree" (for the MALQ, see Appendix E). 6 items out of the MALQ (3, 4, 8, 11, 16, 18) needed to be reverse-coded; therefore, they were reverse-coded before the data analysis.

Table 7

Definition of the Factors in the MALQ (Vandergrift et al., 2006, p. 450-451)

Factors in the MALQ	Definitions	Items in the MALQ
Planning and Evaluation	Strategies used by listeners to make themselves ready for listening and assess their understanding in their listening performance	1, 10, 14, 20, 21
Problem Solving	Strategies listeners apply to make inferences about the parts they have trouble comprehending during listening and check those inferences	5,7,9,13,17,19
Directed Attention	Strategies listeners apply to sustain their concentration on listening tasks	2, 6, 12, 16
Mental Translation	Use of translation into L1 carefully and ungenerously. Strategies that listeners need to give up using and overcome.	4, 11, 18
Person Knowledge	L2 listeners' beliefs about task difficulty and their perceived capabilities in L2 listening	3, 8, 15

The MALQ is a reliable and valid instrument: it has a strong relationship with L2 listening comprehension success, and almost 13% variance in listening performance could be attributed to metacognitive instruction (Vandergrift et al., 2006). Table 8 demonstrates the relationship between the 5 subdimensions of the MALQ and their Cronbach's alpha values to reveal the MALQ's reliability and validity. Besides, the MALQ has been used in several studies in the field and proved its reliability and validity (e.g., Vandergrift & Tafaghodtari, 2010; Goh & Taib, 2006; Cross, 2009; Graham & Macaro, 2008; Coşkun, 2010; Zeng, 2007, 2012).

Furthermore, the MALQ on listening comprehension skill could be used for different purposes (Vandergrift et al., 2006): it can be a consciousness-raising tool for students; therefore, they can be aware of the process of listening, and consequently, they raise their metacognitive awareness and learn how to listen

better and become self-regulated learners; it can be a diagnostic and teaching tool; hence, teachers determine students' current metacognitive awareness and evaluate their preferences of using metacognitive strategies; and it can be a research tool, researchers could conduct it as a pretest-posttest to determine metacognitive awareness of listening of participants. For this study, it was used as a research tool but served for all purposes during the process.

Furthermore, the MALQ is originally in English and suitable for the intermediate level of English proficiency listeners, which means that the current study participants are supposed to understand it due to their B2 level of English. However, to prevent any misunderstanding while conducting it, the examiner translated it, then the translated form was checked by another expert, and the Turkish meaning of each statement was stated under each statement in English. The pilot study of the version, including English and Turkish statements, was conducted and found reliable (see Data Collection for details).

Table 8

The Reliability Analysis of the MALQ (Vandergrift et al., 2006, p.446)

Factor	Exploratory Factor Analysis					Confirmatory Factor Analysis				
	1	2	3	4	5	1	2	3	4	5
1. Problem Solving	0.74									
2. Planning & Evaluation	0.42	0.75				0.53	1			
3. Mental Translation	0.23	0.16	0.78			0.20	0.12	1		
4. Person Knowledge	0.009	-0.05	0.42	0.74		0.10	0.90	0.46	1	
5. Directed Attention	0.47	0.42	-0.14	-0.30	0.68	0.57	0.49	0.20	0.36	

Stimulated-recall protocols. Among introspective methods, stimulated-recall protocols, known as retrospective interviews, in which participants try to recall and express what happened in their minds during the process after they completed a task in order to reach unseen mental processes was conducted (Dörnyei, 2007). In that method, some kind of stimulus is used as a reminder to achieve successful recall, for example; showing a video from the interview, listening to an audio text from the interview, or giving the participants their written materials to see the changes they made (Gass & Mackey, 2000). For the present study, the responses to the MALQ were used as a reminder to stimulate recall so as to the extent that the participants could verbalize possible reasons for possible changes in their

responses by comparing their first and second MALQ responses during the first protocol and their second and final MALQ responses during the second MALQ.

The purposes of the stimulated-recall protocols were explained to the participants from the experimental group, and five of them were volunteered to participate in these sessions. Therefore, they were chosen on a volunteer basis. However, one of the five students left the study during the process; the real implementation of the stimulated recall protocols took place in the study with four volunteers. Firstly, after applying the second MALQ, the researcher and the participants determined a date to interview, and the protocols were conducted within two days immediately after the second MALQ. In the first protocol, the researcher talked about the participants' first and second MALQ responses and discussed important contradictory points in their responses and possible reasons. For the second stimulated-recall protocol, after applying the final MALQ, the researcher and the participants determined a date to interview, and the protocols were fulfilled within one day immediately after the final MALQ. The researcher talked with the participants on possible reasons for other differences in their final MALQ responses by comparing them to the second MALQ; they discussed what it could or should be done for better listening comprehension performance.

All sessions were carried out in L1 of the participants and translated for the data analysis with the help of an expert. The sessions lasted 10-13 minutes in the first stimulated protocol process and 13-20 minutes in the second stimulated recall protocol process. They were tape-recorded, transcribed word-by-word, put into code, and analyzed via the *Atlas.ti* software program.

The protocols started with an introduction to the interview; the researcher greeted the volunteer and explained the aim of the interview, the duration, privacy, and gave information on the recording. Since there was a statement in the consent form received at the beginning of the process on voluntarily participating in the stimulated recall protocols, there was no need to get it again. Therefore, after information was given, it continued comparing each response to each statement of the first MALQ with the second and the second MALQ with the third, respectively, after giving the participants their MALQs. If there was a change on one of the responses, the researcher asked to the participants the following questions:

- There is a difference in your response to statement X. In the first/second/third MALQ, you chose, for example, “strongly agree”; however, in the first/ second/third MALQ, you changed it with “slightly agree.” There is a change here. What were you thinking here? What were you thinking there? Let’s think over these points.
- What could the reasons for this change be? Can you remember what you were thinking while you were choosing that statement?
- What did you realize about your listening performance when choosing this option? What could they be? Why did you choose it?
- (At the end of the second stimulated-recall protocol after comparing all responses of the second and third MALQ) What can you suggest for a better listening performance? How can you improve your listening comprehension skill? What could be done to be a more successful listener?

Above is the English version of the questions, but their Turkish version was conducted to the participants during the implementation of the study. (See Appendix F). The questions were composed in Turkish, and the researcher translated them into English, then the English version was checked by an expert. The Turkish form was controlled in a pilot study, and the third question and the final question of the fourth question were added to the question list.

The order of data collection was listed below:

1. Pre-test Listening Proficiency Test
2. The First MALQ
3. The 8-week Metacognitive Instruction started
4. The Second MALQ in the middle of the 8-week Metacognitive Instruction
5. The First Stimulated-Recall Protocol after the Second MALQ
6. The 8-week Metacognitive Instruction finished
7. Post-test Listening Proficiency Test
8. The Final MALQ at the end of the 8-week Metacognitive Instruction

9. The Second Stimulated-Recall Protocol after the Final MALQ

Table 9

Data Collection Instruments

Research Questions	Data Collection Instrument
Question 1.1.	A – IELTS Listening Test
Question 1.2.	A – IELTS Listening Test
Question 1.3.	A - IELTS Listening Test
Question 2.1.	B – The MALQ
Question 2.2.	B- The MALQ
Question 2.3.	C – Stimulated Recall Protocols

Data Analysis

Quantitative data analysis. The analysis of the quantitative data collected through the IELTS test and the MALQ was carried out through the software Statistical Package for Social Sciences (SPSS 25.0). Firstly, the pretest-posttest numerical data on listening comprehension performance and the numerical data of the first, second, and final MALQs on metacognitive awareness of listening were transferred to the SPSS and prepared to be analyzed.

The first main research question with its sub-questions was quantitatively analyzed using the data from the IELTS listening test while the second main research question with its first and second sub-questions was quantitatively analyzed by using the data from the MALQ itself and its subscales, which are Planning and Evaluation, Problem Solving, Mental Translation, Directed Attention and Person Knowledge.

Firstly, preliminary descriptive statistics were carried out. No missing values were found. In order to decide what parametric or non-parametric methods of data analysis to be applied, the normality test was conducted to control whether the data were normally distributed. For this purpose, the Kolmogorov-Smirnov and Shapiro-Wilk tests were applied. Lewis-Beck, Bryman, and Futing Liao (2004) assert that if the sample size is under 50, the values of Shapiro-Wilk are examined rather than Kolmogorov-Smirnov. Because the sample size of this study was 37, the values of

the Shapiro-Wilk test was checked. If the p-value, as the significant level of the statistics, is more than .05, it is accepted that the data are distributed normally. Firstly, the normal distribution of the pretest and the posttest was checked in terms of the treatment group, the level of listening proficiency, and preparatory school background. It was found out that the values of Shapiro-Wilk of them were above .05, and therefore the data of the pretest and the posttest were normally distributed between the groups, and also in terms of preparatory school background except for regarding the level of listening proficiency. Nevertheless, when the outliers were analyzed, it was found out there were four outliers out of 18 in the posttest scores of the treatment group as well as six outliers out of 19 in the control group's pretest scores. Considering the outliers' mean scores, which were not very similar to the trimmed mean, the Mann-Whitney U test was applied to see any statistical differences in the L2 listeners' test scores comparing the groups. Besides, the Wilcoxon signed ranks test was applied to determine any statistical differences in L2 listeners' pretest and posttest scores over metacognitive instruction by the groups.

Secondly, for the analysis of the second main research question with its first and second sub-questions, the normal distribution of the MALQ and its subscales for three implementations of each was checked in terms of the treatment group, the level of listening proficiency, and preparatory school background. Due to the non-normal distribution of the data, the Friedman test, the non-parametric version of repeated-measured of ANOVA, was carried on to analyze any statistical differences in L2 listeners' metacognitive awareness of listening over the instruction. After that, to understand the source of any difference, the Wilcoxon signed ranks test was conducted. Also, the Mann-Whitney U tests were conducted to compare the groups. Thus, the quantitative analysis of the research was completed.

Qualitative data analysis. The third sub-research questions of the second main research question were qualitatively analyzed through using the data obtained from the stimulated recall protocols based on possible changes of the participants' responses to the MALQ itself and its subscales, which are Planning and Evaluation, Problem Solving, Mental Translation, Directed Attention and Person Knowledge. For the stimulated recall protocol analysis, the thematic content analysis (Braun & Clarke, 2012) was applied, which is about specifying common patterns out of the

qualitative data and emerging specific themes and subthemes supported by each other. While coding, the researcher used the earlier studies (Vandergrift, 2000, Vandergrift et al., 2006). Due to the fact that in the stimulated recall protocols, the MALQs was used as a reminder to be able to discuss over the given responses by the participants, and that each statement was evaluated one by one with the researcher and the participants during the protocols, determining common codes and putting them into themes and subthemes were done based on the related studies. The researcher composed of themes under the title of each subscale, which were Planning and Evaluation, Problem Solving, Mental translation, Directed Attention, and Person Knowledge.

In addition, since the stimulated recall protocols were carried out in Turkish, the researcher translated the data into English, and an expert controlled the translation. After finishing the coding, the researcher analyzed the same data and coded them again one month later. The researcher then compared the earlier coding with the latter, calculated the intra-rater reliability, and found a good agreement ($K = .60$), which showed good agreement (Peat et al., 2001). The researcher then decided on the latest version of the coding by consulting an expert by showing both coding schemes.

Reliability and Validity of the Study

The current research adopted a mixed-methods study design. This methodology aimed to compare, elaborate the results obtained from the quantitative analysis with the qualitative findings, and validate both findings with each other. Therefore, preferring the mixed-methods study design enhanced the reliability and the validity of the current research.

Besides, before implementing the present study, the pilot studies of each data collection tool were conducted so as to uncover any possible threats to the study. The pilot studies of the reliable and valid IELTS listening test and the reliable and valid MALQ were tested in terms of reliability and validity, and the Cronbach Alpha's coefficients of them are found to be reliable. At the same time, the stimulated recall protocol's questions were presented to an expert with the intent of controlling regarding clarity, serving the aim of the study before implementing the pilot study, which was explained in detail as under the title of Pilot Study.

Also, the Cronbach Alpha's coefficients were estimated to determine the study's internal consistency reliability.

Firstly, to determine any differences in listening comprehension performance of L2 listeners over metacognitive instruction, the IELTS listening test was administered as pretest and posttest. As explained before, the researcher used the objective answer key to score the test in the pilot study and found no fault in the results because it is an objective test with an objective answer key. For this reason, in the real implementation, the researcher scored them at first; after two weeks, the researcher scored the tests again. Later, the researcher compared two score sheets and saw no differences between the grades. Because the researcher used an objective answer key, there were no differences in the scores. Therefore, the reliability of the IELTS tests was ensured with an objective application. After the data analysis, for the internal validity of the present research, the Cronbach's Alpha internal consistency coefficients were calculated to be .78 for the IELTS listening pretest and .79 for the IELTS listening posttest with the help of the SPSS 25, which was considered as good reliability (Pallant, 2011).

Secondly, to determine any changes in the metacognitive awareness of listening of L2 listeners over metacognitive instruction, the MALQ was conducted third times. The internal consistency reliability for the MALQ with its subscales was identified one by one. For the current research, the Cronbach's Alpha values were calculated via the SPSS 25 as follows: the first MALQ itself $\alpha = .67$, the second MALQ itself $\alpha = .76$, the third MALQ itself $\alpha = .81$. It showed that the reliability of three implementations of the overall MALQ was preferable (Pallant, 2011). It can show that the increase in the reliability levels of the MALQ itself is consistent with the process of metacognitive instruction. However, when the subscales of the MALQ were analyzed in terms of the Cronbach's alpha value, it was revealed that respectively the Cronbach Alpha's values of the subscales of Person Knowledge (first $\alpha = .73$, second $\alpha = .80$, third $\alpha = .81$), Mental Translation (respectively first $\alpha = .58$, second $\alpha = .76$, third $\alpha = .78$) and Problem Solving (respectively first $\alpha = .56$, second $\alpha = .78$, third $\alpha = .85$) was acceptable because the value from .5 to .7 could also be accepted as moderate reliability (Aron & Aron, 1999, especially in studies in the field of psychology) while the first, second and third Cronbach Alpha's values of Planning and Evaluation (respectively $\alpha = .28$, $\alpha = .36$, $\alpha = .45$) and Directed Attention

(respectively $\alpha = .20$, $\alpha = .39$, $\alpha = .54$) were unacceptable (Pallant, 2011), but it was seen that it gradually increased within the process of metacognitive instruction though.

Nonetheless, according to earlier studies, it was stated that the numbers of items in a scale could affect and cause the Cronbach Alpha value to be small (Nunnally, 1978; Peterson, 1994, Churchill & Peter, 1994, Dall'oglio et al., 2010); this can explain those small values in Planning and Evaluation and Directed Attention because the item number of those subscales is lower than six. On the other hand, it was recommended to check inter-item correlations for Planning and Evaluation and Directed Attention in those situations. Clark and Watson (1995) emphasized that average inter-item correlations should be between .15 and .50. On the other hand, Briggs and Cheek (1986) stated that the mean inter-item correlations for the items were ranged between .2 and .4. When analyzed the mean inter-item correlations of Planning and Evaluation and Directed Attention, it was found as follows: the first Planning and Evaluation was .09, the second Planning and Evaluation was .10, the third Planning and Evaluation was .15, which is acceptable, while the first Directed Attention was .09, the second Directed Attention was .13, the third Directed Attention was .20, which is acceptable. The final MALQ could be accepted as reliable in terms of its all subscales. Nevertheless, considering all these, Schmitt (1994) claims that in some circumstances, finding out low levels of Cronbach's Alpha value could be useful, and this may show that there is an undesirable interaction in the subscale of Planning and Evaluation and Directed Attention in terms of its first and second implementation for this study. In the findings, those low values could be explained meaningfully due to the variances in those constructs (Peterson, 1994) as a result of the study.

In addition, the stimulated recall protocols with four participants from the experimental group were carried out twice: once after the second MALQ and once after the final MALQ to discuss possible changes in the participants' responses to the MALQs over the process of metacognitive instruction. The researcher transcribed the data on the *Atlas.ti* program and did initial coding. The researcher coded the same data one month later. For the intra-rater reliability of the stimulated recall protocol analysis, a Kappa Measurement of Agreement showed enough harmony with the value as $K=.60$ (Peat et al., 2001).

Given the information on the analysis above, the study's internal validity is moderately satisfied. On the other hand, the study's external validity is quite limited in generalizing the results to other contexts because the number of participants is quite small by the nature of experimental studies. However, these issues are discussed in the limitation part of the study.



Chapter 4

Findings

This section aims to explain the findings obtained from the data analysis. Initially, the first and second main research questions with their first and second sub-questions are analyzed quantitatively. The qualitative findings are then explained by answering the third sub-question of the second research question through thematic content analysis to elaborate on the quantitative findings and validate all findings. The findings are accepted as statistically significant if the p -value is below .05, and the strength of any differences in the variables is interpreted according to the guideline of Cohen (1988) that indicates the effect size estimate $r = 0.1$ to 0.3 = small effect; 0.3 to 0.5 = medium effect; 0.5 and higher = large effect size which is estimated by a converted z -score (Rosenthal, 1991, as cited in Field, 2009, p. 550).

Quantitative findings

R. Q. 1. To what extent does metacognitive instruction affect L2 listeners' listening comprehension performance? The Wilcoxon signed ranks tests were conducted to determine any changes in the IELTS listening pretest and posttest performances for each group. The descriptive statistics on these scores were presented in Table 10.

Table 10

Descriptive Statistics of the Pretest and Posttest Scores by the Groups

Group		N	Mean	Std. Deviation	Min	Max	Median
Experimental	Pretest	18	20.0000	5.30261	12.00	30.00	21.0000
	Posttest	18	22.8333	5.63863	11.00	32.00	22.5000
Control	Pretest	19	19.4211	5.95647	9.00	30.00	20.0000
	Posttest	19	19.4211	6.54315	11.00	32.00	17.5000

Second, the Wilcoxon signed ranks tests were conducted to determine the effect of metacognitive instruction on the IELTS listening pretest and posttest scores of L2 listeners in the experimental and control groups. The findings on it within the groups were presented in Table 11.

Table 11

Wilcoxon Signed Ranks Tests on Pretest and Posttest within the Groups

	N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
--	---	-----------	--------------	---	------------------------

Experimental Group	Posttest-Pretest	Negative Ranks	2 ^a	8.00	16.00	-2.695 ^d	0.007
		Positive Ranks	14 ^b	8.57	120.00		
		Ties	2 ^c				
Control Group	Posttest-Pretest	Negative Ranks	8 ^a	8.25	66.00	-.104 ^d	0.917
		Positive Ranks	8 ^b	8.75	70.00		
		Ties	3 ^c				

a. posttest < pretest
b. posttest > pretest
c. posttest = pretest
d. Based on negative ranks.

According to the Wilcoxon signed ranks tests' results, it was found that there was a statistically significant difference between the IELTS listening pretest and posttest scores by the experimental group, $z = -2.695$, $p < .05$, with a medium effect size ($r = .45$) (Cohen, 1988). The median score on the IELTS listening comprehension performance of the experimental group increased from pre-metacognitive instruction ($Mdn = 21$) to post-metacognitive instruction ($Mdn = 22.5$). However, no significant differences between the IELTS listening pretest ($Mdn = 20$) and posttest scores ($Mdn = 17.5$) by the control group were found, $z = -.104$, $p = .917$, with a small effect size ($r = .02$) (Cohen, 1988).

Then, the Mann-Whitney U tests were conducted to see any changes in the scores of the pretest and posttest of IELTS listening between the experimental and control groups. Table 12 shows the findings of the Mann-Whitney U tests applied to the pretest and posttest scores by separately comparing the experimental group to the control group.

Table 12

Mann-Whitney U Test on Pretest and Posttest between the Groups

	Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Pretest	Experimental	18	19.36	348.50	164.500	-0.199	0.843
	Control	19	18.66	354.50			
	Total	37					
Posttest	Experimental	18	22.28	401.00	112.000	-1.797	0.072
	Control	19	15.89	302.00			
	Total	37					

According to the Mann-Whitney U test analysis, no significant differences were found between the experimental group ($Mdn = 21$, $n = 18$) and the control

group ($Mdn = 20$, $n = 19$) in their listening pretest scores, $U = 164.500$, $z = -0.199$, $p = .843$, $r = .03$, with a small effect size as well as between the experimental group ($Mdn = 22.5$, $n = 18$) and the control group ($Mdn = 17.5$, $n = 19$) in their listening posttest scores, $U = 112.000$, $z = -1.797$, $p = .072$, $r = .30$, with a medium effect size.

R. Q. 1. 2. Are there any statistically significant differences in L2 listeners' listening comprehension performance over metacognitive instruction regarding the level of listening proficiency? To answer the question above, at first, the Wilcoxon signed ranks tests were conducted to determine the effect of metacognitive instruction on the IELTS listening pretest and posttest scores of less-successful and more-successful L2 listeners in the experimental and control groups. Firstly, Table 13 presents the descriptive statistics of the pretest and posttest scores by the groups regarding the level of listening proficiency.

Table 13

Descriptive Statistics of the Pretest and Posttest by the Groups regarding Level of Listening Proficiency

Group	Level of Listening Proficiency		N	Mean	Std. Deviation	Min	Max	Median
Experimental	Less-successful	Pretest	8	15.2500	2.25198	12.00	18.00	15.5000
		Posttest	8	18.6250	4.43807	11.00	23.00	20.5000
	More-successful	Pretest	10	23.8000	3.61478	21.00	30.00	21.5000
		Posttest	10	26.2000	4.04969	21.00	32.00	26.5000
Control	Less-successful	Pretest	10	15.2000	4.15799	9.00	20.00	17.0000
		Posttest	10	15.8000	3.67575	11.00	22.00	15.5000
	More-successful	Pretest	9	24.1111	3.65529	20.00	30.00	22.0000
		Posttest	9	23.4444	6.82113	13.00	32.00	25.0000

Secondly, Table 14 shows the Wilcoxon signed ranks tests' findings of the pretest and posttest scores within the groups regarding the level of listening proficiency.

Table 14

Wilcoxon Signed Ranks Tests on Pretest and Posttest within the Groups regarding Level of Listening Proficiency

Group	Level of Listening Proficiency	N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
-------	--------------------------------	---	-----------	--------------	---	------------------------

Experimental group	Less-successful	Posttest-Pretest	Negative Ranks	2 ^a	2.00	4.00	-1.970 ^d	0.049
			Positive Ranks	6 ^b	5.33	32.00		
			Ties	0 ^c				
			Total	8				
	More-successful	Posttest-Pretest	Negative Ranks	0 ^a	0.00	0.00	-2.546 ^d	0.011
			Positive Ranks	8 ^b	4.50	36.00		
			Ties	2 ^c				
			Total	10				
Control group	Less-successful	Posttest-Pretest	Negative Ranks	4 ^a	4.38	17.50	-.595 ^d	0.552
			Positive Ranks	5 ^b	5.50	27.50		
			Ties	1 ^c				
			Total	10				
	More-successful	Posttest-Pretest	Negative Ranks	4 ^a	4.25	17.00	-.511 ^e	0.610
			Positive Ranks	3 ^b	3.67	11.00		
			Ties	2 ^c				
			Total	9				

a. posttest_listening < pretest_listening
 b. posttest_listening > pretest_listening
 c. posttest_listening = pretest_listening
 d. Based on negative ranks.
 e. Based on positive ranks.

According to the findings, it was found that there was a statistically significant difference between the IELTS listening pretest and posttest scores of less-successful L2 listeners in the experimental group, $z = -1.970$, $p < .05$, with a medium effect size ($r = .49$) (Cohen, 1988). The median score on the IELTS listening comprehension performance of less-successful L2 listeners in the experimental group increased from pre-metacognitive instruction ($Mdn = 15.5$) to post-metacognitive instruction ($Mdn = 20.5$). Similarly, it was concluded that there was a statistically significant difference between the IELTS listening pretest ($Mdn = 21.5$) and posttest ($Mdn = 26.5$) scores of more-successful L2 listeners in the experimental group, $z = -2.546$, $p < .05$, with a large effect size ($r = .57$) (Cohen, 1988). It showed that the median score on the IELTS listening comprehension performance of more-

successful L2 listeners in the experimental group increased from pre-metacognitive instruction ($Mdn = 21.5$) to post-metacognitive instruction ($Mdn = 26.5$).

On the other hand, in the control group, there were no statistically significant differences in the IELTS listening pretest ($Mdn = 17$) and posttest ($Mdn = 15.5$) scores of less-successful L2 listeners, $z = -.595$, $p = .552$, $r = .13$, with a small effect size as well as in the IELTS listening pretest ($Md = 22$) and posttest ($Md = 25$) scores of more-successful L2 listeners, $z = -.511$, $p = .610$, $r = .12$, with a small effect size (Cohen, 1988).

After that, the Mann-Whitney U tests were conducted to determine any differences in the pretest and posttest scores of IELTS listening between less-successful and more-successful L2 listeners within the experimental group and the control group regarding the level of listening proficiency. Table 15 shows the Mann-Whitney U tests' findings applied to the pretest and posttest scores by the experimental group and the control group regarding the level of listening proficiency.

Table 15

Mann-Whitney U Test on Pretest and Posttest within the Groups regarding Level of Listening Proficiency

Group		Level of Listening Proficiency	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Experimental	Pretest	Less-successful	8	4.50	36.00	0.000	-3.603	0.000
		More-successful	10	13.50	135.00			
		Total	18					
	Posttest	Less-successful	8	5.50	44.00	8.000	-2.858	0.004
		More-successful	10	12.70	127.00			
		Total	18					
Control	Pretest	Less-successful	10	5.55	55.50	0.500	-3.643	0.000
		More-successful	9	14.94	134.50			
		Total	19					
	Posttest	Less-successful	10	7.15	71.50	16.500	-2.332	0.020
		More-successful	9	13.17	118.50			
		Total	19					

According to the Mann-Whitney U test analysis, in the experimental group, there was a significant difference between less-successful L2 listeners ($Mdn = 15.5$, $n = 8$) and more-successful L2 listeners ($Mdn = 21.5$, $n = 10$) in their pretest scores, $U = .000$, $z = -3.603$, $p < .05$, $r = .85$, with a large effect size. The IELTS listening pretest scores of more-successful L2 listeners in the experimental group were significantly higher ranked than the less-successful L2 listeners' pretest scores at the beginning of metacognitive instruction. Also, the posttest scores in the experimental group showed a statistically significant difference between less-successful L2 listeners ($Mdn = 20.5$, $n = 8$) and more-successful L2 listeners ($Mdn = 26.5$, $n = 10$), $U = 8.000$, $z = -2.858$, $p < .05$, $r = .67$, with a large effect size (Cohen, 1988). The IELTS listening posttest scores of more-successful L2 listeners in the experimental group were significantly higher ranked than the less-successful L2 listeners' posttest scores at the end of metacognitive instruction.

Similarly, in the control group, there was a statistically significant difference between less-successful L2 listeners ($Mdn = 17$, $n = 10$) and more-successful L2 listeners ($Mdn = 22$, $n = 9$) in their pretest scores, $U = .500$, $z = -3.643$, $p < .05$, $r = .84$ with a large effect size. The IELTS listening pretest scores of more-successful L2 listeners in the control group were significantly higher ranked than the less-successful L2 listeners' pretest scores at the beginning of metacognitive instruction. Likewise, the posttest scores of the control group showed a statistically significant difference between less-successful L2 listeners ($Mdn = 15.5$, $n = 10$) and more-successful L2 listeners ($Mdn = 25$, $n = 9$), $U = 16.500$, $z = -2.332$, $p < .05$, $r = .53$, with a large effect size (Cohen, 1988). The IELTS listening posttest scores of more-successful L2 listeners in the control group were significantly higher ranked than the less-successful L2 listeners' posttest scores at the end of metacognitive instruction.

After that analysis, another Mann-Whitney U test was conducted to explore any differences between the experimental group and the control group in their pretest scores and posttest scores regarding the level of listening proficiency. Table 16 shows the Mann-Whitney U test findings applied to the pretest and posttest scores to compare the experimental group with the control group regarding their listening proficiency level.

Table 16

Mann-Whitney U Test on Pretest and Posttest between the Groups regarding Level of Listening Proficiency

Level of Listening Proficiency		Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Less-successful	Pretest	Experimental	8	8.88	71.00	35.500	-0.451	0.652
		Control	10	10.00	100.00			
		Total	18					
	Posttest	Experimental	8	11.31	90.50	25.500	-1.298	0.194
		Control	10	8.05	80.50			
		Total	18					
More-successful	Pretest	Experimental	10	9.75	97.50	42.500	-0.210	0.834
		Control	9	10.28	92.50			
		Total	19					
	Posttest	Experimental	10	111.00	111.00	34.000	-0.901	0.367
		Control	9	79.00	79.00			
		Total	19					

According to the Mann-Whitney U test analysis, no significant differences were found between the experimental group ($Mdn = 15.5$, $n = 8$) and the control group ($Mdn = 17$, $n = 10$) in the listening pretest scores of less-successful L2 listeners, $U = 35.500$, $z = -.451$, $p = .652$, $r = .11$, with a small effect size as well as between the experimental group ($Mdn = 20.5$, $n = 8$) and the control group ($Mdn = 15.5$, $n = 10$) in their listening posttest scores of less-successful L2 listeners, $U = 25.500$, $z = -1.298$, $p = .194$, $r = .31$ with a medium effect size (Cohen, 1988).

In addition, according the findings of the Mann-Whitney U test, no significant differences were found between the experimental group ($Mdn = 21.5$, $n = 10$) and the control group ($Mdn = 22$, $n = 9$) in the listening pretest scores of more-successful L2 listeners, $U = 42.500$, $z = .210$, $p = .834$, $r = .05$, with a small effect size as well as between the experimental group ($Mdn = 26.5$, $n = 10$) and the control group ($Mdn = 25$, $n = 9$) in the listening posttest scores of more-successful L2 listeners, $U = 34.000$, $z = -.901$, $p = .367$, $r = .21$, with a small effect size (Cohen, 1988).

R. Q. 1. 2. Are there any statistically significant differences in L2 listeners' listening comprehension performance over metacognitive instruction regarding different preparatory school backgrounds?. To answer the question above, at first, the Wilcoxon signed ranks tests were conducted to find out any differences in the scores of the pretest and posttest of IELTS listening by the experimental group and the control group regarding different preparatory school

backgrounds. Firstly, Table 17 shows the descriptive statistics of the pretest and posttest scores by the groups regarding different preparatory school backgrounds.

Table 17

Descriptive Statistics of the Pretest and Posttest by the Groups regarding Preparatory School Backgrounds

Group	Preparatory school background		Mean	N	Std. Deviation	Min	Max	Median
Experimental	Yes	Pretest	23.6667	6	6.08824	14.00	30.00	24.5000
		Posttest	25.5000	6	7.71362	11.00	32.00	28.5000
	No	Pretest	18.1667	12	3.95045	12.00	25.00	17.5000
		Posttest	21.5000	12	4.03395	14.00	30.00	21.5000
Control	Yes	Pretest	20.3333	9	5.31507	11.00	28.00	21.0000
		Posttest	18.3333	9	5.63471	11.00	28.00	17.0000
	No	Pretest	18.6000	10	6.65332	9.00	30.00	18.0000
		Posttest	20.4000	10	7.42668	11.00	32.00	19.0000

Secondly, Table 18 shows the Wilcoxon signed ranks tests' findings applied to the pretest and posttest scores within the groups regarding different preparatory school backgrounds.

Table 18

Wilcoxon Signed Ranks Tests on Pretest and Posttest within the Groups regarding Preparatory School Background

Preparatory school background	Group			N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
Yes	Experimental Group	Posttest-Pretest	Negative Ranks	1 ^a	5.00	5.00	-1.166 ^d	0.244
			Positive Ranks	5 ^b	3.20	16.00		
			Ties	0 ^c				
	Control Group	Posttest-Pretest	Negative Ranks	6 ^a	4.00	200	-1.706 ^e	0.088
			Positive Ranks	1 ^b	4.00	4.00		
			Ties	2 ^c				
No	Experimental Group	Posttest-Pretest	Negative Ranks	1 ^a	3.50	3.50	-2.451 ^d	0.014
			Positive Ranks	9 ^b	5.72	51.50		
			Ties	2 ^c				
	Control Group	Posttest-Pretest	Negative Ranks	2 ^a	5.50	11.00	-1.372 ^e	0.170
			Positive Ranks	7 ^b	4.86	34.00		
			Ties	1 ^c				

a. posttest_listening < pretest_listening

b. posttest_listening > pretest_listening

- c. $\text{posttest_listening} = \text{pretest_listening}$
 - d. Based on negative ranks.
 - e. Based on positive ranks.
-

The Wilcoxon signed ranks tests were conducted to find out the impact of metacognitive instruction on the IELTS listening pretest and posttest scores of L2 listeners by the experimental group and the control group regarding different preparatory school backgrounds. It was found that there was a statistically significant difference between the IELTS listening pretest and posttest scores of L2 listeners in the experimental group who did not study at the preparatory school before, $z = -2.451$, $p < .05$, with a large effect size ($r = .50$) (Cohen, 1988). The median score on the IELTS listening comprehension performance for L2 listeners in the experimental group who did not study at the preparatory school before increased from pre-metacognitive instruction ($Mdn = 17.5$) to post-metacognitive instruction ($Mdn = 21.5$). However, no significant differences between the IELTS listening pretest ($Mdn = 24.5$) and posttest ($Mdn = 28.5$) scores of L2 listeners in the experimental group who studied at the preparatory school before were concluded, $z = -1.166$, $p = .244$, with a medium effect size ($r = .34$) (Cohen, 1988).

On the other hand, for the control group, there were no statistically significant differences in the IELTS listening pretest ($Mdn = 21$) and posttest ($Mdn = 17$) scores for L2 listeners who studied at the preparatory school before, $z = -1.706$, $p = .088$, $r = .40$, with a medium effect size as well as in the IELTS listening pretest ($Md = 18$) and posttest ($Md = 19$) scores for L2 listeners who did not study at the preparatory school before, $z = -1.372$, $p = .17$, $r = .31$, with a medium effect size (Cohen, 1988).

After that, the Mann-Whitney U tests were conducted to reveal any changes in the pretest and posttest scores of IELTS listening within the experimental group and the control group regarding different preparatory school backgrounds. Table 19 shows the Mann-Whitney U tests' findings applied to the pretest and posttest scores compared with each other for the experimental and control groups in terms of their preparatory school backgrounds.

Table 19

Mann-Whitney U Test on Pretest and Posttest within the Groups regarding Preparatory School Background

Group		Preparatory school background	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Experimental	Pretest	Yes	6	12.75	76.50	16.500	-1.851	0.064
		No	12	7.88	94.50			
		Total	18					
	Posttest	Yes	6	12.58	75.50	17.500	-1.742	0.082
		No	12	7.96	95.50			
		Total	18					
Control	Pretest	Yes	9	11.33	102.00	33.000	-0.982	0.326
		No	10	8.80	88.00			
		Total	19					
	Posttest	Yes	9	9.28	83.50	38.500	-0.532	0.595
		No	10	10.65	106.50			
		Total	19					

According to the Mann-Whitney U test analysis, in the experimental group, no significant differences were found between L2 listeners who studied at the preparatory school before ($Mdn = 24.5$, $n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 17.5$, $n = 12$) in their pretest scores, $U = 16.500$, $z = -1.851$, $p = .064$, $r = .44$, with a medium effect size, and also, the posttest scores in the experimental group did not differ between L2 listeners who studied at the preparatory school before ($Mdn = 28.5$, $n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 21.5$, $n = 12$), $U = 17.500$, $z = -1.742$, $p = .082$, $r = .41$, with a medium effect size (Cohen, 1988).

Similarly, in the control group, there were no significant differences between L2 listeners who studied at the preparatory school before ($Mdn = 21$, $n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 18$, $n = 10$) in their pretest scores, $U = 33.000$, $z = -.982$, $p = .326$, $r = .23$, with a small effect size, and also, the posttest scores of the control group did not show any significant differences between L2 listeners who studied at the preparatory school before ($Mdn = 17$, $n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 19.00$, $n = 10$), $U = 38.500$, $z = -.532$, $p = .595$, $r = .12$, with a small effect size (Cohen, 1988).

After that analysis, another Mann-Whitney U test was conducted to explore any differences between the experimental and control groups in their pretest scores and posttest scores regarding different preparatory school backgrounds. Table 20 shows the Mann-Whitney U test findings applied to the pretest and posttest scores

to compare the experimental group to the control group regarding their different preparatory school backgrounds.

Table 20

Mann-Whitney U Test on Pretest and Posttest between the Groups regarding Preparatory School Background

Preparatory School Experience		Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Yes	Pretest	Experimental	6	9.58	57.50	17.500	-1.130	0.259
		Control	9	6.94	62.50			
		Total	15					
	Posttest	Experimental	6	10.67	64.00	11.000	-1.891	0.059
		Control	9	6.22	56.00			
		Total	15					
No	Pretest	Experimental	12	11.38	136.50	58.500	-0.100	0.920
		Control	10	11.65	116.50			
		Total	22					
	Posttest	Experimental	12	12.29	147.50	50.500	-0.630	0.529
		Control	10	10.55	105.50			
		Total	22					

According to the Mann-Whitney U test analysis, no significant differences were found between the experimental group ($Mdn = 24.5$, $n = 6$) and the control group ($Mdn = 21$, $n=9$) in the listening pretest scores by L2 listeners who studied at the preparatory school before, $U = 17.500$, $z = -1.130$, $p = .259$, $r = .30$, with a medium effect size as well as between the experimental group ($Mdn = 28.5$, $n =6$) and the control group ($Mdn = 17$, $n = 9$) in their listening posttest scores of L2 listeners who studied at the preparatory school before, $U = 11.000$, $z = -1,891$, $p = .059$, $r = .49$ with a medium effect size.

In addition, according the findings of the Mann-Whitney U test, no significant differences were found between the experimental group ($Mdn = 17.5$, $n = 12$) and the control group ($Mdn = 18$, $n = 10$) in the listening pretest scores of L2 listeners who did not study at the preparatory school before, $U = 58.500$, $z = .100$, $p = .920$, $r = .02$, with a small effect size as well as between the experimental group ($Mdn = 21.5$, $n = 12$) and the control group ($Mdn = 19$, $n = 10$) in their listening posttest scores, $U = 50.500$, $z = -.630$, $p = .529$, $r =.13$, with a small effect size (Cohen, 1988).

R. Q. 2. To what extent does metacognitive instruction affect L2 listeners' metacognitive awareness of listening?. To answer the research

question above, the Friedman test as the nonparametric version of the mixed between-within subjects analysis of variance (ANOVA) was administered to show any differences in L2 listeners' metacognitive awareness within and between the groups over the process of metacognitive instruction because the data did not meet the assumptions of that type of ANOVA, except for the assumptions of normality and homogeneity of variances.

The independent variables are Group (experimental group and control group), which is the between-subjects factor, and Time including the measures of MALQ at the beginning of (time point 1), at the midpoint of (time point 2), and at the end of (time point 3) the metacognitive instruction, which is the within-subjects factor. To investigate the growth in metacognitive awareness of L2 listeners, which is the dependent variable, the total values of the overall MALQ and the subdimensions of the MALQ including Planning and Evaluation (PE), Problem Solving (PS), Directed Attention (DA), Mental Translation (MT), and Person Knowledge (PK) for each time point were calculated.

Since the Friedman test does not provide a post hoc test, a Wilcoxon signed ranks test with Bonferroni adjustment is conducted to see the source of those differences if there are any differences in the results. After analyzing the Friedman test within the groups, to explore any differences between the groups in the scores of MALQ over metacognitive instruction, the Mann-Whitney U tests were applied.

Analysis of the differences in the MALQ scores within the groups: the Friedman test. Firstly, the descriptive statistics of the MALQ scores for the experimental and control groups are presented in Table 21.

Table 21

Descriptive Statistics of the MALQ for the Groups

Group	MALQ	N	Std.				
			Mean	Deviation	Min.	Max.	Median
Experimental	First MALQ	18	91.3333	10.04109	77.00	111.00	88.5000
	Mid MALQ	18	89.1111	11.27667	72.00	111.00	91.0000
	Final MALQ	18	90.1111	11.63104	68.00	118.00	89.5000
	First PS	18	28.8333	4.01834	1.00	36.00	30.0000
	Mid PS	18	29.0000	4.71543	19.00	36.00	31.0000
	Final PS	18	29.0000	5.00588	18.00	36.00	29.5000
	First PE	18	21.0000	3.81945	11.00	26.00	21.5000
	Mid PE	18	19.8333	3.22217	14.00	27.00	19.5000

Control	Final PE	18	21.1667	3.31219	17.00	30.00	20.0000
	First MT	18	11.1111	3.75561	5.00	17.00	10.5000
	Mid MT	18	10.7222	4.12746	5.00	17.00	10.5000
	Final MT	18	10.3333	4.69042	4.00	18.00	10.0000
	First DA	18	19.1111	2.32351	15.00	23.00	19.0000
	Mid DA	18	18.3333	2.70076	14.00	22.00	19.0000
	Final DA	18	19.3889	2.99291	14.00	24.00	19.0000
	First PK	18	11.2778	3.40943	4.00	16.00	11.5000
	Mid PK	18	11.2222	3.60646	5.00	16.00	11.0000
	Final PK	18	10.2222	4.02281	4.00	16.00	10.5000
	First MALQ	19	86.1579	9.08150	65.00	104.00	87.0000
	Mid MALQ	19	84.7368	10.41282	63.00	101.00	86.0000
	Final MALQ	19	82.0526	12.26320	56.00	105.00	84.0000
	First PS	19	27.0000	3.36650	18.00	32.00	27.0000
	Mid PS	19	27.0526	3.97874	17.00	33.00	27.0000
	Final PS	19	25.4737	4.67105	16.00	35.00	25.0000
	First PE	19	19.8421	2.87254	14.00	24.00	21.0000
	Mid PE	19	20.1579	3.18439	15.00	26.00	21.0000
	Final PE	19	19.5789	3.80520	13.00	26.00	20.0000
	First MT	19	10.8421	2.36322	7.00	15.00	11.0000
	Mid MT	19	10.0000	2.66667	5.00	15.00	9.0000
	Final MT	19	10.6842	2.47325	6.00	15.00	11.0000
	First DA	19	18.6842	2.92599	11.00	22.00	20.0000
	Mid DA	19	17.6316	3.18347	10.00	22.00	18.0000
	Final DA	19	17.0000	3.23179	9.00	21.00	18.0000
	First PK	19	9.7895	3.64507	4.00	16.00	10.0000
	Mid PK	19	9.8947	3.75492	3.00	15.00	10.0000
	Final PK	19	9.3158	3.59092	3.00	15.00	9.0000

Secondly, Table 23 shows the findings of the Friedman test conducted to find out any differences in L2 listeners' MALQ scores across the three-time points (at the beginning, midpoint, and end of instruction) over the process of metacognitive instruction by the experimental group and the control group.

Table 22

Friedman Test on the MALQ within the Groups

Group	MALQ	N	Mean Rank	Friedman χ^2	df	Asymp. Sig.
Experimental	First MALQ	18	2.17	1.043	2	0.593
	Mid MALQ	18	1.83			
	Final MALQ	18	2.00			
	First PS	18	1.92	0.206	2	0.902
	Mid PS	18	2.06			
	Final PS	18	2.03			
	First PE	18	2.00	1.910	2	0.385
	Mid PE	18	1.78			
	Final PE	18	2.22			

Control	First MT	18	2.11	1.298	2	0.523
	Mid MT	18	2.08			
	Final MT	18	1.81			
	First DA	18	2.11	3.733	2	0.155
	Mid DA	18	1.67			
	Final DA	18	2.22			
	First PK	18	2.39	6.421	2	0.040
	Mid PK	18	1.97			
	Final PK	18	1.64			
	First MALQ	19	2.47	10.088	2	0.006
	Mid MALQ	19	2.03			
	Final MALQ	19	1.50			
	First PS	19	2.34	8.355	2	0.015
	Mid PS	19	2.13			
	Final PS	19	1.53			
	First PE	19	2.05	1.130	2	0.568
	Mid PE	19	2.13			
	Final PE	19	1.82			
	First MT	19	2.05	2.577	2	0.276
	Mid MT	19	1.76			
	Final MT	19	2.18			
	First DA	19	2.37	5.681	2	0.058
	Mid DA	19	2.00			
	Final DA	19	1.63			
	First PK	19	2.11	0.393	2	0.821
	Mid PK	19	1.95			
	Final PK	19	1.95			

According to the results of the Friedman test, no significant differences were found by the experimental group in the scores of *overall MALQ* scores across the three time points [beginning (*Mdn* = 88.5), mid (*Mdn* = 91), and end (*Mdn* = 89.5)], [χ^2 (2, *n* = 18) = 1.043, *p* = .593]; *Problem Solving* across the three time points [beginning (*Mdn* = 30), mid (*Mdn* = 31), and end (*Mdn* = 29.5)], [χ^2 (2, *n* = 18) = 0.206, *p* = .902]; *Planning and Evaluation* across the three time points [beginning (*Mdn* = 21.5), mid (*Mdn* = 19.5), and end (*Mdn* = 20)], [χ^2 (2, *n* = 18) = 1.910, *p* = .385]; *Mental Translation* across the three time points [beginning (*Mdn* = 10.5), mid (*Mdn* = 10.5), and end (*Mdn* = 10)], [χ^2 (2, *n* = 18) = 1.298, *p* = .523]; *Directed Attention* across the three time points [beginning (*Mdn* = 19), mid (*Mdn* = 19), and end (*Mdn* = 19)], [χ^2 (2, *n* = 18) = 3.733, *p* = .155].

On the other hand, there was a statistically significant difference in the scores of *Person Knowledge* across the three-time points (beginning, mid, and endpoint of instruction), [χ^2 (2, *n* = 18) = 6.421, *p* < .05]. The median values showed a decrease in the experimental group's metacognitive awareness in terms of Person Knowledge

from the beginning ($Mdn = 11.5$) through the midpoint ($Mdn = 11$) to the end ($Mdn = 10.5$) of metacognitive instruction.

In order to explore at which time points the significant difference in the scores of Person Knowledge for the experimental group occurred, the Wilcoxon signed ranks tests with Bonferroni adjustment ($p = .05/3 = .0167$) were conducted and presented in Table 23.

Table 23

Wilcoxon Signed Ranks Tests on Person Knowledge by the Experimental Group

			N	Mean Rank	Sum of Ranks	Z	Adjusted p (.0167)
Experimental Group	Mid PK - First PK	Negative Ranks	10 ^a	6.05	60.50	-.507	.612
		Positive Ranks	4 ^b	11.13	44.50		
		Ties	4 ^c				
	Final PK - Mid PK	Negative Ranks	8 ^d	5.75	46.00	-	.244
		Positive Ranks	3 ^e	6.67	20.00		
		Ties	7 ^f				
	Final PK - First PK	Negative Ranks	12 ^g	8.21	98.50	-	.113
		Positive Ranks	4 ^h	9.38	37.50		
		Ties	2 ⁱ				

a. Mid PK < First PK d. Final PK < Mid PK g. Final PK < First PK

b. Mid PK > First PK e. Final PK > Mid PK h. Final PK > First PK

c. Mid PK = First PK f. Final PK = Mid PK i. Final PK = First PK

It was revealed that no significant differences were found, according to the adjusted p -value, between the first PK and the mid PK, $z = -.507$, $p = .612$, with a small effect size ($r = .08$); between the final PS and the mid PS, $z = -1.164$, $p = .244$, with a small effect size ($r = .19$); between the final PS and the first PS, $z = -1.584$, $p = .11$, with a small effect size ($r = .26$) (Cohen, 1988). Although it was not significant, it was clear that there was a decrease in the metacognitive awareness of the experimental group in terms of Person Knowledge from the beginning ($Mdn = 11.5$) to the midpoint ($Mdn = 11$) and end ($Mdn = 10.5$) of metacognitive instruction.

Besides, the results of the Friedman test showed that no significant differences were found by the control group in the scores of *Planning and Evaluation* across the three time points [beginning ($Mdn = 21$), mid ($Mdn = 21$), and end ($Mdn = 20$)], [$\chi^2 (2, n = 19) = 1.130$, $p = .568$]; *Mental Translation* across the three time

points [beginning ($Mdn = 11$), mid ($Mdn = 9$), and end ($Mdn = 11$)], [$\chi^2 (2, n = 19) = 2.577, p = .276$]; *Directed Attention* across the three time points [beginning ($Mdn = 20$), mid ($Mdn = 18$), and end ($Mdn = 18$)], [$\chi^2 (2, n = 19) = 5.681, p = .058$]; Person Knowledge across the three time points [beginning ($Mdn = 10$), mid ($Mdn = 10$), and end ($Mdn = 9$)], [$\chi^2 (2, n = 19) = .393, p = .821$].

On the other hand, there was a statistically significant difference in the overall metacognitive awareness of the control group across the three-time points (beginning, mid, and endpoint of instruction), [$\chi^2 (2, n = 19) = 10.088, p < .05$]. The median values showed a decrease in overall metacognitive awareness from the beginning ($Mdn = 87$) through the midpoint ($Mdn = 86$) and the end ($Mdn = 84$) of metacognitive instruction.

In order to explore at which time points the significant difference in the overall MALQ scores for the control group was found, the Wilcoxon signed ranks tests with Bonferroni adjustment ($p = .05/3 = .0167$) were conducted and presented in Table 24.

Table 24

Wilcoxon Signed Ranks Tests on the Overall MALQ for the Control Group

			N	Mean Rank	Sum of Ranks	Z	Adjusted p (0.0167)
Control Group	Mid MALQ-First MALQ	Negative Ranks	11 ^a	8.73	96.00	-.925 ^b	0.355
		Positive Ranks	6 ^b	9.50	57.00		
		Ties	2 ^c				
	Final MALQ-Mid MALQ	Negative Ranks	11 ^d	9.95	109.50	-53 ^b	0.031
		Positive Ranks	5 ^e	5.30	26.50		
		Ties	3 ^f				
	Final MALQ-First MALQ	Negative Ranks	15 ^g	8.37	125.50	-22 ^b	0,020
		Positive Ranks	2 ^h	13.75	27.50		
		Ties	2 ⁱ				

a. Mid MALQ < First MALQ d. Final MALQ < Mid MALQ g. Final MALQ < First MALQ
b. Mid MALQ > First MALQ e. Final MALQ > Mid MALQ h. Final MALQ > First MALQ
c. Mid MALQ = First MALQ f. Final MALQ = Mid MALQ i. Final MALQ = First MALQ

It was found out that no significant differences were found, according to the adjusted p -value, between the first MALQ and the mid MALQ, $z = -.925$, $p = .355$, with a small effect size ($r = .15$); between the final MALQ and the mid MALQ, $z = -2.153$, $p = .031$, with a medium effect size ($r = .35$); between the final MALQ and the first MALQ, $z = -2.322$, $p = .020$, with a medium effect size ($r = .38$) (Cohen, 1988). However, as mentioned above, there was a decrease in overall metacognitive awareness of listening of the control group from the beginning ($Mdn = 87$), through the midpoint ($Mdn = 86$), and the end ($Mdn = 84$) of metacognitive instruction.

Besides, a statistically significant difference in the scores of the Problem Solving by the control group across the three-time points (beginning, mid, and endpoint of instruction) [$\chi^2 (2, n = 19) = 8.355$, $p < .05$] was concluded. The median values showed a decrease in the control group's metacognitive awareness in terms of Problem Solving from the beginning ($Mdn = 27$) and midpoint ($Mdn = 27$) to the end ($Mdn = 25$) of metacognitive instruction. In order to explore at which time points the difference in the scores of Problem Solving for the control occurred, the Wilcoxon signed ranks tests with Bonferroni adjustment ($p = .05/3 = .0167$) were conducted and presented in Table 25.

Table 25

Wilcoxon Signed Ranks Tests on Problem Solving for the Control Group

			N	Mean Rank	Sum of Ranks	Z	Adjusted p (0.0167)
Control Group	Mid PS-First PS	Negative Ranks	10 ^a	7.50	75.00	-.072	0.943
		Positive Ranks	7 ^b	11.14	78.00		
		Ties	2 ^c				
	Final PS - Mid PS	Negative Ranks	10 ^d	6.90	69.00	-	0.017
		Positive Ranks	2 ^e	4.50	9.00		
		Ties	7 ^f				
	Final PS -First PS	Negative Ranks	13 ^g	8.42	106.50	-	0.031
		Positive Ranks	3 ^h	8.83	26.50		
		Ties	3 ⁱ				

a. Mid PS < First PS d. Final PS < Mid PS g. Final PS < First PS

b. Mid PS > First PS e. Final PS > Mid PS h. Final PS > First PS

c. Mid PS = First PS f. Final PS = Mid PS i. Final PS = First PS

According to the findings, no significant differences were found, regarding the adjusted p -value, between the first PS and the mid PS, $z = -.072$, $p = .943$, with a small effect size ($r = .01$); between the final PS and the first PS, $z = -2.158$, $p = .031$,

with a medium effect size ($r = .35$) (Cohen, 1988). However, to the extent, it could be deduced that there was a slightly significant difference between the final PS and the mid PS, $z = -2.376$, $p = .017$, with a medium effect size ($r = .39$) because the adjusted p -value (.0167) is very close to .0170. Therefore, it can be concluded that there was a decrease in the metacognitive awareness of listening of the control group considering Problem Solving from the beginning ($Mdn = 27$) to the end ($Mdn = 25$) metacognitive instruction.

After analyzing the MALQ across the three-time points of the metacognitive instruction, to explore if there were any differences between the groups in the scores of the MALQ and its subdimensions over metacognitive instruction, the Mann-Whitney U tests were applied.

Analysis of the differences in the MALQ scores between the groups: the Mann-Whitney U Test. Firstly, the descriptive statistics of the MALQ scores between the experimental and the control groups were given in Table 26.

Table 26

Descriptive Statistics of the MALQ between the Groups

MALQ	Group	N	Std.				
			Mean	Deviation	Min.	Max.	Median
First MALQ	Experimental	18	91.3333	10.04109	77.00	111.00	88.5000
	Control	19	86.1579	9.08150	65.00	104.00	87.0000
Mid MALQ	Experimental	18	89.1111	11.27667	72.00	111.00	91.0000
	Control	19	84.7368	10.41282	63.00	101.00	86.0000
Final MALQ	Experimental	18	90.1111	11.63104	68.00	118.00	89.5000
	Control	19	82.0526	12.26320	56.00	105.00	84.0000
First PS	Experimental	18	28.8333	4.01834	21.00	36.00	30.0000
	Control	19	27.0000	3.36650	18.00	32.00	27.0000
Mid PS	Experimental	18	29.0000	4.71543	19.00	36.00	31.0000
	Control	19	27.0526	3.97874	17.00	33.00	27.0000
Final PS	Experimental	18	29.0000	5.00588	18.00	36.00	39.5000
	Control	19	25.4737	4.67105	16.00	35.00	25.0000
First PE	Experimental	18	21.0000	3.81945	11.00	26.00	21.5000
	Control	19	19.8421	2.87254	14.00	24.00	21.0000
Mid PE	Experimental	18	19.8333	3.22217	14.00	27.00	19.5000
	Control	19	20.1579	3.18439	15.00	26.00	21.0000
Final PE	Experimental	18	21.1667	3.31219	17.00	30.00	20.0000
	Control	19	19.5789	3.80520	13.00	26.00	21.0000
First MT	Experimental	18	11.1111	3.75561	5.00	17.00	10.5000
	Control	19	10.8421	2.36322	7.00	15.00	11.0000
Mid MT	Experimental	18	10.7222	4.12746	5.00	17.00	10.5000
	Control	19	10.0000	2.66667	5.00	15.00	9.0000
Final MT	Experimental	18	10.3333	4.69042	4.00	18.00	10.0000

First DA	Control	19	10.6842	2.47325	6.00	15.00	11.0000
	Experimental	18	19.1111	2.32351	15.00	23.00	19.0000
Mid DA	Control	19	18.6842	2.92599	11.00	22.00	20.0000
	Experimental	18	18.3333	2.70076	14.00	22.00	19.0000
Final DA	Control	19	17.6316	3.18347	10.00	22.00	18.0000
	Experimental	18	19.3889	2.99291	14.00	24.00	19.0000
First PK	Control	19	17.0000	3.23179	9.00	21.00	18.0000
	Experimental	18	11.2778	3.40943	4.00	16.00	11.5000
Mid PK	Control	19	9.7895	3.64507	4.00	16.00	10.0000
	Experimental	18	11.2222	3.60646	5.00	16.00	11.0000
Final PK	Control	19	9.8947	3.75492	3.00	15.00	10.0000
	Experimental	18	10.2222	4.02281	4.00	16.00	10.5000
	Control	19	9.3158	3.59092	3.00	15.00	9.0000

Secondly, Table 27 below shows the findings of the Mann-Whitney U tests applied to compare the experimental group to the control group in order to find out if there were any differences in the MALQ scores across the three-time points (at the beginning, midpoint, and end of instruction) over the process of metacognitive instruction.

Table 27

Mann-Whitney U Test on the MALQ between the Groups

	Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
First MALQ	Experimental	18	21.42	385.50	127.500	-1.325	0.185
	Control	19	16.71	317.50			
	Total	37					
Mid MALQ	Experimental	18	21.17	381.00	122.000	-1.187	0.235
	Control	19	16.95	322.00			
	Total	37					
Final MALQ	Experimental	18	22.64	407.50	105.500	-1.992	0.046
	Control	19	15.55	295.50			
	Total	37					
First PS	Experimental	18	21.28	383.00	130.000	-1.253	0.210
	Control	19	16.84	320.00			
	Total	37					
Mid PS	Experimental	18	21.53	387.50	125.500	-1.390	0.165
	Control	19	16.61	315.50			
	Total	37					
Final PS	Experimental	18	22.89	412.00	101.000	-2.134	0.033
	Control	19	15.32	291.00			
	Total	37					
First PE	Experimental	18	21.14	380.50	132.500	-1.178	0.239
	Control	19	16.97	322.50			
	Total	37					
Mid PE	Experimental	18	18.28	329.00	158.000	-.397	0.691
	Control	19	19.68	374.00			
	Total	37					
Final PE	Experimental	18	21.03	378.50	134.500	-1.115	0.265
	Control	19	17.08	324.50			
	Total	37					
First MT	Experimental	18	19.14	344.50	168.000	-0.076	0.939

	Control	19	18.87	358.50			
	Total	37					
Mid MT	Experimental	18	19.86	357.50	155.500	-0.473	0.636
	Control	19	18.18	345.50			
	Total	37					
Final MT	Experimental	18	18.08	325.50	154.500	-.504	0.614
	Control	19	19.87	377.50			
	Total	37					
First DA	Experimental	18	19.25	346.50	166.500	-0.138	0.890
	Control	19	18.76	356.50			
	Total	37					
Mid DA	Experimental	18	20.22	364.00	149.000	-.674	0.500
	Control	19	17.84	339.00			
	Total	37					
Final DA	Experimental	18	22.64	407.50	105.500	-2.004	0.045
	Control	19	15.55	295.50			
	Total	37					
First PK	Experimental	18	21.25	382.50	130.500	-1.236	0.217
	Control	19	16.87	320.50			
	Total	37					
Mid PK	Experimental	18	20.89	376.00	137.000	-1.040	0.298
	Control	19	17.21	327.00			
	Total	37					
Final PK	Experimental	18	20.25	364.50	148.500	-.686	0.493
	Control	19	17.82	338.50			
	Total	37					

According to the Mann-Whitney U test analysis, no significant differences were found in the scores of *the first overall MALQ* between the experimental group ($Mdn = 88.5, n = 18$) and the control group ($Mdn = 87, n = 19$), $U = 127.500, z = -1.325, p = .185, r = .21$, with a small effect size; *the second overall MALQ* between the experimental group ($Mdn = 91, n = 18$) and the control group ($Mdn = 86, n = 19$), $U = 122.000, z = -1.187, p = .235, r = .20$, with a small effect size; *the first PS* between the experimental group ($Mdn = 30, n = 18$) and the control group ($Mdn = 27, n = 19$), $U = 130.000, z = -1.253, p = .210, r = .20$, with a small effect size; *the mid PS* between the experimental group ($Mdn = 31, n = 18$) and the control group ($Mdn = 27, n = 19$), $U = 125.500, z = -1.390, p = .165, r = .23$, with a small effect size; *the first PE* between the experimental group ($Mdn = 21.5, n = 18$) and the control group ($Mdn = 21, n = 19$), $U = 132.500, z = -.178, p = .239, r = .03$, with a small effect size; *the mid PE* between the experimental group ($Mdn = 19.5, n = 18$) and the control group ($Mdn = 21, n = 19$), $U = 158.000, z = -.397, p = .691, r = .07$, with a small effect size; *the final PE* between the experimental group ($Mdn = 20, n = 18$) and the control group ($Mdn = 20, n = 19$), $U = 134.500, z = -1.115, p = .265, r = .18$, with a small effect size; *the first MT* between the experimental group ($Mdn = 10.5, n = 18$) and the control group ($Mdn = 11, n = 19$), $U = 168.000, z = -.076, p = .939, r = .01$, with a small effect size; *the mid MT* between the experimental group

($Mdn = 10.5$, $n = 18$) and the control group ($Mdn = 9$, $n = 19$), $U = 155.500$, $z = -.473$, $p = .636$, $r = .07$, with a small effect size; *the final MT* between the experimental group ($Mdn = 10$, $n = 18$) and the control group ($Mdn = 11$, $n = 19$), $U = 154.500$, $z = -.504$, $p = .614$, $r = .08$, with a small effect size; *the first DA* between the experimental group ($Mdn = 19$, $n = 18$) and the control group ($Mdn = 20$, $n = 19$), $U = 166.500$, $z = -.138$, $p = .890$, $r = .02$, with a small effect size; *the mid DA* between the experimental group ($Mdn = 19$, $n = 18$) and the control group ($Mdn = 18$, $n = 19$), $U = 149.000$, $z = -.674$, $p = .500$, $r = .11$, with a small effect size; *the first PK* between the experimental group ($Mdn = 11.5$, $n = 18$) and the control group ($Mdn = 10$, $n = 19$), $U = 130.500$, $z = -1.236$, $p = .217$, $r = .20$, with a small effect size; *the first PK* between the experimental group ($Mdn = 11$, $n = 18$) and the control group ($Mdn = 10$, $n = 19$), $U = 137.000$, $z = -1.040$, $p = .298$, $r = .17$, with a small effect size; *the first PK* between the experimental group ($Mdn = 10.5$, $n = 18$) and the control group ($Mdn = 9$, $n = 19$), $U = 148.500$, $z = -.686$, $p = .493$, $r = .11$, with a small effect size (Cohen, 1988).

On the other hand, it was found that there was a statistically difference; in the scores of *the final overall MALQ* between the experimental group ($Mdn = 89.5$, $n = 18$) and the control group ($Mdn = 84$, $n = 19$), $U = 105.500$, $z = -1.992$, $p < .05$, $r = .33$, with a medium effect size; in the scores of *the final PS* between the experimental group ($Mdn = 39.5$, $n = 18$) and the control group ($Mdn = 25$, $n = 19$), $U = 101.000$, $z = -2.134$, $p < .05$, $r = .35$, with a medium effect size; in the scores of *the final DA* between the experimental group ($Mdn = 19$, $n = 18$) and the control group ($Mdn = 18$, $n = 19$), $U = 105.500$, $z = -2.004$, $p < .05$, $r = .33$, with a medium effect size (Cohen, 1988). The median score on the metacognitive awareness of the experimental group in terms of the final overall metacognitive awareness, final Problem Solving and final Directed Attention was significantly ranked higher than the control group's median score.

R. Q. 2.1. Are there any statistically significant differences in L2 listeners' metacognitive awareness of listening over metacognitive instruction regarding the level of listening proficiency? To answer the research question above, at first, the Friedman test as the nonparametric version of mixed between-within subjects analysis of variance (ANOVA), because the data did not meet the assumptions of mixed between-within subjects ANOVA, was applied to

show any statistically significant differences in L2 listeners' metacognitive awareness within and between the experimental group and the control group regarding the level of listening proficiency over the process of metacognitive instruction.

The independent variables are Group (experimental group and control group), which is the between-subjects factor, and Time including the measures of MALQ at the beginning (time point 1), at the midpoint (time point 2), and at the end (time point 3) of the metacognitive instruction and Level of Listening Proficiency (less-successful L2 listeners and more-successful L2 listeners), which are the within-subjects factors while the metacognitive awareness of L2 listeners over the process of metacognitive instruction is the dependent factor. In order to measure the metacognitive awareness of L2 listeners, the total values of the overall MALQ and the subdimensions of the MALQ including Planning & Evaluation, Problem Solving, Monitoring, Directed Attention, Mental Translation, and Person Knowledge for each time point were calculated.

Since the Friedman test does not provide a post hoc test, a Wilcoxon signed ranks test with Bonferroni adjustment is conducted to see the source of those differences if there are statistically significant differences in the results.

After the analysis of the Friedman test within the groups, to explore if there were statistically significant differences between the experimental group and the control group regarding the level of listening proficiency in the scores of the MALQ over metacognitive instruction, the Mann-Whitney U tests were applied.

Analysis of the differences in the MALQ scores within the groups regarding the level of listening proficiency: the Friedman test. The findings of the Friedman test applied to find out if there were statistically significant differences in the MALQ scores of less-successful L2 listeners and more-successful L2 listeners across the three-time points (at the beginning, the midpoint, and the end of instruction) over the process of metacognitive instruction were presented by the groups in different tables.

Firstly, the descriptive statistics of the MALQ scores are given in Table 28 for the experimental group.

Table 28

Descriptive Statistics of the MALQ for the Experimental Group regarding Level of Listening Proficiency

Group	Level of Listening Proficiency	MALQ	N	Mean	Std. Deviation	Min.	Max.	Median
Experimental	Less-successful	First MALQ	8	88.5000	11.66190	77.00	111.00	84.0000
		Mid MALQ	8	85.6250	12.50071	72.00	107.00	86.5000
		Final MALQ	8	86.8750	16.21672	68.00	118.00	85.5000
		First PS	8	27.7500	4.83292	21.00	36.00	26.0000
		Mid PS	8	27.1250	5.74301	19.00	35.00	27.5000
		Final PS	8	27.0000	6.36957	18.00	36.00	25.5000
		First PE	8	21.0000	4.86973	11.00	26.00	22.0000
		Mid PE	8	20.7500	3.91882	14.00	27.00	20.0000
		Final PE	8	21.0000	4.27618	17.00	30.00	19.5000
		First MT	8	10.6250	3.06769	5.00	16.00	10.0000
		Mid MT	8	10.6250	3.66206	5.00	17.00	10.5000
		Final MT	8	11.1250	4.48609	5.00	18.00	11.5000
		First DA	8	19.5000	2.61861	16.00	23.00	19.0000
		Mid DA	8	18.2500	2.71241	14.00	22.00	19.0000
		Final DA	8	18.7500	3.95511	14.00	24.00	18.0000
		First PK	8	9.6250	3.96187	4.00	15.00	8.5000
		Mid PK	8	8.8750	2.64237	5.00	13.00	9.0000
		Final PK	8	9.0000	4.14039	4.00	16.00	9.0000
	More-successful	First MALQ	10	93.6000	8.47480	80.00	105.00	95.5000
		Mid MALQ	10	91.9000	9.96048	80.00	111.00	91.0000
		Final MALQ	10	92.7000	5.85093	83.00	102.00	92.0000
		First PS	10	29.7000	3.23351	25.00	36.00	30.0000
		Mid PS	10	30.5000	3.27448	25.00	36.00	31.0000
		Final PS	10	30.6000	3.06232	24.00	35.00	30.5000
		First PE	10	21.0000	3.01846	16.00	26.00	20.5000
		Mid PE	10	19.1000	2.51440	15.00	23.00	19.0000
		Final PE	10	21.3000	2.54078	18.00	27.00	20.5000
		First MT	10	11.5000	4.35252	5.00	17.00	11.5000
		Mid MT	10	10.8000	4.66190	5.00	17.00	11.0000
		Final MT	10	9.7000	4.98999	4.00	17.00	8.5000
		First DA	10	18.8000	2.14994	15.00	22.00	19.0000
		Mid DA	10	18.4000	2.83627	14.00	22.00	19.0000
		Final DA	10	19.9000	2.02485	18.00	24.00	19.0000
		First PK	10	12.6000	2.31900	9.00	16.00	13.0000
		Mid PK	10	13.1000	3.21282	7.00	16.00	14.5000
		Final PK	10	11.2000	3.85285	6.00	16.00	11.5000

Table 29 shows the findings of the Friedman test applied to explore whether there were any changes in the MALQ scores of less-successful and more-successful L2 listeners across the three-time points (at the beginning, midpoint, and end of instruction) over the process of metacognitive instruction by comparing the scores of three-time points by less-successful and more-successful L2 listeners in the experimental group.

Table 29

Friedman Test on the MALQ within the Experimental Group regarding Level of Listening Proficiency

Group	Level of Listening Proficiency	MALQ	N	Mean Rank	Friedman χ^2	df	Asymp. Sig.
Experimental	Less-successful	First MALQ	8	2.25	1.750	2	0.417
		Mid MALQ	8	1.63			
		Final MALQ	8	2.13			
		First PS	8	2.00	0.467	2	0.792
		Mid PS	8	2.33			
		Final PS	8	1.67			
		First PE	8	1.94	0.194	2	0.908
		Mid PE	8	2.13			
		Final PE	8	1.94			
		First MT	8	2.00	1.231	2	0.540
		Mid MT	8	1.75			
		Final MT	8	2.25			
		First DA	8	2.31	2.000	2	0.368
		Mid DA	8	1.69			
		Final DA	8	2.00			
	More-successful	First PK	8	2.56	4.067	2	0.131
		Mid PK	8	1.75			
		Final PK	8	1.69			
		First MALQ	10	2.10	0.216	2	0.898
		Mid MALQ	10	2.00			
		Final MALQ	10	1.90			
		First PS	10	1.80	1.105	2	0.575
		Mid PS	10	2.25			
		Final PS	10	1.95			
		First PE	10	2.05	5.056	2	0.080
		Mid PE	10	1.50			
		Final PE	10	2.45			
		First MT	10	2.20	6.000	2	0.050
		Mid MT	10	2.35			
		Final MT	10	1.45			
		First DA	10	1.95	3.257	2	0.196
		Mid DA	10	1.65			
		Final DA	10	2.40			
		First PK	10	2.25	3.630	2	0.163
		Mid PK	10	2.15			
		Final PK	10	1.60			

According to the results of the Friedman test, no significant differences were found for less-successful L2 listeners in the experimental group in the scores of *overall MALQ* across the three time points [beginning (Mdn = 84), mid (Mdn = 86.5), and end (Mdn = 85.5)], [χ^2 (2, n = 8) = 1.750, p = .417]; *Problem Solving* across the three time points [beginning (Mdn = 26), mid (Mdn = 27.5), and end (Mdn = 25.5)], [χ^2 (2, n = 8) = .467, p = .792]; *Planning and Evaluation* across the three time points [beginning (Mdn = 22), mid (Mdn = 20), and end (Mdn = 19.5)], [χ^2 (2, n = 8) = .194, p = .908]; *Mental Translation* across the three time points [beginning (Mdn = 10), mid (Mdn = 10.5), and end (Mdn = 11.5)], [χ^2 (2, n = 8) = 1.231, p = .540]; *Directed Attention* across the three time points [beginning (Mdn = 19), mid (Mdn = 19), and end (Mdn = 18)], [χ^2 (2, n = 8) = 2.000, p = .368]; *Person Knowledge* across the three time points [beginning (Mdn = 8.5), mid (Mdn = 9), and end (Mdn = 9)], [χ^2 (2, n = 8) = 4.067, p = .131].

Similarly, no significant differences were found for more-successful L2 listeners in the experimental group in the scores of *overall MALQ* across the three time points [beginning (Mdn = 95.5), mid (Mdn = 91), and end (Mdn = 92)], [χ^2 (2, n = 10) = .216, p = .898]; *Problem Solving* across the three time points [beginning (Mdn = 30), mid (Mdn = 31), and end (Mdn = 30.5)], [χ^2 (2, n = 10) = 1.105, p = .575]; *Planning and Evaluation* across the three time points [beginning (Mdn = 20.5), mid (Mdn = 19), and end (Mdn = 20.5)], [χ^2 (2, n = 10) = 5.056, p = .080]; *Mental Translation* across the three time points [beginning (Mdn = 11.5), mid (Mdn = 11), and end (Mdn = 8.5)], [χ^2 (2, n = 10) = 6.000, p = .050]; *Directed Attention* across the three time points [beginning (Mdn = 19), mid (Mdn = 19), and end (Mdn = 19)], [χ^2 (2, n = 10) = 3.257, p = .196]; *Person Knowledge* across the three time points [beginning (Mdn = 13), mid (Mdn = 14.5), and end (Mdn = 11.5)], [χ^2 (2, n = 10) = 3.630, p = .163].

After analyzing the MALQ scores of the experimental group regarding the level of listening proficiency, the descriptive statistics of the MALQ scores are presented in Table 30 for the control group regarding the level of listening proficiency.

Table 30

Descriptive Statistics of the MALQ for the Control Group regarding Level of Listening Proficiency

Group	Level of Listening Proficiency	MALQ	N	Mean	Std. Deviation	Min.	Max.	Median
Control	Less- successful	First MALQ	10	86.9000	7.37036	74.00	97.00	87.0000
		Mid MALQ	10	84.8000	6.23253	72.00	97.00	85.5000
		Final MALQ	10	82.4000	6.50128	68.00	91.00	84.0000
		First PS	10	26.7000	3.62246	18.00	31.00	27.0000
		Mid PS	10	26.2000	4.28952	17.00	32.00	26.5000
		Final PS	10	24.9000	3.92853	16.00	29.00	25.5000
		First PE	10	20.0000	2.44949	15.00	23.00	21.0000
		Mid PE	10	20.0000	2.98142	15.00	23.00	21.0000
		Final PE	10	19.4000	3.43835	13.00	23.00	21.0000
		First MT	10	11.0000	2.53859	7.00	15.00	11.0000
		Mid MT	10	10.5000	2.41523	8.00	15.00	9.5000
		Final MT	10	10.7000	1.76698	8.00	13.00	10.5000
		First DA	10	19.4000	2.06559	15.00	22.00	20.0000
		Mid DA	10	18.4000	2.59058	14.00	22.00	18.5000
		Final DA	10	18.2000	1.54919	16.00	20.00	18.0000
	More- successful	First PK	10	9.8000	3.04777	5.00	15.00	10.0000
		Mid PK	10	9.7000	2.75076	5.00	15.00	9.0000
		Final PK	10	9.2000	2.20101	6.00	13.00	9.0000
		First MALQ	9	85.3333	11.09054	65.00	104.00	87.0000
		Mid MALQ	9	84.6667	14.15097	63.00	101.00	87.0000
		Final MALQ	9	81.6667	17.04406	56.00	105.00	81.0000
		First PS	9	27.3333	3.24037	23.00	32.00	27.0000
		Mid PS	9	28.0000	3.60555	22.00	33.00	28.0000
		Final PS	9	26.1111	5.55528	19.00	35.00	25.0000
		First PE	9	19.6667	3.42783	14.00	24.00	20.0000
		Mid PE	9	20.3333	3.57071	16.00	26.00	20.0000
		Final PE	9	19.7778	4.38115	16.00	26.00	18.0000
		First MT	9	10.6667	2.29129	7.00	14.00	11.0000
		Mid MT	9	9.4444	2.96273	5.00	15.00	9.0000
		Final MT	9	10.6667	3.20156	6.00	15.00	12.0000
		First DA	9	17.8889	3.62093	11.00	22.00	19.0000
		Mid DA	9	16.7778	3.70060	10.00	22.00	18.0000
		Final DA	9	15.6667	4.12311	9.00	21.00	17.0000
		First PK	9	9.7778	4.40959	4.00	16.00	11.0000
		Mid PK	9	10.1111	4.80740	3.00	15.00	12.0000
		Final PK	9	9.4444	4.85054	3.00	15.00	10.0000

Table 31 shows the findings of the Friedman test applied to find out if there were any differences in the three MALQ scores of less-successful and more-successful L2 listeners in the control group across the three-time points (at the beginning, midpoint, and end of instruction) over the process of metacognitive instruction by comparing the MALQ scores of three-time points.

Table 31

Friedman Test on the MALQ within the Control Group regarding Level of Listening Proficiency

Group	Level of Listening Proficiency	MALQ	N	Mean Rank	Friedman χ^2	df	Asymp. Sig.
Control	Less-successful	First MALQ	10	2.50	7.515	2	0.023
		Mid MALQ	10	2.10			
		Final MALQ	10	1.40			
		First PS	10	2.45	6.242	2	0.044
		Mid PS	10	2.10			
		Final PS	10	1.45			
		First PE	10	2.15	1.588	2	0.452
		Mid PE	10	2.15			
		Final PE	10	1.70			
		First MT	10	2.05	0.636	2	0.727
		Mid MT	10	1.85			
		Final MT	10	2.10			
		First DA	10	2.40	3.167	2	0.205
		Mid DA	10	1.95			
		Final DA	10	1.65			
		First PK	10	2.05	0.065	2	0.968
		Mid PK	10	1.95			
		Final PK	10	2.00			
	More-successful	First MALQ	9	2.44	3.257	2	0.196
		Mid MALQ	9	1.94			
		Final MALQ	9	1.61			
		First PS	9	2.22	2.552	2	0.279
		Mid PS	9	2.17			
		Final PS	9	1.61			
		First PE	9	1.94	0.171	2	0.918
		Mid PE	9	2.11			
		Final PE	9	1.94			
		First MT	9	2.06	2.067	2	0.356
		Mid MT	9	1.67			
		Final MT	9	2.28			
		First DA	9	2.33	2.606	2	0.272
		Mid DA	9	2.06			
		Final DA	9	1.61			
		First PK	9	2.17	0.467	2	0.792
		Mid PK	9	1.94			
		Final PK	9	1.89			

According to the Friedman test, no significant differences were found between the scores of the three MALQs of less-successful L2 listeners in the control group in the scores of *Planning and Evaluation* across the three time points [beginning (Mdn = 21), mid (Mdn = 21), and end (Mdn = 21)], [χ^2 (2, n = 10) = 1.588,

$p = .452$]; *Mental Translation* across the three time points [beginning (Mdn = 11), mid (Mdn = 9.5), and end (Mdn = 10.5)], [$\chi^2 (2, n = 10) = .636, p = .727$]; *Directed Attention* across the three time points [beginning (Mdn = 20), mid (Mdn = 18.5), and end (Mdn = 18)], [$\chi^2 (2, n = 10) = 3.167, p = .205$]; *Person Knowledge* across the three time points [beginning (Mdn = 10), mid (Mdn = 9), and end (Mdn = 9)], [$\chi^2 (2, n = 10) = .065, p = .968$].

However, it was found that there was a significant difference in the scores of the overall MALQ scores across the three time points [beginning (Mdn = 87), mid (Mdn = 85.5), and end (Mdn = 84)], [$\chi^2 (2, n = 10) = 7.515, p < .05$] of less-successful L2 listeners in the control group. The median values showed a decrease in the control group's overall metacognitive awareness from the beginning through the midpoint and the end of the process.

In order to explore in which time points the significant difference in the scores of the overall MALQ for less-successful L2 listeners in the control group occurred, the Wilcoxon signed ranks test with Bonferroni adjustment ($p = .05/3 = .0167$) was conducted and presented in Table 32.

Table 32

Wilcoxon Signed Ranks Tests on Overall MALQ for Less-Successful Listeners in the Control Group

	Level of Listening Proficiency	MALQ	N	Mean Rank	Sum of Ranks	Z	Adjusted p (0.0167)
Control Group	Less-successful listeners	Mid MALQ- First MALQ	6 ^b	4.42	26.50	-1.192 ^a	0.233
		Negative Ranks					
		Positive Ranks	2 ^c	4.75	9.50		
		Ties	2 ^d				
		Final MALQ- Mid MALQ	7 ^e	4.86	34.00	-2.254 ^a	0.024
		Negative Ranks					
		Positive Ranks	1 ^f	2.00	2.00		
		Ties	2 ^g				
		Final MALQ - First MALQ	7 ^h	4.29	30.00	-1.682 ^a	0.092
		Negative Ranks					
		Positive Ranks	1 ⁱ	6.00	6.00		
		Ties	2 ^j				

a. Based on positive ranks
 b. Mid MALQ < First MALQ e. Final MALQ < Mid MALQ h. Final MALQ < First MALQ
 c. Mid MALQ > First MALQ f. Final MALQ > Mid MALQ i. Final MALQ > First MALQ

According to the Wilcoxon signed ranks with Bonferroni adjustment, no significant differences were found, according to the adjusted p-value (.0167), between the first MALQ and the mid MALQ, $z = -1.192$, $p = .233$, with a small effect size ($r = .27$); between the mid MALQ and the final MALQ, $z = -2.254$, $p = .024$, with a medium effect size ($r = .50$); between the first MALQ and the final MALQ, $z = -1.682$, $p = .092$, with a medium effect size ($r = .38$) (Cohen, 1988). Although the median values on overall metacognitive awareness for less-successful L2 listeners in the control group showed a decreased from beginning through the midpoint and to the end of the study process [beginning (Mdn = 87), mid (Mdn = 85.5), and end (Mdn = 84)].

Besides, it was found that there was a significant difference in the scores of Problem Solving across the three-time points [beginning (Mdn = 27), mid (Mdn = 26.5), and end (Mdn = 25.5)], [$\chi^2 (2, n = 10) = 6.242$, $p < .05$] of less-successful L2 listeners in the control group. The median values on Problem Solving for less-successful L2 listeners in the control group decreased from the beginning through the midpoint and the end of the study process.

In order to explore in which time points the significant differences in the scores of Problem Solving for less-successful L2 listeners in the control group, the Wilcoxon signed ranks tests with Bonferroni adjustment ($p = .05/3 = .0167$) were conducted and presented in Table 33.

Table 33

Wilcoxon Signed Ranks Tests on Problem Solving for Less-Successful Listeners in the Control Group

	Level of Listening Proficiency	of MALQ		N	Mean Rank	Sum of Ranks	Z	Adjusted p (0.0167)
Control Group	Less-successful	Mid PS- First PS	Negative Ranks	6 ^a	4,50	27,00	-.535 ^a	0.593
			Positive Ranks	3 ^b	6,00	18,00		
			Ties	1 ^c				
		Final PS- Mid PS	Negative Ranks	6 ^d	4,33	26,00	-2.047 ^a	0.041
			Positive Ranks	1 ^e	2,00	2,00		
			Ties	3 ^f				

Final PS - First PS	Negative Ranks	7 ^g	4,36	30,50	-1.774 ^a	0.076
	Positive Ranks	1 ^h	5,50	5,50		
	Ties	2 ⁱ				

a. Based on positive ranks

b. Mid PS < First PS e. Final PS < Mid PS h. Final PS < First PS

c. Mid PS > First PS f. Final PS > Mid PS i. Final PS > First PS

d. Mid PS = First PS g. Final PS = Mid PS j. Final PS = First PS

According to the Wilcoxon signed ranks with Bonferroni adjustment, it was concluded that no significant differences were found, according to the adjusted p-value (.0167), between the first PS and the mid PS, $z = -.535$, $p = .593$, with a small effect size ($r = .12$); between the mid PS and the final PS, $z = -2.047$, $p = .041$, with a medium effect size ($r = .46$); between the first PS and the final PS, $z = -1.774$, $p = .076$, with a medium effect size ($r = .40$) (Cohen, 1988) even though the median values on Problem Solving for less-successful L2 listeners in the control group showed a decrease from the beginning through the midpoint and to the end of the study process [beginning (Mdn = 27), mid (Mdn = 26.5), and end (Mdn = 25.5)].

On the other hand, according to the Friedman test, no significant differences were found for more-successful L2 listeners in the control group in the scores of *the overall MALQ* across the three time points [beginning (Mdn = 87), mid (Mdn = 87), and end (Mdn = 81)], [$\chi^2(2, n = 9) = 3.257$, $p = .196$]; *Problem Solving* across the three time points [beginning (Mdn = 27), mid (Mdn = 28), and end (Mdn = 25)], [$\chi^2(2, n = 9) = 2.552$, $p = .279$]; *Planning and Evaluation* across the three time points [beginning (Mdn = 20), mid (Mdn = 20), and end (Mdn = 18)], [$\chi^2(2, n = 9) = .171$, $p = .918$]; *Mental Translation* across the three time points [beginning (Mdn = 11), mid (Mdn = 9), and end (Mdn = 12)], [$\chi^2(2, n = 9) = 2.067$, $p = .356$]; *Directed Attention* across the three time points [beginning (Mdn = 19), mid (Mdn = 18), and end (Mdn = 17)], [$\chi^2(2, n = 9) = 2.606$, $p = .272$]; *Person Knowledge* across the three time points [beginning (Mdn = 11), mid (Mdn = 12), and end (Mdn = 10)], [$\chi^2(2, n = 9) = .467$, $p = .792$].

After comparing the three MALQ scores for less-successful and more-successful L2 listeners separately within the groups, the three MALQ scores were also analyzed to find out the differences in the MALQ scores by comparing less-successful and more-successful L2 listeners with each other within the groups.

Analysis of the differences in the MALQ scores within the groups regarding the level of listening proficiency: the Mann-Whitney U test. The Mann-Whitney U tests were used to explore any changes in the three MALQ scores measured at the beginning, at the midpoint, and at the end of the instruction within the experimental and control groups' less-successful L2 listeners with more-successful listeners.

Table 34 shows the Mann-Whitney U tests' findings applied to the MALQ scores to compare less-successful and most-successful L2 listeners in the experimental group.

Table 34

Mann-Whitney U Test on the MALQ for the Experimental Group regarding Level of Listening Proficiency

Group		Level of Listening Proficiency	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Experimental	First MALQ	Less-successful	8	7.81	62.50	26.500	-1.200	0.230
		More-successful	10	10.85	108.50			
		Total	18					
	Mid MALQ	Less-successful	8	8.25	66.00	30.000	-0.892	0.373
		More-successful	10	10.50	105.00			
		Total	18					
	Final MALQ	Less-successful	8	7.56	60.50	24.500	-1.381	0.167
		More-successful	10	11.05	110.50			
		Total	18					
	First PS	Less-successful	8	8.19	65.50	29.500	-0.947	0.344
		More-successful	10	10.55	105.50			
		Total	18					
	Mid PS	Less-successful	8	7.50	60.00	24.000	-1.439	0.150
		More-successful	10	11.10	111.00			
		Total	18					
	Final PS	Less-successful	8	7.50	60.00	24.000	-1.428	0.153
		More-successful	10	11.10	111.00			
		Total	18					
	First PE	Less-successful	8	9.94	79.50	36.500	-0.313	0.754
		More-successful	10	9.15	91.50			
		Total	18					
	Mid PE	Less-successful	8	11.06	88.50	27.500	-1.120	0.263
		More-successful	10	8.25	82.50			
		Total	18					
	Final PE	Less-successful	8	8.38	67.00	31.000	-0.811	0.418
		More-successful	10	10.40	104.00			
		Total	18					
	First MT	Less-successful	8	8.63	69.00	33.000	-0.631	0.528
		More-successful	10	10.20	102.00			
		Total	18					
	Mid MT	Less-successful	8	9.50	76.00	40.000	0.000	1.000
		More-successful	10	9.50	95.00			
		Total	18					
		Less-successful	8	10.56	84.50	31.500	-0.758	0.448

Final MT	More-successful	10	8.65	86.50			
	Total	18					
First DA	Less-successful	8	10.19	81.50	34.500	-0.494	0.621
	More-successful	10	8.95	89.50			
	Total	18					
Mid DA	Less-successful	8	9.44	75.50	39.500	-0.045	0.964
	More-successful	10	9.55	95.50			
	Total	18					
Final DA	Less-successful	8	8.63	69.00	33.000	-0.626	0.531
	More-successful	10	10.20	102.00			
	Total	18					
First PK	Less-successful	8	7.06	56.50	20.500	-1.745	0.081
	More-successful	10	11.45	114.50			
	Total	18					
Mid PK	Less-successful	8	5.94	47.50	11.500	-2.559	0.011
	More-successful	10	12.35	123.50			
	Total	18					
Final PK	Less-successful	8	7.94	63.50	27.500	-1.115	0.265
	More-successful	10	10.75	107.50			

According to the Mann-Whitney U test analysis, in the experimental group, no significant differences were found in the scores of *the first overall MALQ* between less-successful L2 listeners ($Mdn = 84, n = 8$) and more-successful L2 listeners ($Mdn = 95.5, n = 10$), $U = 26.500, z = -1.200, p = .230, r = .28$, with a small effect size; *the second overall MALQ* between less-successful L2 listeners ($Mdn = 86.5, n = 8$) and more-successful L2 listeners ($Mdn = 91, n = 10$), $U = 30.000, z = -.892, p = .373, r = .21$, with a small effect size; *the final overall MALQ* between less-successful L2 listeners ($Mdn = 85.5, n = 8$) and more-successful L2 listeners ($Mdn = 92, n = 10$), $U = 24.500, z = -1.381, p = .167, r = .33$ with a medium effect size; *the first PS* between less-successful L2 listeners ($Mdn = 26, n = 8$) and more-successful L2 listeners ($Mdn = 30, n = 10$), $U = 29.500, z = -.947, p = .344, r = .22$, with a small effect size; *the mid PS* between less-successful L2 listeners ($Mdn = 27.5, n = 8$) and more-successful L2 listeners ($Mdn = 31, n = 10$), $U = 24.000, z = -1.439, p = .150, r = .34$, with a medium effect size; *the final PS* between less-successful L2 listeners ($Mdn = 25.5, n = 8$) and more-successful L2 listeners ($Mdn = 30.5, n = 10$), $U = 24.000, z = -1.428, p = .153, r = .34$ with a medium effect size; *the first PE* between less-successful L2 listeners ($Mdn = 22, n = 8$) and more-successful L2 listeners ($Mdn = 20.5, n = 10$), $U = 36.500, z = -.313, p = .754, r = .07$, with a small effect size; *the mid PE* between less-successful L2 listeners ($Mdn = 20, n = 8$) and more-successful L2 listeners ($Mdn = 19, n = 10$), $U = 27.500, z = -1.120, p = .263, r = .26$, with a small effect size; *the final PE* between less-successful L2

listeners ($Mdn = 19.5, n = 8$) and more-successful L2 listeners ($Mdn = 20.5, n = 10$), $U = 31.000, z = -.811, p = .418, r = .19$, with a small effect size; *the first MT* between less-successful L2 listeners ($Mdn = 10, n = 8$) and more-successful L2 listeners ($Mdn = 11.5, n = 10$), $U = 33.000, z = -.631, p = .528, r = .15$, with a small effect size; *the mid MT* between less-successful L2 listeners ($Mdn = 10.5, n = 8$) and more-successful L2 listeners ($Mdn = 11, n = 10$), $U = 40.000, z = .000, p = .1000, r = .00$, with a small effect size; *the final MT* between less-successful L2 listeners ($Mdn = 11.5, n = 8$) and more-successful L2 listeners ($Mdn = 8.5, n = 10$), $U = 31.500, z = -.758, p = .448, r = .18$, with a small effect size; *the first DA* between less-successful L2 listeners ($Mdn = 19, n = 8$) and more-successful L2 listeners ($Mdn = 19, n = 10$), $U = 34.500, z = -.494, p = .621, r = .12$, with a small effect size; *the mid DA* between less-successful L2 listeners ($Mdn = 19, n = 8$) and more-successful L2 listeners ($Mdn = 19, n = 10$), $U = 39.500, z = -.045, p = .964, r = .01$, with a small effect size; *the final DA* between less-successful L2 listeners ($Mdn = 18, n = 8$) and more-successful L2 listeners ($Mdn = 19, n = 10$), $U = 33.000, z = -.626, p = .531, r = .15$, with a small effect size; *the first PK* between less-successful L2 listeners ($Mdn = 8.5, n = 8$) and more-successful L2 listeners ($Mdn = 13, n = 10$), $U = 20.500, z = -1.745, p = .081, r = .41$, with a medium effect size; *the final PK* between less-successful L2 listeners ($Mdn = 9, n = 8$) and more-successful L2 listeners ($Mdn = 11.5, n = 10$), $U = 27.500, z = -1.115, p = .265, r = .26$, with a small effect size (Cohen, 1988).

On the other hand, it was found that there was a statistically difference in the scores of the *mid PK* between less-successful L2 listeners ($Mdn = 9, n = 8$) and more-successful L2 listeners in the experimental group ($Mdn = 14.5, n = 10$), $U = 11.500, z = -2.559, p < .05, r = .60$, with a large effect size (Cohen, 1988). The median score of the mid MALQ on *Person Knowledge* of more-successful L2 listeners ($Mdn = 14.5$) was significantly ranked higher than less-successful L2 listeners ($Mdn = 9$) in the experimental group.

Table 35 shows the Mann-Whitney U tests' findings applied to the MALQ scores to compare less-successful L2 listeners and more-successful L2 listeners in the control group.

Table 35

Mann-Whitney U Test on the MALQ for the Control Group regarding Level of Listening Proficiency

Group	MALQ	Level of Listening Proficiency	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Control	First MALQ	Less-successful	10	10.55	105.50	39.500	-0.454	0.650
		More-successful	9	9.39	84.50			
		Total	19					
	Mid MALQ	Less-successful	10	9.65	96.50	41.500	-0.287	0.774
		More-successful	9	10.39	93.50			
		Total	19					
	Final MALQ	Less-successful	10	10.05	100.50	44.500	-0.041	0.967
		More-successful	9	9.94	89.50			
		Total	19					
	First PS	Less-successful	10	9.75	97.50	42.500	-0.208	0.835
		More-successful	9	10.28	92.50			
		Total	19					
	Mid PS	Less-successful	10	9.00	90.00	35.500	-0.820	0.412
		More-successful	9	11.11	100.00			
		Total	19					
	Final PS	Less-successful	10	10.00	100.00	45.000	0.000	1.000
		More-successful	9	10.00	90.00			
		Total	19					
	First PE	Less-successful	10	10.15	101.50	43.500	-0.124	0.901
		More-successful	9	9.83	88.50			
		Total	19					
	Mid PE	Less-successful	10	9.70	97.00	42.000	-0.246	0.805
		More-successful	9	10.33	93.00			
		Total	19					
	Final PE	Less-successful	10	9.95	99.50	44.500	-0.041	0.967
		More-successful	9	10.06	90.50			
		Total	19					
	First MT	Less-successful	10	10.40	104.00	41.000	-0.330	0.741
		More-successful	9	9.56	86.00			
		Total	19					
	Mid MT	Less-successful	10	11.10	111.00	34.000	-0.908	0.364
		More-successful	9	8.78	79.00			
		Total	19					
	Final MT	Less-successful	10	9.85	98.50	43.500	-0.124	0.901
		More-successful	9	10.17	91.50			
		Total	19					
	First DA	Less-successful	10	10.85	108.50	36.500	-0.702	0.482
		More-successful	9	9.06	81.50			
		Total	19					
	Mid DA	Less-successful	10	11.10	111.00	34.000	-0.905	0.366
		More-successful	9	8.78	79.00			
		Total	19					
	Final DA	Less-successful	10	11.60	116.00	29.000	-1.323	0.186
		More-successful	9	8.22	74.00			
		Total	19					
	First PK	Less-successful	10	9.75	97.50	42.500	-0.206	0.837
		More-successful	9	10.28	92.50			
		Total	19					
	Mid PK	Less-successful	10	9.25	92.50	37.500	-0.616	0.538
		More-successful	9	10.83	97.50			
		Total	19					

Final PK	Less-successful	10	9.55	95.50	40.500	-0.369	0.712
	More-successful	9	10.50	94.50			

According to the Mann-Whitney U test analysis, in the control group, no significant differences were found in the scores of *the first overall MALQ* between less-successful L2 listeners ($Mdn = 87, n = 10$) and more-successful L2 listeners ($Mdn = 87, n = 9$), $U = 39.500, z = -.454, p = .650, r = .10$, with a small effect size; *the second overall MALQ* between less-successful L2 listeners ($Mdn = 85.5, n = 10$) and more-successful L2 listeners ($Mdn = 87, n = 9$), $U = 41.500, z = -.287, p = .774, r = .07$, with a small effect size; *the final overall MALQ* between less-successful L2 listeners ($Mdn = 84, n = 10$) and more-successful L2 listeners ($Mdn = 81, n = 9$), $U = 44.500, z = -.041, p = .967, r = .01$, with a small effect size; *the first PS* between less-successful L2 listeners ($Mdn = 27, n = 10$) and more-successful L2 listeners ($Mdn = 27, n = 9$), $U = 42.500, z = -.208, p = .835, r = .05$, with a small effect size; *the mid PS* between less-successful L2 listeners ($Mdn = 26.5, n = 10$) and more-successful L2 listeners ($Mdn = 28, n = 9$), $U = 35.500, z = -.820, p = .412, r = .19$, with a small effect size; *the final PS* between less-successful L2 listeners ($Mdn = 25.5, n = 10$) and more-successful L2 listeners ($Mdn = 25, n = 9$), $U = 45.000, z = .000, p = .1000, r = .00$, with a small effect size *the first PE* between less-successful L2 listeners ($Mdn = 21, n = 10$) and more-successful L2 listeners ($Mdn = 20, n = 9$), $U = 43.500, z = -.124, p = .901, r = .03$, with a small effect size; *the mid PE* between less-successful L2 listeners ($Mdn = 21, n = 10$) and more-successful L2 listeners ($Mdn = 20, n = 9$), $U = 42.000, z = -.246, p = .805, r = .06$, with a small effect size; *the final PE* between less-successful L2 listeners ($Mdn = 21, n = 10$) and more-successful L2 listeners ($Mdn = 18, n = 9$), $U = 44.500, z = -.041, p = .967, r = .01$, with a small effect size; *the first MT* between less-successful L2 listeners ($Mdn = 11, n = 10$) and more-successful L2 listeners ($Mdn = 11, n = 9$), $U = 41.000, z = -.330, p = .741, r = .08$, with a small effect size; *the mid MT* between less-successful L2 listeners ($Mdn = 9.5, n = 10$) and more-successful L2 listeners ($Mdn = 9, n = 9$), $U = 34.000, z = -.908, p = .364, r = .21$, with a small effect size; *the final MT* between less-successful L2 listeners ($Mdn = 10.5, n = 10$) and more-successful L2 listeners ($Mdn = 12, n = 9$), $U = 43.500, z = -.124, p = .901, r = .03$, with a small effect size; *the first DA* between less-successful L2 listeners ($Mdn = 20, n = 10$) and more-successful L2 listeners ($Mdn = 19, n = 9$), $U = 36.500, z = -.702, p = .482, r = .16$, with a small effect size; *the mid DA* between less-successful L2 listeners ($Mdn =$

18.5, $n = 10$) and more-successful L2 listeners ($Mdn = 18$, $n = 9$), $U = 34.00$, $z = -.905$, $p = .366$, $r = .21$, with a small effect size; *the final DA* between less-successful L2 listeners ($Mdn = 18$, $n = 10$) and more-successful L2 listeners ($Mdn = 17$, $n = 9$), $U = 29.000$, $z = -1.323$, $p = .186$, $r = .30$, with a medium effect size; *the first PK* between less-successful L2 listeners ($Mdn = 10$, $n = 10$) and more-successful L2 listeners ($Mdn = 11$, $n = 9$), $U = 42.500$, $z = -.206$, $p = .837$, $r = .05$, with a small effect size; *the mid PK* between less-successful L2 listeners ($Mdn = 9$, $n = 10$) and more-successful L2 listeners ($Mdn = 12$, $n = 9$), $U = 37.500$, $z = -.616$, $p = .538$, $r = .14$, with a small effect size; *the final PK* between less-successful L2 listeners ($Mdn = 9$, $n = 10$) and more-successful L2 listeners ($Mdn = 10$, $n = 9$), $U = 40.500$, $z = -.369$, $p = .712$, $r = .08$, with a small effect size.

As a result, it was concluded that the median score on the MALQ scores for each measurement did not show any statistically significant difference between less-successful and more-successful L2 listeners in the control group.

Analysis of the differences in the MALQ scores between the groups regarding the level of listening proficiency: the Mann-Whitney U test. The Mann-Whitney U tests were conducted to find out if there were any differences in the three MALQ scores measured at the beginning, at the midpoint, and at the end of the instruction between the experimental group and the control group by comparing them with each other in terms of the level of listening proficiency.

Table 36 shows the Mann-Whitney U tests' findings applied to the MALQ scores to compare the experimental group to the control group regarding less-successful L2 listeners.

Table 36

Mann-Whitney U Test on the MALQ Scores of Less-Successful L2 listeners between the Groups

Level of Listening Ability	MALQ	Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Less-successful	First MALQ	Experimental	8	9.13	73.00	37.000	-0.267	0.789
		Control	10	9.80	98.00			
		Total	18					
	Mid MALQ	Experimental	8	9.63	77.00	39.000	-0.089	0.929
		Control	10	9.40	94.00			
		Total	18					

Final MALQ	Experimental	8	10.06	80.50	35.500	-0.401	0.688
	Control	10	9.05	90.50			
	Total	18					
First PS	Experimental	8	9.44	75.50	39.500	-0.045	0.964
	Control	10	9.55	95.50			
	Total	18					
Mid PS	Experimental	8	9.88	79.00	37.000	-0.268	0.789
	Control	10	9.20	92.00			
	Total	18					
Final PS	Experimental	8	9.94	79.50	36.500	-0.313	0.754
	Control	10	9.15	91.50			
	Total	18					
First PE	Experimental	8	10.94	87.50	28.500	-1.033	0.302
	Control	10	8.35	83.50			
	Total	18					
Mid PE	Experimental	8	9.88	79.00	37.000	-0.269	0.788
	Control	10	9.20	92.00			
	Total	18					
Final PE	Experimental	8	10.31	82.50	33.500	-0.582	0.561
	Control	10	8.85	88.50			
	Total	18					
First MT	Experimental	8	9.38	75.00	39.000	-0.090	0.928
	Control	10	9.60	96.00			
	Total	18					
Mid MT	Experimental	8	9.75	78.00	38.000	-0.180	0.857
	Control	10	9.30	93.00			
	Total	18					
Final MT	Experimental	8	9.69	77.50	38.500	-0.134	0.893
	Control	10	9.35	93.50			
	Total	18					
First DA	Experimental	8	9.63	77.00	39.000	-0.090	0.929
	Control	10	9.40	94.00			
	Total	18					
Mid DA	Experimental	8	9.50	76.00	40.000	0.000	1.000
	Control	10	9.50	95.00			
	Total	18					
Final DA	Experimental	8	9.56	76.50	39.500	-0.045	0.964
	Control	10	9.45	94.50			
	Total	18					
First PK	Experimental	8	9.13	73.00	37.000	-0.268	0.789
	Control	10	9.80	98.00			
	Total	18					
Mid PK	Experimental	8	8.63	69.00	33.000	-0.636	0.525
	Control	10	10.20	102.00			
	Total	18					
Final PK	Experimental	8	9.13	73.00	37.000	-0.268	0.789
	Control	10	9.80	98.00			
	Total	18					

According to the Mann-Whitney U test analysis, no statistically significant differences were found by less-successful L2 listeners in the scores of *the first overall MALQ* between the experimental group ($Mdn = 84$, $n = 8$) and the control group ($Mdn = 87$, $n = 10$), $U = 37.000$, $z = -.267$, $p = .789$, $r = .06$, with a small effect size; *the second overall MALQ* between the experimental group ($Mdn = 86.5$, $n = 8$) and the control group ($Mdn = 85.5$, $n = 10$), $U = 39.000$, $z = -.089$, $p = .929$, $r = .02$, with a small effect size; *the final overall MALQ* between the experimental group

($Mdn = 85.5, n = 8$) and the control group ($Mdn = 84, n = 10$), $U = 35.500, z = -.401, p = .688, r = .09$, with a small effect size; *the first PS* between the experimental group ($Mdn = 26, n = 8$) and the control group ($Mdn = 27, n = 10$), $U = 39.500, z = -.045, p = .964, r = .01$, with a small effect size; *the mid PS* between the experimental group ($Mdn = 27.5, n = 8$) and the control group ($Mdn = 26.5, n = 10$), $U = 37.000, z = -.268, p = .789, r = .06$, with a small effect size; *the final PS* between the experimental group ($Mdn = 25.5, n = 8$) and the control group ($Mdn = 25.5, n = 10$), $U = 36.500, z = -.313, p = .754, r = .07$, with a small effect size; *the first PE* between the experimental group ($Mdn = 22, n = 8$) and the control group ($Mdn = 21, n = 10$), $U = 28.500, z = -1.033, p = .332, r = .24$, with a small effect size; *the mid PE* between the experimental group ($Mdn = 20, n = 8$) and the control group ($Mdn = 21, n = 10$), $U = 37.000, z = -.269, p = .788, r = .06$, with a small effect size; *the final PE* between the experimental group ($Mdn = 19.5, n = 8$) and the control group ($Mdn = 21, n = 10$), $U = 33.500, z = -.582, p = .561, r = .14$, with a small effect size; *the first MT* between the experimental group ($Mdn = 10, n = 8$) and the control group ($Mdn = 11, n = 10$), $U = 39.000, z = -.090, p = .928, r = .02$, with a small effect size; *the mid MT* between the experimental group ($Mdn = 10.5, n = 8$) and the control group ($Mdn = 9.5, n = 10$), $U = 38.000, z = -.180, p = .857, r = .04$, with a small effect size; *the final MT* between the experimental group ($Mdn = 11.5, n = 8$) and the control group ($Mdn = 10.5, n = 10$), $U = 38.500, z = -.134, p = .893, r = .03$, with a small effect size; *the first DA* between the experimental group ($Mdn = 19, n = 8$) and the control group ($Mdn = 20, n = 10$), $U = 39.000, z = -.090, p = .929, r = .02$, with a small effect size; *the mid DA* between the experimental group ($Mdn = 19, n = 8$) and the control group ($Mdn = 18.5, n = 10$), $U = 40.000, z = .000, p = 1.000, r = .00$, with a small effect size; *the final DA* between the experimental group ($Mdn = 18, n = 8$) and the control group ($Mdn = 18, n = 10$), $U = 39.500, z = -.045, p = .964, r = .01$, with a small effect size; *the first PK* between the experimental group ($Mdn = 8.5, n = 8$) and the control group ($Mdn = 9, n = 10$), $U = 37.000, z = -.268, p = .789, r = .06$, with a small effect size; *the mid PK* between the experimental group ($Mdn = 9, n = 8$) and the control group ($Mdn = 9, n = 10$), $U = 33.000, z = -.636, p = .525, r = .15$, with a small effect size; *the final PK* between the experimental group ($Mdn = 9, n = 8$) and the control group ($Mdn = 7, n = 10$), $U = 37.000, z = -.268, p = .789, r = .06$, with a small effect size.

As a result, it was concluded that there were no statistically significant differences in the median scores of the MALQ of less-successful L2 listeners between the experimental and control groups.

Table 37 presents the Mann-Whitney U test analysis conducted to explore any differences in the MALQ scores of more-successful L2 listeners between the experimental group and the control group over the process of metacognitive instruction.

Table 37

Mann-Whitney U Test on the MALQ Scores of More-Successful L2 Listeners between the Groups

Level of Listening Proficiency	MALQ	Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
More-successful	First MALQ	Experimental	10	12.15	121.50	23.500	-1.760	0.078
		Control	9	7.61	68.50			
		Total	19					
	Mid MALQ	Experimental	10	11.40	114.00	31.000	-1.144	0.253
		Control	9	8.44	76.00			
		Total	19					
	Final MALQ	Experimental	10	11.65	116.50	28.500	-1.350	0.177
		Control	9	8.17	73.50			
		Total	19					
	First PS	Experimental	10	11.75	117.50	27.500	-1.446	0.148
		Control	9	8.06	72.50			
		Total	19					
	Mid PS	Experimental	10	11.45	114.50	30.500	-1.207	0.227
		Control	9	8.39	75.50			
		Total	19					
	Final PS	Experimental	10	12.00	120.00	25.000	-1.638	0.101
		Control	9	7.78	70.00			
		Total	19					
	First PE	Experimental	10	10.95	109.50	35.500	-0.783	0.434
		Control	9	8.94	80.50			
		Total	19					
	Mid PE	Experimental	10	9.20	92.00	37.000	-0.656	0.512
		Control	9	10.89	98.00			
		Total	19					
	Final PE	Experimental	10	11.45	114.50	30.500	-1.201	0.230
		Control	9	8.39	75.50			
		Total	19					
	First MT	Experimental	10	10.50	105.00	40.000	-0.413	0.680
		Control	9	9.44	85.00			
		Total	19					
	Mid MT	Experimental	10	10.60	106.00	39.000	-0.491	0.623
		Control	9	9.33	84.00			
		Total	19					
	Final MT	Experimental	10	9.35	93.50	38.500	-0.533	0.594
		Control	9	10.72	96.50			

	Total	19					
First DA	Experimental	10	10.30	103.00	42.000	-0.247	0.805
	Control	9	9.67	87.00			
	Total	19					
Mid DA	Experimental	10	11.15	111.50	33.500	-0.946	0.344
	Control	9	8.72	78.50			
	Total	19					
Final DA	Experimental	10	13.05	130.50	14.500	-2.511	0.012
	Control	9	6.61	59.50			
	Total	19					
First PK	Experimental	10	11.65	116.50	28.500	-1.355	0.175
	Control	9	8.17	73.50			
	Total	19					
Mid PK	Experimental	10	12.15	121.50	23.500	-1.773	0.076
	Control	9	7.61	68.50			
	Total	19					
Final PK	Experimental	10	11.15	111.50	33.500	-0.944	0.345
	Control	9	8.72	78.50			

According to the Mann-Whitney U test analysis, no significant differences were found for more-successful L2 listeners in the scores of *the first overall MALQ* between the experimental group ($Mdn = 95.5, n = 10$) and the control group ($Mdn = 87, n = 9$), $U = 23.500, z = -1.760, p = .078, r = .40$, with a medium effect size; *the second overall MALQ* between the experimental group ($Mdn = 91, n = 10$) and the control group ($Mdn = 87, n = 9$), $U = 31.000, z = -1.144, p = .253, r = .26$, with a small effect size; *the final overall MALQ* between the experimental group ($Mdn = 92, n = 10$) and the control group ($Mdn = 81, n = 9$), $U = 28.500, z = -1.350, p = .177, r = .31$, with a medium effect size; *the first PS* between the experimental group ($Mdn = 30, n = 10$) and the control group ($Mdn = 27, n = 9$), $U = 27.500, z = -1.446, p = .148, r = .33$, with a medium effect size; *the mid PS* between the experimental group ($Mdn = 31, n = 10$) and the control group ($Mdn = 28, n = 9$), $U = 30.500, z = -1.207, p = .227, r = .28$, with a small effect size; *the final PS* between the experimental group ($Mdn = 30.5, n = 10$) and the control group ($Mdn = 25, n = 9$), $U = 25.000, z = -1.638, p = .101, r = .36$, with a medium effect size; *the first PE* between the experimental group ($Mdn = 20.5, n = 10$) and the control group ($Mdn = 20, n = 9$), $U = 35.500, z = -.783, p = .434, r = .18$, with a small effect size; *the mid PE* between the experimental group ($Mdn = 19, n = 10$) and the control group ($Mdn = 20, n = 9$), $U = 37.000, z = -.656, p = .512, r = .15$, with a small effect size; *the final PE* between the experimental group ($Mdn = 20.5, n = 10$) and the control group ($Mdn = 18, n = 9$), $U = 30.500, z = -1.201, p = .230, r = .28$, with a small effect size; *the first MT* between the experimental group ($Mdn = 11.5, n = 10$) and the control group ($Mdn = 11, n = 9$), $U = 40.000, z = -.413, p = .680, r = .09$, with a small effect size; *the mid*

MT between the experimental group ($Mdn = 11, n = 10$) and the control group ($Mdn = 9, n = 9$), $U = 39.000, z = -.491, p = .623, r = .11$, with a small effect size; *the final MT* between the experimental group ($Mdn = 8.5, n = 10$) and the control group ($Mdn = 12, n = 9$), $U = 38.500, z = -.533, p = .594, r = .12$, with a small effect size; *the first DA* between the experimental group ($Mdn = 19, n = 10$) and the control group ($Mdn = 19, n = 9$), $U = 42.000, z = -.247, p = .805, r = .06$, with a small effect size; *the mid DA* between the experimental group ($Mdn = 19, n = 10$) and the control group ($Mdn = 18, n = 9$), $U = 33.500, z = -.946, p = .344, r = .22$, with a small effect size; *the first PK* between the experimental group ($Mdn = 13, n = 10$) and the control group ($Mdn = 11, n = 9$), $U = 28.500, z = -1.355, p = .175, r = .31$, with a small effect size; *the mid PK* between the experimental group ($Mdn = 14.5, n = 10$) and the control group ($Mdn = 12, n = 9$), $U = 23.500, z = -1.773, p = .076, r = .41$, with a medium effect size; *the final PK* between the experimental group ($Mdn = 11.5, n = 10$) and the control group ($Mdn = 10, n = 9$), $U = 33.500, z = -.944, p = .345, r = .22$, with a small effect size (Cohen,1988).

On the other hand, it was found out there was a statistically significant difference in the scores of *the final DA* of more-successful L2 listeners between the experimental group ($Mdn = 19, n = 10$) and the control group ($Mdn = 17, n = 9$), $U = 14.500, z = -2.511, p < .05, r = .58$, with a large effect size. Most-successful L2 listeners in the experimental group's median score on Directed Attention measured at the end of the metacognitive process was significantly higher ranked than most-successful L2 listeners in the control group's median score.

R. Q. 2. 2. Are there any statistical differences in L2 listeners' metacognitive awareness of listening over metacognitive instruction regarding preparatory school backgrounds? To answer the research question above, at first, the Friedman test as the nonparametric version of mixed between-within subjects analysis of variance (ANOVA), because the data did not meet the assumptions of mixed between-within subjects ANOVA except for the assumptions of normality and homogeneity of variances, was conducted to show if there were any differences in L2 listeners' metacognitive awareness within and between the experimental group and the control group regarding different preparatory school backgrounds over the process of metacognitive instruction.

The independent variables are Group (experimental group and control group), which is the between-subjects factor, and Time including the measures of MALQ at the beginning (time point 1), at the midpoint (time point 2), and at the end (time point 3) of the metacognitive instruction and Preparatory school background (yes and no), which are the within-subjects factors, while the metacognitive awareness of L2 listeners over the process of metacognitive instruction is the dependent factor. In order to measure their metacognitive awareness of listening, the total values of the overall MALQ and the subdimensions of the MALQ including Planning & Evaluation, Problem Solving, Monitoring, Directed Attention, Mental Translation, and Person Knowledge for each time point were calculated.

Since the Friedman test does not provide a post hoc test, a Wilcoxon signed ranks test with Bonferroni adjustment is conducted to see the source of those differences if there are any differences in the results.

After analyzing the Friedman test within the groups, to explore if there were any differences between the experimental group and the control group regarding different preparatory school backgrounds in the scores of the MALQ over metacognitive instruction, the Mann-Whitney U tests were applied.

Analysis of the differences in the MALQ scores within the groups regarding preparatory school background: the Friedman test. Firstly, the descriptive statistics of the MALQ scores are presented in Table 38 for the experimental group.

Table 38

Descriptive Statistics of the MALQ for the Experimental Group regarding Preparatory School Background

Group	Preparatory school background	MALQ	N	Mean	Std. Deviation	Min.	Max.	Median
Experimental	Yes	First MALQ	6	91.1667	7.30525	83.00	99.00	91.5000
		Mid MALQ	6	87.0000	5.29150	81.00	94.00	86.5000
		Final MALQ	6	88.6667	10.55778	68.00	96.00	92.0000
		First PS	6	28.8333	3.18852	25.00	33.00	30.0000
		Mid PS	6	29.3333	3.01109	25.00	32.00	31.0000
		Final PS	6	28.0000	5.96657	18.00	34.00	30.0000
		First PE	6	20.3333	2.58199	18.00	24.00	19.5000
		Mid PE	6	18.0000	1.54919	15.00	19.00	18.5000
		Final PE	6	20.0000	1.09545	18.00	21.00	20.0000
		First MT	6	11.6667	2.73252	8.00	16.00	11.5000
		Mid MT	6	11.6667	3.32666	7.00	16.00	12.0000

No	Final MT	6	11.6667	3.61478	7.00	16.00	12.0000
	First DA	6	18.0000	2.09762	15.00	20.00	18.5000
	Mid DA	6	16.6667	2.06559	14.00	19.00	16.5000
	Final DA	6	18.3333	1.63299	16.00	21.00	18.0000
	First PK	6	12.3333	2.94392	8.00	16.00	13.0000
	Mid PK	6	11.3333	4.13118	7.00	16.00	11.5000
	Final PK	6	10.6667	5.04645	4.00	16.00	11.5000
	First MALQ	12	91.4167	11.46900	77.00	111.00	88.5000
	Mid MALQ	12	90.1667	13.42205	72.00	111.00	91.0000
	Final MALQ	12	90.8333	12.51787	73.00	118.00	89.0000
	First PS	12	28.8333	4.50925	21.00	36.00	28.5000
	Mid PS	12	28.8333	5.49104	19.00	36.00	30.0000
	Final PS	12	29.5000	4.66125	23.00	36.00	29.5000
	First PE	12	21.3333	4.37624	11.00	26.00	22.0000
	Mid PE	12	20.7500	3.49350	14.00	27.00	20.5000
	Final PE	12	21.7500	3.91094	17.00	30.00	21.0000
	First MT	12	10.8333	4.26046	5.00	17.00	10.0000
	Mid MT	12	10.2500	4.53522	5.00	17.00	9.5000
	Final MT	12	9.6667	5.15811	4.00	18.00	9.0000
	First DA	12	19.6667	2.30940	16.00	23.00	19.5000
	Mid DA	12	19.1667	2.65718	14.00	22.00	19.5000
	Final DA	12	19.9167	3.42340	14.00	24.00	20.5000
	First PK	12	10.7500	3.62128	4.00	15.00	11.0000
	Mid PK	12	11.1667	3.51188	5.00	16.00	11.0000
	Final PK	12	10.0000	3.64318	5.00	16.00	10.5000

Secondly, the findings of the Friedman test applied to show if there were any changes in the MALQ scores of L2 listeners who studied at the preparatory school before and who did not study at the preparatory school across the three-time points (at the beginning, midpoint, and end of instruction) over the process of metacognitive instruction were presented by the groups in different tables.

Table 39 shows the findings of the Friedman test applied to demonstrate whether there were any changes in the MALQ scores of L2 listeners in the experimental group across the three-time points (at the beginning, midpoint, and end of instruction) regarding different preparatory school backgrounds.

Table 39

Friedman Test on the MALQ within the Experimental Group regarding Preparatory School Background

Group	Preparatory school background	MALQ	N	Mean Rank	Friedman χ^2	df	Asymp. Sig.
Experimental	Yes	First MALQ	6	2.33	1.826	2	0.401
		Mid MALQ	6	1.58			

No	Final MALQ	6	2.08			
	First PS	6	2.00	1.333	2	0.513
	Mid PS	6	2.33			
	Final PS	6	1.67			
	First PE	6	2.00	2.273	2	0.321
	Mid PE	6	1.58			
	Final PE	6	2.42			
	First MT	6	2.08	0.700	2	0.705
	Mid MT	6	2.17			
	Final MT	6	1.75			
	First DA	6	2.08	4.095	2	0.129
	Mid DA	6	1.42			
	Final DA	6	2.50			
	First PK	6	2.50	3.647	2	0.161
	Mid PK	6	1.92			
	Final PK	6	1.58			
	First MALQ	12	2.08	0.130	2	0.937
	Mid MALQ	12	1.96			
	Final MALQ	12	1.96			
	First PS	12	1.88	0.864	2	0.649
	Mid PS	12	1.92			
	Final PS	12	2.21			
	First PE	12	2.00	0.400	2	0.819
	Mid PE	12	1.88			
	Final PE	12	2.13			
	First MT	12	2.13	0.703	2	0.704
	Mid MT	12	2.04			
	Final MT	12	1.83			
	First DA	12	2.13	0.974	2	0.614
	Mid DA	12	1.79			
	Final DA	12	2.08			
	First PK	12	2.33	3.200	2	0.202
	Mid PK	12	2.00			
	Final PK	12	1.67			

The Friedman test presented that no significant differences were found between L2 listeners in the experimental group who studied at the preparatory school before in the scores of *overall MALQ* across the three time points [beginning ($Mdn = 91.5$), mid ($Mdn = 86.5$), and end ($Mdn = 92$)], [$\chi^2 (2, n = 6) = 1.826, p = .401$]; of *Problem Solving* across the three time points [beginning ($Mdn = 30$), mid ($Mdn = 31$), and end ($Mdn = 30$)], [$\chi^2 (2, n = 6) = 1.333, p = .513$]; of *Planning and Evaluation* across the three time points [beginning ($Mdn = 19.5$), mid ($Mdn = 18.5$), and end ($Mdn = 20$)], [$\chi^2 (2, n = 6) = 2.273, p = .321$]; of *Mental Translation* across the three time points [beginning ($Mdn = 11.5$), mid ($Mdn = 12$), and end ($Mdn = 12$)], [$\chi^2 (2, n = 6) = .700, p = .705$]; of *Directed Attention* across the three time points

[beginning (*Mdn* = 18.5), mid (*Mdn* = 16.5), and end (*Mdn* = 18)], [χ^2 (2, *n* = 6) = 4.095, *p* = .129]; of *Person Knowledge* across the three time points [beginning (*Mdn* = 13), mid (*Mdn* = 11.5), and end (*Mdn* = 11.5)], [χ^2 (2, *n* = 6) = 3.647, *p* = .161].

Similarly, no significant differences were found between L2 listeners in the experimental group who did not study at the preparatory school before in the scores of *overall MALQ* across the three time points [beginning (*Mdn* = 88.5), mid (*Mdn* = 91), and end (*Mdn* = 89)], [χ^2 (2, *n* = 12) = .130, *p* = .937]; *Problem Solving* across the three time points [beginning (*Mdn* = 28.5), mid (*Mdn* = 30), and end (*Mdn* = 29.5)], [χ^2 (2, *n* = 12) = .864, *p* = .649]; *Planning and Evaluation* across the three time points [beginning (*Mdn* = 22), mid (*Mdn* = 20.5), and end (*Mdn* = 21)], [χ^2 (2, *n* = 12) = .400, *p* = .819]; *Mental Translation* across the three time points [beginning (*Mdn* = 10), mid (*Mdn* = 9.5), and end (*Mdn* = 9)], [χ^2 (2, *n* = 12) = .703, *p* = .704]; *Directed Attention* across the three time points [beginning (*Mdn* = 19.5), mid (*Mdn* = 19.5), and end (*Mdn* = 20.5)], [χ^2 (2, *n* = 12) = .974, *p* = .614]; *Person Knowledge* across the three time points [beginning (*Mdn* = 11), mid (*Mdn* = 11), and end (*Mdn* = 10.5)], [χ^2 (2, *n* = 12) = 3.200, *p* = .202].

After analyzing the experimental group, the descriptive statistics of the MALQ scores are presented in Table 40 for the control group regarding different preparatory school backgrounds.

Table 40

Descriptive Statistics of the MALQ for the Control Group regarding Preparatory School Background

Group	Preparatory school background	MALQ	N	Mean	Std. Deviation	Min.	Max.	Median
Control	Yes	First MALQ	9	81.7778	8.51143	65.00	94.00	82.0000
		Mid MALQ	9	81.2222	11.41028	63.00	101.00	83.0000
		Final MALQ	9	77.4444	14.40582	56.00	105.00	78.0000
		First PS	9	25.8889	2.31541	23.00	29.00	26.0000
		Mid PS	9	26.7778	2.94863	22.00	32.00	26.0000
		Final PS	9	24.6667	4.76970	19.00	35.00	23.0000
		First PE	9	18.7778	2.86259	14.00	23.00	18.0000
		Mid PE	9	20.5556	2.18581	17.00	23.00	21.0000
		Final PE	9	18.8889	3.91933	15.00	26.00	16.0000
		First MT	9	10.6667	1.93649	7.00	13.00	11.0000
		Mid MT	9	9.0000	2.00000	5.00	12.00	9.0000

No	Final MT	9	10.6667	2.39792	6.00	13.00	11.0000
	First DA	9	17.7778	3.45607	11.00	21.00	20.0000
	Mid DA	9	16.0000	3.60555	10.00	22.00	15.0000
	Final DA	9	15.7778	4.17665	9.00	21.00	17.0000
	First PK	9	8.6667	4.18330	4.00	15.00	10.0000
	Mid PK	9	8.8889	4.91031	3.00	15.00	8.0000
	Final PK	9	7.4444	3.60940	3.00	14.00	7.0000
	First MALQ	10	90.1000	8.00625	74.00	104.00	91.0000
	Mid MALQ	10	87.9000	8.81224	72.00	100.00	87.0000
	Final MALQ	10	86.2000	8.71525	68.00	101.00	86.0000
	First PS	10	28.0000	3.94405	18.00	32.00	29.0000
	Mid PS	10	27.3000	4.87739	17.00	33.00	28.5000
	Final PS	10	26.2000	4.70933	16.00	33.00	25.0000
	First PE	10	20.8000	2.65832	15.00	24.00	21.0000
	Mid PE	10	19.8000	3.96653	15.00	26.00	20.0000
	Final PE	10	20.2000	3.79473	13.00	25.00	20.5000
	First MT	10	11.0000	2.78887	7.00	15.00	11.0000
	Mid MT	10	10.9000	2.96086	7.00	15.00	10.5000
	Final MT	10	10.7000	2.66875	7.00	15.00	11.0000
	First DA	10	19.5000	2.22361	15.00	22.00	19.5000
	Mid DA	10	19.1000	1.91195	16.00	22.00	19.0000
	Final DA	10	18.1000	1.59513	16.00	20.00	18.0000
	First PK	10	10.8000	2.93636	6.00	16.00	10.0000
	Mid PK	10	10.8000	2.20101	8.00	15.00	10.5000
	Final PK	10	11.0000	2.74874	7.00	15.00	11.0000

For the control group, the findings of the Friedman test applied to determine any changes in the MALQ scores of L2 listeners across the three-time points (at the beginning, midpoint, and end of instruction) regarding different preparatory school backgrounds were presented in Table 41.

Table 41

Friedman Test on the MALQ within the Control Group regarding Preparatory School Background

Group	Preparatory school background	MALQ	N	Mean Rank	Friedman χ^2	df	Asymp. Sig.
Control	Yes	First MALQ	9	2.22	2.387	2	0.303
		Mid MALQ	9	2.17			
		Final MALQ	9	1.61			
		First PS	9	2.22	3.200	2	0.202
		Mid PS	9	2.22			
		Final PS	9	1.56			
		First PE	9	1.94	3.455	2	0.178
		Mid PE	9	2.44			
		Final PE	9	1.61			
		First MT	9	2.22	6.200	2	0.045

No	Mid MT	9	1.39			
	Final MT	9	2.39			
	First DA	9	2.56	4.688	2	0.096
	Mid DA	9	1.72			
	Final DA	9	1.72			
	First PK	9	2.17	0.897	2	0.639
	Mid PK	9	2.06			
	Final PK	9	1.78			
	First MALQ	10	2.70	9.297	2	0.010
	Mid MALQ	10	1.90			
	Final MALQ	10	1.40			
	First PS	10	2.45	5.688	2	0.058
	Mid PS	10	2.05			
	Final PS	10	1.50			
	First PE	10	2.15	0.500	2	0.779
	Mid PE	10	1.85			
	Final PE	10	2.00			
	First MT	10	1.90	0.364	2	0.834
	Mid MT	10	2.10			
	Final MT	10	2.00			
	First DA	10	2.20	3.297	2	0.192
	Mid DA	10	2.25			
	Final DA	10	1.55			
	First PK	10	2.05	0.437	2	0.804
	Mid PK	10	1.85			
	Final PK	10	2.10			

According to the Friedman test, no significant differences were determined between L2 listeners in the control group who studied at the preparatory school before in the scores of *Overall MALQ* across the three time points [beginning (*Mdn* = 82), mid (*Mdn* = 83), and end (*Mdn* = 78)], [χ^2 (2, *n* = 9) = 2.387, *p* = .303]; *Problem Solving* across the three time points [beginning (*Mdn* = 26), mid (*Mdn* = 26), and end (*Mdn* = 23)], [χ^2 (2, *n* = 9) = 3.200, *p* = .202]; *Planning and Evaluation* across the three time points [beginning (*Mdn* = 18), mid (*Mdn* = 21), and end (*Mdn* = 16)], [χ^2 (2, *n* = 9) = 3.455, *p* = .178]; *Directed Attention* across the three time points [beginning (*Mdn* = 20), mid (*Mdn* = 15), and end (*Mdn* = 17)], [χ^2 (2, *n* = 9) = 4.688, *p* = .096]; *Person Knowledge* across the three time points [beginning (*Mdn* = 10), mid (*Mdn* = 8), and end (*Mdn* = 7)], [χ^2 (2, *n* = 9) = .897, *p* = .639].

However, it was found that there was a significant difference in the scores of *Mental Translation* across the three-time points [beginning (*Mdn* = 11), mid (*Mdn* = 9), and end (*Mdn* = 11)], [χ^2 (2, *n* = 9) = 6.200, *p* = .045] between L2 listeners in the control group who studied at the preparatory school before. The median values

showed a decrease in the control group's metacognitive awareness in terms of mental translation from beginning to the midpoint of the process metacognitive instruction and then increased from the midpoint to end of the process.

In order to explore in which time points the significant difference in the scores of Mental Translation for L2 listeners in the control group who had a preparatory school background took place, the Wilcoxon signed ranks tests with Bonferroni adjustment ($p = .05/3 = .0167$) were conducted and presented in Table 42.

Table 42

Wilcoxon Signed Ranks Tests on Mental Translation for the Control Group regarding Different Preparatory School Background

	Preparatory school background	MALQ		N	Mean Rank	Sum of Ranks	Z	Adjusted p (0.0167)
Control Group	Yes	Mid MT-First MT	Negative Ranks	6 ^b	4.42	26.50	-2.124 ^a	0.034
			Positive Ranks	1 ^c	1.50	1.50		
			Ties	2 ^d				
		Final MT - Mid MT	Negative Ranks	1 ^e	2.50	2.50	-2.200 ^k	0.028
			Positive Ranks	7 ^f	4.79	33.50		
			Ties	1 ^g				
		Final MT -First MT	Negative Ranks	3 ^h	4.83	14.50	-.085 ^a	0.932
			Positive Ranks	4 ⁱ	3.38	13.50		
			Ties	2 ^j				

a. Based on negative ranks. k. Based on positive ranks.
 b. Mid MT < First MT e. Final MT < Mid MT h. Final MT < First MT
 c. Mid MT > First MT f. Final MT > Mid MT i. Final MT > First MT
 d. Mid MT = First MT g. Final MT = Mid MT j. Final MT = First MT

According to the Wilcoxon signed ranks with Bonferroni adjustment, it was concluded that no significant differences were found, according to the adjusted p value (.0167), between the first MT and the mid MT, $z = -2.124$, $p = .034$, with a medium effect size ($r = .50$); between the mid MT and the final MT, $z = -2.200$, $p = .028$, with a large effect size ($r = .52$); between the first MT and the final MT, $z = -.085$, $p = .932$, with a small effect size ($r = .02$) (Cohen, 1988).

On the other hand, according to the Friedman test, no significant differences were found between L2 listeners in the control group who did not study at the preparatory school before in the scores of; *Problem Solving* across the three time

points [beginning (*Mdn* = 29), mid (*Mdn* = 28.5), and end (*Mdn* = 25)], [χ^2 (2, *n* = 10) = 5.688 *p* = .058]; *Planning and Evaluation* across the three time points [beginning (*Mdn* = 21), mid (*Mdn* = 20), and end (*Mdn* = 20.5)], [χ^2 (2, *n* = 10) = .500, *p* = .779]; *Mental Translation* across the three time points [beginning (*Mdn* = 11), mid (*Mdn* = 10.5), and end (*Mdn* = 11)], [χ^2 (2, *n* = 10) = .364, *p* = .834]; *Directed Attention* across the three time points [beginning (*Mdn* = 19.5), mid (*Mdn* = 19), and end (*Mdn* = 18)], [χ^2 (2, *n* = 10) = 3.297, *p* = .192]; *Person Knowledge* across the three time points [beginning (*Mdn* = 10), mid (*Mdn* = 10.5), and end (*Mdn* = 11)], [χ^2 (2, *n* = 10) = .437, *p* = .804].

Nevertheless, it was found that there was a significant difference in the scores of the *overall MALQ* across the three-time points [beginning (*Mdn* = 91), mid (*Mdn* = 87), and end (*Mdn* = 86)], [χ^2 (2, *n* = 10) = 9.297, *p* < .05] between L2 listeners in the control group who did not study at the preparatory school before.

In order to explore in which time points the difference in the scores of the *overall MALQ* for L2 listeners in the control group who did not have any preparatory school background took place, the Wilcoxon signed ranks tests with Bonferroni adjustment (*p* = .05/3 = 0.0167) were conducted and presented in Table 43.

Table 43

Wilcoxon Signed Ranks Tests of Overall MALQ for the Control Group regarding Preparatory School Background

	Preparatory school background	MALQ		N	Mean Rank	Sum of Ranks	Adjusted Z	p (0.0167)
Control Group	No	Mid MALQ-First MALQ	Negative Ranks	7 ^b	5.00	35.00	-1.486 ^a	0.137
			Positive Ranks	2 ^c	5.00	10.00		
		Final MALQ - Mid MALQ	Ties	1 ^d				
			Negative Ranks	6 ^e	5.92	35.50	-1.548 ^a	0.122
		Final MALQ -First MALQ	Positive Ranks	3 ^f	3.17	9.50		
			Ties	1 ^g				
			Negative Ranks	9 ^h	5.00	45.00	-2.684 ^a	0.007
			Positive Ranks	0 ⁱ	0.00	0.00		
			Ties	1 ^j				

a. Based on positive ranks.
b. Mid MALQ < First MALQ e. Final MALQ < Mid MALQ h. Final MALQ < First MALQ
MALQ

c. Mid MALQ > First MALQ f. Final MALQ > Mid MALQ i. Final MALQ > First MALQ
d. Mid MALQ = First MALQ g. Final MALQ = Mid MALQ j. Final MALQ = First MALQ

According to the Wilcoxon signed ranks with Bonferroni adjustment, no significant differences were found, according to the adjusted p -value, between the first MALQ and the mid MALQ, $z = -1.486$, $p = .137$, with a medium effect size ($r = .33$); between the mid MALQ and the final MALQ, $z = -1.548$, $p = .122$, with a medium effect size ($r = .35$) (Cohen, 1988). However, it could be deduced that there was a significant difference between the first MALQ and the final MALQ, $z = -2.684$, $p < .0167$, with a large effect size ($r = .60$), and therefore, it was concluded that there was a decrease in the overall metacognitive awareness of the L2 listeners in the control group who did not study at the preparatory school before from the beginning of ($Mdn = 91$) to the midpoint ($Mdn = 87$) and end ($Mdn = 86$) of metacognitive instruction.

Analysis of the differences in the MALQ scores within the groups regarding preparatory school backgrounds: the Mann-Whitney U test. The Mann-Whitney U tests were conducted to reveal any differences in the three MALQ scores measured at the beginning, midpoint, and end of the instruction within the experimental group and the control group by comparing L2 listeners with different preparatory school backgrounds.

Table 44 shows the Mann-Whitney U tests' findings applied on the MALQ scores compared to L2 listeners who studied at the preparatory school before and L2 listeners who did not study at the preparatory school in the experimental group.

Table 44

Mann-Whitney U Test on the MALQ of the Experimental Group regarding Preparatory School Background

Group		Preparatory school background	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Experimental	First MALQ	Yes	6	9.67	58.00	35.000	-0.094	0.928
		No	12	9.42	113.00			
		Total	18					
	Mid MALQ	Yes	6	8.83	53.00	32.000	-0.376	0.707
		No	12	9.83	118.00			
		Total	18					
	Final MALQ	Yes	6	9.58	57.50	35.500	-0.047	0.963
		No	12	9.46	113.50			
		Total	18					

First PS	Yes	6	9.17	55.00	34.000	-0.190	0.849
	No	12	9.67	116.00			
	Total	18					
Mid PS	Yes	6	9.67	58.00	35.000	-0.095	0.924
	No	12	9.42	113.00			
	Total	18					
Final PS	Yes	6	8.92	53.50	32.500	-0.329	0.742
	No	12	9.79	117.50			
	Total	18					
First PE	Yes	6	7.83	47.00	26.000	-0.943	0.345
	No	12	10.33	124.00			
	Total	18					
Mid PE	Yes	6	5.58	33.50	12.500	-2.219	0.026
	No	12	11.46	137.50			
	Total	18					
Final PE	Yes	6	8.33	50.00	29.000	-0.665	0.506
	No	12	10.08	121.00			
	Total	18					
First MT	Yes	6	10.42	62.50	30.500	-0.522	0.601
	No	12	9.04	108.50			
	Total	18					
Mid MT	Yes	6	10.75	64.50	28.500	-0.705	0.481
	No	12	8.88	106.50			
	Total	18					
Final MT	Yes	6	11.08	66.50	26.500	-0.893	0.372
	No	12	8.71	104.50			
	Total	18					
First DA	Yes	6	7.25	43.50	22.500	-1.278	0.201
	No	12	10.63	127.50			
	Total	18					
Mid DA	Yes	6	6.08	36.50	15.500	-1.948	0.051
	No	12	11.21	134.50			
	Total	18					
Final DA	Yes	6	7.17	43.00	22.000	-1.319	0.187
	No	12	10.67	128.00			
	Total	18					
First PK	Yes	6	10.92	65.50	27.500	-0.802	0.423
	No	12	8.79	105.50			
	Total	18					
Mid PK	Yes	6	9.58	57.50	35.500	-0.047	0.962
	No	12	9.46	113.50			
	Total	18					
Final PK	Yes	6	10.00	60.00	33.000	-0.282	0.778
	No	12	9.25	111.00			

According to the Mann-Whitney U test analysis, in the experimental group, no significant differences were obtained in the scores of *the first overall MALQ* between L2 listeners who studied at the preparatory school before ($Mdn = 91.5$, $n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 88.5$, $n = 12$), $U = 35.500$, $z = -0.094$, $p = .928$, $r = .02$, with a small effect size; *the second overall MALQ* between L2 listeners who studied at the preparatory school before ($Mdn = 86.5$, $n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 91$, $n = 12$), $U = 32.000$, $z = -.376$, $p = .707$, $r = .09$, with a small effect size; *the final overall MALQ* between L2 listeners who studied at the preparatory

school before ($Mdn = 92, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 89, n = 12$), $U = 35.500, z = -.047, p = .963, r = .01$, with a small effect size; *the first PS* between L2 listeners who studied at the preparatory school before ($Mdn = 30, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 28.5, n = 12$), $U = 34.000, z = -.190, p = .849, r = .05$, with a small effect size; *the mid PS* between L2 listeners who studied at the preparatory school before ($Mdn = 31, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 30, n = 12$), $U = 35.000, z = -.095, p = .924, r = .22$, with a small effect size; *the final PS* between L2 listeners who studied at the preparatory school before ($Mdn = 30, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 29.5, n = 12$), $U = 32.500, z = -.329, p = .742, r = .08$, with a small effect size; *the first PE* between L2 listeners who studied at the preparatory school before ($Mdn = 19.5, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 22, n = 12$), $U = 26.000, z = -.943, p = .345, r = .22$, with a small effect size; *the final PE* between L2 listeners who studied at the preparatory school before ($Mdn = 20, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 21, n = 12$), $U = 29.000, z = -.665, p = .506, r = .17$, with a small effect size; *the first MT* between L2 listeners who studied at the preparatory school before ($Mdn = 11.5, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 10, n = 12$), $U = 30.500, z = -.522, p = .601, r = .12$, with a small effect size; *the mid MT* between L2 listeners who studied at the preparatory school before ($Mdn = 12, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 9.5, n = 12$), $U = 28.500, z = -.705, p = .481, r = .17$, with a small effect size; *the final MT* between L2 listeners who studied at the preparatory school before ($Mdn = 12, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 9, n = 12$), $U = 26.500, z = -.893, p = .372, r = .21$, with a small effect size; *the first DA* between L2 listeners who studied at the preparatory school before ($Mdn = 18.5, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 19.5, n = 12$), $U = 22.500, z = -1.278, p = .201, r = .30$, with a medium effect size; *the final DA* between L2 listeners who studied at the preparatory school before ($Mdn = 18, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 20.5, n = 12$), $U = 22.000, z = -1.319, p = .187, r = .31$, with a medium effect size; *the first PK* between L2 listeners who studied at the preparatory school before ($Mdn = 13, n = 6$) and L2 listeners who did not study

at the preparatory school before ($Mdn = 11, n = 12$), $U = 27.500, z = -.802, p = .423, r = .19$, with a small effect size; *the mid PK* between L2 listeners who studied at the preparatory school before ($Mdn = 11.5, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 11, n = 12$), $U = 35.500, z = -.047, p = .962, r = .01$, with a small effect size; *the final PK* between L2 listeners who studied at the preparatory school before ($Mdn = 11.5, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 10.5, n = 12$), $U = 33.500, z = -.282, p = .778, r = .12$, with a small effect size (Cohen, 1988).

On the other hand, it was found that there was a statistically significant difference in the scores of *the mid PE* between L2 listeners who studied at the preparatory school before ($Mdn = 18.5, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 20.5, n = 12$), $U = 12.500, z = -2.219, p < .05, r = .52$, with a large effect size (Cohen, 1988).

Also, although the difference in the scores of the mid DA between L2 listeners who studied at the preparatory school before ($Mdn = 16.5, n = 6$) and L2 listeners who did not study at the preparatory school before ($Mdn = 19.5, n = 12$) was not quite statistically significant, $U = 15.500, z = -1.948, p = .051, p > .50, r = .46$, with medium effect size, it could be concluded that the median score on the metacognitive awareness of L2 listeners who did not study at the preparatory school ($Mdn = 16.5$) was ranked lower than L2 listeners who studied at the preparatory school before ($Mdn = 19.5$) in terms of the mid DA because of the small difference between p-values.

After the Mann-Whitney U tests which were conducted to find out if there were any differences in the three MALQ scores measured at the beginning, at the midpoint, and end of the instruction within the experimental group by comparing L2 listeners with different preparatory school backgrounds, another Mann-Whitney U test was conducted for the control group.

Table 45 shows the Mann-Whitney U tests' findings applied on the MALQ scores comparing L2 listeners who studied at the preparatory school before and L2 listeners who did not study at the preparatory school before in the control group.

Table 45

Mann-Whitney U Test on the MALQ for the Control Group regarding Preparatory School Background

Group	MALQ	Preparatory school background	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Control	First MALQ	Yes	9	7.28	65.50	20.500	-2.022	0.043
		No	10	12.45	124.50			
		Total	19					
	Mid MALQ	Yes	9	8.28	74.50	29.500	-1.272	0.203
		No	10	11.55	115.50			
		Total	19					
	Final MALQ	Yes	9	7.56	68.00	23.000	-1.799	0.072
		No	10	12.20	122.00			
		Total	19					
	First PS	Yes	9	7.33	66.00	21.000	-1.997	0.046
		No	10	12.40	124.00			
		Total	19					
	Mid PS	Yes	9	9.28	83.50	38.500	-0.533	0.594
		No	10	10.65	106.50			
		Total	19					
	Final PS	Yes	9	8.67	78.00	33.000	-0.985	0.324
		No	10	11.20	112.00			
		Total	19					
	First PE	Yes	9	7.94	71.50	26.500	-1.527	0.127
		No	10	11.85	118.50			
		Total	19					
	Mid PE	Yes	9	10.61	95.50	39.500	-0.451	0.652
		No	10	9.45	94.50			
		Total	19					
	Final PE	Yes	9	8.72	78.50	33.500	-0.946	0.344
		No	10	11.15	111.50			
		Total	19					
	First MT	Yes	9	9.56	86.00	41.000	-0.330	0.741
		No	10	10.40	104.00			
		Total	19					
	Mid MT	Yes	9	8.39	75.50	30.500	-1.197	0.231
		No	10	11.45	114.50			
		Total	19					
	Final MT	Yes	9	10.17	91.50	43.500	-0.124	0.901
		No	10	9.85	98.50			
		Total	19					
	First DA	Yes	9	8.67	78.00	33.000	-0.992	0.321
		No	10	11.20	112.00			
		Total	19					
	Mid DA	Yes	9	7.28	65.50	20.500	-2.016	0.044
		No	10	12.45	124.50			
		Total	19					
	Final DA	Yes	9	8.61	77.50	32.500	-1.033	0.301
		No	10	11.25	112.50			
		Total	19					
	First PK	Yes	9	8.50	76.50	31.500	-1.111	0.267
		No	10	11.35	113.50			
		Total	19					
	Mid PK	Yes	9	8.89	80.00	35.000	-0.821	0.412
		No	10	11.00	110.00			

	Total	19					
Final PK	Yes	9	7.33	66.00	21.000	-1.970	0.049
	No	10	12.40	124.00			

According to the Mann-Whitney U test analysis, in the control group, no significant differences were found in the scores of *the second overall MALQ* between L2 listeners who studied at the preparatory school before ($Mdn = 83, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 87, n = 10$), $U = 29.500, z = -1.272, p = .203, r = .29$, with a small effect size; *the final overall MALQ* between L2 listeners who studied at the preparatory school before ($Mdn = 78, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 86, n = 10$), $U = 23.000, z = -1.799, p = .072, r = .41$, with an almost medium effect size; *the mid PS* between L2 listeners who studied at the preparatory school before ($Mdn = 26, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 28.5, n = 10$), $U = 38.500, z = -.533, p = .594, r = .12$, with a small effect size; *the final PS* between L2 listeners who studied at the preparatory school before ($Mdn = 23, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 25, n = 10$), $U = 33.000, z = -.985, p = .324, r = .23$, with a small effect size; *the first PE* between L2 listeners who studied at the preparatory school before ($Mdn = 18, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 21, n = 10$), $U = 26.500, z = -1.527, p = .127, r = .35$, with a medium effect size; *the mid PE* between L2 listeners who studied at the preparatory school before ($Mdn = 21, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 20, n = 10$), $U = 39.500, z = -.451, p = .652, r = .10$, with a small effect size; *the final PE* between L2 listeners who studied at the preparatory school before ($Mdn = 16, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 20.5, n = 10$), $U = 33.500, z = -.946, p = .344, r = .22$, with a small effect size; *the first MT* between L2 listeners who studied at the preparatory school before ($Mdn = 11, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 11, n = 10$), $U = 41.000, z = -.330, p = .741, r = .08$, with a small effect size; *the mid MT* between L2 listeners who studied at the preparatory school before ($Mdn = 9, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 10.5, n = 10$), $U = 30.500, z = -1.197, p = .231, r = .27$, with a small effect size; *the final MT* between L2 listeners who studied at the preparatory school before ($Mdn = 11, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 11, n = 10$), $U = 43.500, z = -.124, p = .901, r = .03$, with a small effect size;

the first DA between L2 listeners who studied at the preparatory school before ($Mdn = 20, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 19.5, n = 10$), $U = 33.000, z = -.992, p = .321, r = .23$, with a small effect size; *the final DA* between L2 listeners who studied at the preparatory school before ($Mdn = 17, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 18, n = 10$), $U = 32.500, z = -1.033, p = .301, r = .24$, with a small effect size; *the first PK* between L2 listeners who studied at the preparatory school before ($Mdn = 10, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 10, n = 10$), $U = 31.500, z = -1.111, p = .267, r = .25$, with a small effect size; *the mid PK* between L2 listeners who studied at the preparatory school before ($Mdn = 8, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 10.5, n = 10$), $U = 35.000, z = -.821, p = .412, r = .19$, with a small effect size.

On the other hand, it was found that there was a statistically difference in the scores of; *the first overall MALQ* between L2 listeners who studied at the preparatory school before ($Mdn = 82, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 91, n = 110$), $U = 20.500, z = -2.022, p < .05, r = .46$, with a medium effect size; *the first PS* between L2 listeners who studied at the preparatory school before ($Mdn = 26, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 29, n = 10$), $U = 21.000, z = -1.997, p < .05, r = .46$, with a medium effect size; *the mid DA* between L2 listeners who studied at the preparatory school before ($Mdn = 15, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 19, n = 10$), $U = 20.500, z = -2.016, p < .05, r = .46$, with a medium effect size; *the final PK* between L2 listeners who studied at the preparatory school before ($Mdn = 7, n = 9$) and L2 listeners who did not study at the preparatory school before ($Mdn = 11, n = 10$), $U = 21.000, z = -1.970, p < .05, r = .45$, with a medium effect size (Cohen, 1988).

As a result, it was concluded that the median score on the first overall MALQ, first PS, mid DA, and final PK of L2 listeners who did not study at the preparatory school was ranked higher than L2 listeners who studied at the preparatory school before in the control group.

Analysis of the differences in the MALQ scores between the groups regarding preparatory school background: the Mann-Whitney U test. The Mann-Whitney U tests were conducted to determine any changes in the three MALQ

scores measured at the beginning, midpoint, and end of the instruction between the experimental and control groups with each other in terms of different preparatory school backgrounds.

Table 46 shows the Mann-Whitney U tests' findings applied to the MALQ scores by comparing the experimental group to the control group regarding L2 listeners who studied at the preparatory school before.

Table 46

Mann-Whitney U Test on the MALQ between the Groups for L2 listeners who studied at the Preparatory School Before

Preparatory school background	MALQ	Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2-tailed)
Yes	First MALQ	Experimental	6	10.75	64.50	10.500	-1.966	0.049
		Control	9	6.17	55.50			
		Total	15					
	Mid MALQ	Experimental	6	9.58	57.50	17.500	-1.122	0.262
		Control	9	6.94	62.50			
		Total	15					
	Final MALQ	Experimental	6	10.58	63.50	11.500	-1.832	0.067
		Control	9	6.28	56.50			
		Total	15					
	First PS	Experimental	6	10.67	64.00	11.000	-1.904	0.057
		Control	9	6.22	56.00			
		Total	15					
	Mid PS	Experimental	6	9.83	59.00	16.000	-1.315	0.188
		Control	9	6.78	61.00			
		Total	15					
	Final PS	Experimental	6	9.67	58.00	17.000	-1.181	0.238
		Control	9	6.89	62.00			
		Total	15					
	First PE	Experimental	6	9.50	57.00	18.000	-1.082	0.279
		Control	9	7.00	63.00			
		Total	15					
	Mid PE	Experimental	6	5.25	31.50	10.500	-1.973	0.049
		Control	9	9.83	88.50			
		Total	15					
	Final PE	Experimental	6	8.67	52.00	23.000	-0.480	0.632
		Control	9	7.56	68.00			
		Total	15					
	First MT	Experimental	6	8.83	53.00	22.000	-0.598	0.550
		Control	9	7.44	67.00			
		Total	15					
	Mid MT	Experimental	6	10.17	61.00	14.000	-1.549	0.121
		Control	9	6.56	59.00			
		Total	15					
	Final MT	Experimental	6	8.67	52.00	23.000	-0.475	0.635
		Control	9	7.56	68.00			
		Total	15					
	First DA	Experimental	6	7.42	44.50	23.500	-0.422	0.673
		Control	9	8.39	75.50			
		Total	15					

Mid DA	Experimental	6	8.75	52.50	22.500	-0.535	0.593
	Control	9	7.50	67.50			
	Total	15					
Final DA	Experimental	6	9.67	58.00	17.000	-1.202	0.229
	Control	9	6.89	62.00			
	Total	15					
First PK	Experimental	6	10.25	61.50	13.500	-1.602	0.109
	Control	9	6.50	58.50			
	Total	15					
Mid PK	Experimental	6	9.75	58.50	16.500	-1.245	0.213
	Control	9	6.83	61.50			
	Total	15					
Final PK	Experimental	6	9.75	58.50	16.500	-1.246	0.213
	Control	9	6.83	61.50			
	Total	15					

According to the Mann-Whitney U test analysis, there were no significant differences for L2 listeners who studied at the preparatory school before in the scores of *the second overall MALQ* between the experimental group ($Mdn = 86.5$, $n = 6$) and the control group ($Mdn = 83$, $n = 9$), $U = 17.500$, $z = -1.122$, $p = .262$, $r = .29$, with a small effect size; *the final overall MALQ* between the experimental group ($Mdn = 92$, $n = 6$) and the control group ($Mdn = 78$, $n = 9$), $U = 11.500$, $z = -1.832$, $p = .067$, $r = .47$, with a medium effect size; *the first PS* between the experimental group ($Mdn = 30$, $n = 6$) and the control group ($Mdn = 26$, $n = 9$), $U = 11.000$, $z = -1.904$, $p = .057$, $r = .49$, with a medium effect size; *the mid PS* between the experimental group ($Mdn = 31$, $n = 6$) and the control group ($Mdn = 26$, $n = 9$), $U = 16.000$, $z = -1.315$, $p = .188$, $r = .34$, with a medium effect size; *the final PS* between the experimental group ($Mdn = 30$, $n = 6$) and the control group ($Mdn = 23$, $n = 9$), $U = 17.000$, $z = -1.181$, $p = .238$, $r = .31$, with a medium effect size; *the first PE* between the experimental group ($Mdn = 19.5$, $n = 6$) and the control group ($Mdn = 18$, $n = 9$), $U = 18.000$, $z = -1.082$, $p = .279$, $r = .28$, with a small effect size; *the final PE* between the experimental group ($Mdn = 20$, $n = 6$) and the control group ($Mdn = 16$, $n = 9$), $U = 23.000$, $z = -.480$, $p = .632$, $r = .12$, with a small effect size; *the first MT* between the experimental group ($Mdn = 11.5$, $n = 6$) and the control group ($Mdn = 11$, $n = 9$), $U = 22.000$, $z = -.598$, $p = .550$, $r = .15$, with a small effect size; *the mid MT* between the experimental group ($Mdn = 12$, $n = 6$) and the control group ($Mdn = 9$, $n = 9$), $U = 14.000$, $z = -1.549$, $p = .121$, $r = .40$, with a medium effect size; *the final MT* between the experimental group ($Mdn = 12$, $n = 6$) and the control group ($Mdn = 11$, $n = 9$), $U = 23.000$, $z = -.475$, $p = .635$, $r = .12$, with a small effect size; *the first DA* between the experimental group ($Mdn = 18.5$, $n = 6$) and the control group ($Mdn = 20$, $n = 9$), $U = 23.500$, $z = -.422$, $p = .673$, $r = .11$, with a small effect size; *the mid DA* between the experimental group ($Mdn = 16.5$, $n = 6$) and the control

group ($Mdn = 15$, $n = 9$), $U = 22.500$, $z = -.535$, $p = .593$, $r = .14$, with a small effect size; *the final DA* between the experimental group ($Mdn = 18$, $n = 6$) and the control group ($Mdn = 17$, $n = 9$), $U = 17.000$, $z = -1.202$, $p = .229$, $r = .31$, with a medium effect size; *the first PK* between the experimental group ($Mdn = 13$, $n = 6$) and the control group ($Mdn = 10$, $n = 9$), $U = 13.500$, $z = -1.602$, $p = .109$, $r = .41$, with a medium effect size; *the mid PK* between the experimental group ($Mdn = 11.5$, $n = 6$) and the control group ($Mdn = 8$, $n = 9$), $U = 16.500$, $z = -1.245$, $p = .213$, $r = .32$, with a medium effect size; *the final PK* between the experimental group ($Mdn = 11.5$, $n = 6$) and the control group ($Mdn = 7$, $n = 9$), $U = 16.500$, $z = -1.246$, $p = .213$, $r = .32$, with a medium effect size;

On the other hand, it was demonstrated that there was a statistically significant difference for L2 listeners who studied at the preparatory school before in the scores of: *the first overall MALQ* between the experimental group ($Mdn = 91.5$, $n = 6$) and the control group ($Mdn = 82$, $n = 9$), $U = 10.500$, $z = -1.966$, $p < .05$, $r = .51$, with a large effect size; *the mid PE* between the experimental group ($Mdn = 18.5$, $n = 6$) and the control group ($Mdn = 21$, $n = 9$), $U = 10.500$, $z = -1.973$, $p < .05$, $r = .51$, with a large effect size (Cohen, 1988).

As a result, it was concluded that the median score on the first overall MALQ of L2 listeners in the experimental group who studied at the preparatory school was ranked higher than L2 listeners in the control group who studied at the preparatory school before while L2 listeners in the control group who studied at the preparatory school was ranked higher than the experimental group's median score on the mid PE of MALQ.

Table 47 presents the Mann-Whitney U test analysis conducted to explore any differences in the MALQ scores over the process of metacognitive instruction of L2 listeners who did not study at the preparatory school between the experimental group and the control group.

Table 47

Mann-Whitney U Test on the MALQ between the Groups for L2 listeners who did not study at the Preparatory School Before

Preparatory school background	MALQ	Group	N	Mean Rank	Sum of Ranks	U	Z	Asymp. Sig. (2- tailed)
-------------------------------------	------	-------	---	--------------	--------------------	---	---	-------------------------------

No	First MALQ	Experimental	12	11.54	138.50	59.500	-0.033	0.974
		Control	10	11.45	114.50			
		Total	22					
	Mid MALQ	Experimental	12	12.25	147.00	51.000	-0.595	0.552
		Control	10	10.60	106.00			
		Total	22					
	Final MALQ	Experimental	12	12.33	148.00	50.000	-0.660	0.509
		Control	10	10.50	105.00			
		Total	22					
	First PS	Experimental	12	11.42	137.00	59.000	-0.066	0.947
		Control	10	11.60	116.00			
		Total	22					
	Mid PS	Experimental	12	12.46	149.50	48.500	-0.761	0.447
		Control	10	10.35	103.50			
		Total	22					
	Final PS	Experimental	12	13.25	159.00	39.000	-1.398	0.162
		Control	10	9.40	94.00			
		Total	22					
	First PE	Experimental	12	12.29	147.50	50.500	-0.632	0.527
		Control	10	10.55	105.50			
		Total	22					
	Mid PE	Experimental	12	12.13	145.50	52.500	-0.497	0.620
		Control	10	10.75	107.50			
		Total	22					
	Final PE	Experimental	12	12.33	148.00	50.000	-0.663	0.507
		Control	10	10.50	105.00			
		Total	22					
	First MT	Experimental	12	11.17	134.00	56.000	-0.266	0.790
		Control	10	11.90	119.00			
		Total	22					
	Mid MT	Experimental	12	10.88	130.50	52.500	-0.497	0.691
		Control	10	12.25	122.50			
		Total	22					
	Final MT	Experimental	12	10.50	126.00	48.000	-0.794	0.427
		Control	10	12.70	127.00			
		Total	22					
	First DA	Experimental	12	11.63	139.50	58.500	-0.100	0.920
		Control	10	11.35	113.50			
		Total	22					
	Mid DA	Experimental	12	12.00	144.00	54.000	-0.400	0.689
		Control	10	10.90	109.00			
		Total	22					
	Final DA	Experimental	12	13.38	160.50	37.500	-1.493	0.135
		Control	10	9.25	92.50			
		Total	22					
	First PK	Experimental	12	11.71	140.50	57.500	-0.166	0.868
		Control	10	11.25	112.50			
		Total	22					
	Mid PK	Experimental	12	11.96	143.50	54.000	-0.367	0.713
		Control	10	10.95	109.50			
		Total	22					
	Final PK	Experimental	12	10.54	126.50	48.500	-0.761	0.447
		Control	10	12.65	126.50			
		Total	22					

According to the Mann-Whitney U test analysis, no significant differences were found for L2 listeners who did not study at the preparatory school before in the scores of: *the first overall MALQ* between the experimental group (Mdn = 88.5, n =

12) and the control group (Mdn = 91, n = 10), $U = 59.500$, $z = -.033$, $p = .974$, $r = .01$, with a small effect size; *the second overall MALQ* between the experimental group (Mdn = 91, n = 12) and the control group (Mdn = 87, n = 10), $U = 51.000$, $z = -.595$, $p = .552$, $r = .13$, with a small effect size; *the final overall MALQ* between the experimental group (Mdn = 89, n = 12) and the control group (Mdn = 86, n = 10), $U = 50.000$, $z = -.660$, $p = .509$, $r = .14$, with a small effect size; *the first PS* between the experimental group (Mdn = 28.5, n = 12) and the control group (Mdn = 29, n = 10), $U = 59.000$, $z = -.066$, $p = .947$, $r = .01$, with a small effect size; *the mid PS* between the experimental group (Mdn = 30, n = 12) and the control group (Mdn = 28.5, n = 10), $U = 48.500$, $z = -.761$, $p = .447$, $r = .16$, with a small effect size; *the final PS* between the experimental group (Mdn = 29.5, n = 12) and the control group (Mdn = 25, n = 10), $U = 39.000$, $z = -1.398$, $p = .162$, $r = .30$, with a small effect size; *the first PE* between the experimental group (Mdn = 22, n = 12) and the control group (Mdn = 21, n = 10), $U = 50.500$, $z = -.632$, $p = .527$, $r = .13$, with a small effect size; *the mid PE* between the experimental group (Mdn = 20.5, n = 12) and the control group (Mdn = 20, n = 10), $U = 52.500$, $z = -.497$, $p = .620$, $r = .11$, with a small effect size; *the final PE* between the experimental group (Mdn = 21, n = 12) and the control group (Mdn = 20.5, n = 10), $U = 50.000$, $z = -.663$, $p = .507$, $r = .14$, with a small effect size; *the first MT* between the experimental group (Mdn = 10, n = 12) and the control group (Mdn = 11, n = 10), $U = 56.000$, $z = -.266$, $p = .790$, $r = .06$, with a small effect size; *the mid MT* between the experimental group (Mdn = 9.5, n = 12) and the control group (Mdn = 10.5, n = 10), $U = 52.500$, $z = -.497$, $p = .691$, $r = .11$, with a small effect size; *the final MT* between the experimental group (Mdn = 9, n = 12) and the control group (Mdn = 11, n = 10), $U = 48.000$, $z = -.794$, $p = .427$, $r = .17$, with a small effect size; *the first DA* between the experimental group (Mdn = 19.5, n = 12) and the control group (Mdn = 19.5, n = 10), $U = 58.500$, $z = -.100$, $p = .920$, $r = .02$, with a small effect size; *the mid DA* between the experimental group (Mdn = 19.5, n = 12) and the control group (Mdn = 19, n = 10), $U = 54.000$, $z = -.400$, $p = .689$, $r = .09$, with a small effect size; *the final DA* between the experimental group (Mdn = 20.5, n = 12) and the control group (Mdn = 18, n = 10), $U = 37.500$, $z = -1.493$, $p = .135$, $r = .32$, with a medium effect size; *the first PK* between the experimental group (Mdn = 11, n = 12) and the control group (Mdn = 10, n = 10), $U = 57.500$, $z = -.166$, $p = .868$, $r = .04$, with a small effect size; *the mid PK* between the experimental group (Mdn = 11, n = 12) and the control group (Mdn

= 10.5, n = 10), $U = 54.500$, $z = -.367$, $p = .713$, $r = .08$, with a small effect size; *the final PK* between the experimental group (Mdn = 10.5, n = 12) and the control group (Mdn = 11, n = 10), $U = 48.500$, $z = -.761$, $p = .447$, $r = .16$, with a small effect size (Cohen,1988).

Qualitative Findings

R. Q. 2.3. What are the perceptions of L2 listeners on possible changes in metacognitive awareness of listening over the process of metacognitive instruction? The data collected through the stimulated recall protocols of four volunteers from the experimental group who joined to discuss possible changes in their responses to the first, mid and final MALQs, was transcribed and analyzed via a qualitative analysis program called *Atlas.ti*. The participants were named Alex (less-successful-less-successful), Taylor (less-successful-more-successful), Paul (less-successful-more-successful), and Martin (more-successful-more-successful), and their descriptive information was given in Table 48. The participants' names were followed by a parenthesis in which their proficiency level at the beginning and the end of the study was respectively written; therefore, while reading and analyzing this study, the progress in listening performance of those participants could easily be realized.

Table 48

Descriptive Information of the Participants of Stimulated Recall Protocols

Participants	Gender	Level of Listening Proficiency at the beginning of the study	Level of Listening Proficiency at the end of the study	Preparatory school background
Alex	Male	Less-successful	Less-successful	No
Taylor	Female	Less-successful	More-successful	No
Paul	Male	Less-successful	More-successful	No
Martin	Male	More-successful	More-successful	Yes

The analysis was carried out using the four L2 listeners' responses to the MALQ during the first and second stimulated protocols to determine positive and negative changes in their metacognitive awareness of listening over metacognitive instruction. Themes were constructed based on the subdimensions of the MALQ, which are Planning and Evaluation, Problem Solving, Mental Translation, Directed Attention, Person Knowledge, and related subtitles of the subdimensions. The data collected was analyzed respectively based on the order of the stimulated recall protocols, the direction of changes (positive or negative), themes on the

subdimensions of the MALQ, and related subthemes on the subdimensions. Then, significant quotations related to the themes and related subthemes on the subdimensions of the MALQ were presented to demonstrate existing positive and negative changes in the metacognitive awareness of the participants over metacognitive instruction.

Positive changes from the first stimulated recall protocol. At first, when analyzed the first stimulated recall protocol in which four participants from the experimental group joined to discuss possible changes in their responses to the first and the second MALQs, it was found out that positive changes in the responses to the second MALQ were observed regarding Planning and Evaluation, Problem Solving, Mental Translation, and Directed Attention, while no change was found related to Person Knowledge.

Related Codes:

– is part of → • Metacognitive Awareness

Considering Planning and Evaluation, positive changes in following certain listening goals and making ready to listen through applying planning, prediction, self-check, and evaluation strategies were observed.

Regarding following certain listening goals,

Quotations

Taylor (less-successful-more-successful): “At first, I thought that I just listened and kept going like that; this wasn’t a problem for me, but I focused more on listening, understanding, and developing this skill. That’s why having a goal has become more important for me; that’s why the response on having a goal in mind became definite.”

Related codes:

– is associated with → • following certain listening goals

Regarding making ready to listen through applying planning, prediction, self-check, and evaluation strategies,

Quotations

Alex (less-successful-less-successful): “After the first listening test, when I learned that I had many mistakes in the test ... I just focused on the general context of the text, I mean, only listening. Then, I made a plan as I would focus on keywords; that’s why I changed my response.

Paul (less-successful-more-successful): “While listening, what I had thought about the title of the listening text came to my mind. In the beginning, I had not written much on the part of My Prediction, now I have started to think and write much more on the part of My Prediction; thus, I matched, compared my predictions with the listening text much more.

Taylor (less-successful-more-successful): “Umm, I mean, now I know better how to listen to concentrate, umm, I mean, topics on concentration. For example, there is an introduction part of the listening text; generally, I don’t focus on that part or don’t take notes. I focused on the part in which the main subject was talked, I focused on certain parts, so I have learned how to do it.

Related codes:

– is associated with → • making ready to listen through applying planning, prediction, self-check, and evaluation strategies

Considering Problem Solving, positive changes in correctly evaluating one’s understanding while listening and reviewing one’s interpretation when needed, guessing the meaning of unknown words based on the meaning of known words, and regularly controlling the reasonability of one’s interpretation based on background knowledge on the topic were observed.

Regarding correctly evaluating understanding while listening and reviewing one’s interpretation when needed,

Quotations

Paul (less-successful-more-successful): “At the beginning of the practice, I did not assess, check or edit my understanding, but then I started to do it; this practice guided me to do it.”

Related codes:

– is associated with → • correctly evaluating understanding while listening and reviewing one’s interpretation when needed

Regarding guessing the meaning of unknown words based on the meaning of known words,

Quotations

Paul (less-successful-more-successful): "I mean, as I realize that I completely understand a sentence while listening, I guess other words. I make some guesses saying, 'There are some words that I know; what other words could be?'"

Related codes:

– is associated with → • guessing the meaning of unknown words based on the meaning of known words

Regarding regularly controlling the reasonability of one's interpretation based on background knowledge on the topic,

Quotations

Paul (less-successful-more-successful): "While listening, what I had thought about the title of the listening text came to my mind while listening at that moment."

Related codes:

– is associated with → • regularly controlling the reasonability of one's interpretation based on background knowledge on the topic,

Considering Mental Translation, positive changes on the dependence on translation were observed:

Regarding the dependence on translation,

Quotations

Taylor (less-successful-more-successful): "I don't do translations anymore; I keep listening as long as I understand it at that moment."

Related codes:

– is associated with → • the dependence on translation

Considering Directed Attention, positive changes in going on listening over difficulties in understanding and sustaining concentration while listening were explained.

Regarding going on listening over difficulties in understanding,

Quotations

Paul (less-successful-more-successful): “In the beginning, I thought that ‘I do not understand it, why am I keeping listening?’... Sometimes, because of the accent of the speaker, I had given up, but now I am not giving up, I am keeping going listening.”

Related codes:

– is associated with → ● going on listening over difficulties in understanding

Regarding sustaining concentration while listening,

Quotations

Paul (less-successful-more-successful): “Firstly, I had absolutely difficulties in concentration. When I lost track of it while listening, I would completely give up. Now, it is getting easier to recover my concentration; still, it is difficult for me. My mind is generally gone very easily while listening, but it is good; it is better now.”

Related codes:

– is associated with → ● sustaining concentration while listening

A scheme of all the themes, the subdimensions, and the subtitles on positive changes of the first stimulated recall protocol is given in Figure 2.

Negative changes from first stimulated recall protocol. It was found out that in the subdimensions of Planning and Evaluation, Problem Solving, Mental Translation, Person Knowledge, and Directed Attention, negative changes in the responses to the second MALQ were observed.

Related codes:

– is part of → ● Metacognitive Awareness

Considering Planning and Evaluation, negative changes on being conscious of one’s level of understanding while listening, bringing to mind one’s knowledge or previous experiences on the subject of the listening text, following certain listening goals, constant reflection of the reasons for listening problems were found out.

Regarding being conscious of one’s level of understanding while listening,

Quotations

Paul (less-successful-more-successful): “My self-confidence in listening is improving day by day. That’s why the response has been changed as partly asking myself my level of understanding while listening.”



Figure 2. Positive changes from first stimulated recall protocol

Related codes:

– is associated with → • being conscious of one’s level of understanding while listening

Regarding bringing to mind one’s knowledge or previous experiences on the subject of the listening text,

Quotations

Paul (less-successful-more-successful): "Because I don't believe that it is quite useful to do it. I mean, they are not much similar to earlier listening subjects I had listened to before. All of them are generally different."

Taylor (less-successful-more-successful): "Umm, this is about which I found out different subjects when I considered earlier experiences. I deduced nonsense conclusions from the different subjects I knew, that's why I should focus on only that text or mainly some certain points of that text."

Taylor (less-successful-more-successful): "The part of My Prediction sometimes negatively affects. The subject can be different from my predictions, or for example, I can have an experience like I watched a TV-series or a documentary about the subject, but it would go against it, and my predictions wouldn't work and so on."

Related codes:

– is associated with → • bringing to mind one's knowledge or previous experiences on the subject of the listening text

Regarding following certain listening goals,

Quotations

Alex (less-successful-less-successful): "I thought that I should listen to only the existing text, not focus on anything; therefore, my concentration would not be lost. I mean, for example, if I had a goal like translating word-by-word, I would lose my concentration on other parts of the text. I don't set any listening goals, and I do as much as I can."

Paul (less-successful-more-successful): "I mean, if I strongly agreed with it, I would focus on only that goal. When I did not understand anything on that goal, I lost track of other listening text parts. That's why I am not set listening goals anymore. That's why I have changed my response."

Related codes:

– is associated with → • following certain listening goals

Regarding constant reflection of the reasons of listening problems,

Alex (less-successful-less-successful): "At the beginning of the practice, frankly, I thought about my listening experience, what change I would do in the future listening because I was doing it for the first time. I wanted to try different listening methods. However, later, because I gradually found out my method, I started not to think of what I would do differently for the next listening. That's why my response has been changed negatively."

Related codes:

– is associated with → • constant reflection of the reasons of listening problems,

Considering Problem Solving, negative changes in correctly evaluating one's understanding while listening and reviewing one's interpretation when needed, using one's background knowledge and experiences to interpret understanding while listening, regularly controlling the reasonability of one's interpretation were explained.

Regarding correctly evaluating one's understanding while listening and immediately reviewing one's interpretation when needed,

Quotations

Alex (less-successful-less-successful): "Maybe the word "quickly" caused my response to be changed. Maybe, it is because when I tried to adjust my interpretation quickly, I did not understand the rest of the text. It was not quick, I mean, I put it at the back of my mind, I mean, I tried to adjust it at the end of the listening practice, I mean, that's why the response has been changed."

Related codes:

– is associated with → • correctly evaluating one's understanding while listening and immediately reviewing one's interpretation when needed

Regarding using one's background knowledge and experiences to interpret understanding while listening,

Quotations

Taylor (less-successful-more-successful): "When the subject of the listening text changes, in fact, my idea changes. I mean, the subjects become different, I

cannot use my knowledge, I do not have any background knowledge about that subject, that's why I had to go on using the conclusions I deduced from the title, from the text completely sometimes... The part of My Prediction sometimes negatively affects me. The subject can be different from my predictions, or for example, I can have an experience like I watched a TV-series or a documentary about the subject, but it would go against it, and my predictions wouldn't work and so on."

Related codes:

– is associated with → • using one's background knowledge and experiences to interpret one's understanding while listening

Regarding regularly controlling the reasonability of one's interpretation based on background knowledge on the topic,

Quotations

Taylor (less-successful-more-successful): "I mean, I don't need to make comparisons between what I comprehended and what I already knew about the text anymore because my ideas are generally close to them."

Related codes:

– is associated with → • regularly controlling the reasonability of one's interpretation

Considering Mental Translation, negative changes in making key words translation consciously were shown.

Regarding making key words translation consciously,

Quotations

Alex (less-successful-less-successful): "Here, I tried to apply a different method. Generally, I tried to focus on a sentence completely. During the first listening practice, I tried to deduce the meaning of the sentence as a whole, not word-by-word, but I wanted to do a thing during the last listening activity to hear every word. I mean, I cannot translate every word, but I would predict the sentence's meaning even if I heard each word in a sentence. For this reason, I don't do word-

by-word translation, but I choose some words and try to do their translation in my mind, so I could deduce its general meaning.”

Related codes:

– is associated with → • making key words translation consciously,

Considering Directed Attention, negative changes in sustaining concentration while listening were observed.

Regarding sustaining concentration while listening,

Alex (less-successful-less-successful): “I don’t know; it generally changes. It doesn’t happen like that every time. Sometimes, I recover my concentration; sometimes, I do not. That why I gave a different response.”

Alex (less-successful-less-successful): “I thought that I should listen to only the existing text, not focus on anything. Therefore, my concentration would not be lost, mean, for example, if I had a goal like translating word-by-word, I would lose my concentration on other parts of the text. I do not set any listening goals, and I do as much as I can.”

Martin (more-successful-more-successful): “I get bored very easily. I don’t like being unsuccessful in something because of my personality. Even if it is a basic listening activity, I lose my concentration very easily. If I had difficulty in understanding a word or an accent, it becomes challenging to recover my concentration from that moment. And frankly, knowing a chance for the second listening also affects this situation. Especially during the first listening, losing my concentration happens a lot. Because I know a chance to recover it, I do not care about the first listening. I just concentrate on keywords. And during the second listening, I make some guesses.”

Taylor (less-successful-more-successful): “Yeah, actually, this is because of the accent. It is a personal problem. My attention is getting lost in time. I mean, it is because I have too much in my head. If not, I would focus on it. I mean, I would recover it very easily. However, I could not do it now. I mean, it is completely a personal thing... My mind was sometimes gone; still, I was not aware of it at that moment.”

Related codes:

– is associated with → • sustaining concentration while listening

Considering Person Knowledge, negative changes in higher levels of difficulties in L2 listening and feeling nervous while listening were found.

Regarding higher levels of L2 listening difficulties,

Quotations

Taylor (less-successful-more-successful): “This response has negatively been changed somehow. I have seen reading texts. I have read a lot in English, or I have started to write an essay in English. They are easier for me, but listening has become more difficult, maybe because of the listening texts. I can get the general meaning of the texts, but ... I couldn’t get every detail, you know, that’s why I think listening sounds more challenging than other skills. At the beginning of the process, I was not aware of it because they all had the same level of difficulty for me.”

Martin (more-successful-more-successful): “...One of the reasons for it is the accent of the speaker.”

Related codes:

– is associated with → • higher levels of difficulties in L2 listening

Regarding feeling nervous while listening,

Quotations

Martin (more-successful-more-successful): “If I am in a situation of face-to-face communication, I never feel nervous because I have an opportunity to ask the speaker about the parts I couldn’t get. However, the listening text is listened through a tape, and I have no chance to replay it several times; yeah, I sometimes feel nervous.”

Taylor (less-successful-more-successful): “Because now I am afraid of making incorrect deductions, misunderstanding. At first, I was not aware of it; I assumed that I could listen and understand in English. However, day by day, I see I am not good at it, but now I feel scared of not understanding it completely; you know, it is like I never listen, speak and understand in English. I am scared of it.”

Related codes:

– is associated with → • feeling nervous while listening

A scheme of all the themes, the subdimensions, and the subtitles on negative changes from the first stimulated recall protocol is given in Figure 3.

Positive changes from the second stimulated recall protocol. The second stimulated recall protocol in which the same four participants from the experimental group joined to discuss possible changes on their responses to the second and final MALQs and to make suggestions on having a better L2 listening comprehension skill was analyzed in terms of the subdimensions of Planning and Evaluation, Problem Solving, Mental Translation, and Directed Attention and related subtitles.

Related codes:

– is part of → • Metacognitive Awareness

Considering Planning and Evaluation, positive changes in following certain listening goals and making ready to listen through applying planning, prediction, self-check and evaluation strategies, being conscious of one's level of understanding while listening, bringing to mind one's knowledge or previous experiences on the subject of the listening text and constant reflection of the reasons of listening problems were found out.

Regarding following certain listening goals,

Quotations

Martin (more-successful-more-successful): “Yes, academically, you know, when listening is practiced, the goal can be preparing for next activities, I mean, for reflection, for exams; however, while listening in other times, the goal in mind can be changed according to the listening text. While listening to a podcast, yes, we can have a goal, but while listening to music, some songs, it does not happen.”

Related codes:

– is associated with → • following certain listening goals,

Regarding making ready to listen through applying planning, prediction, self-check, and evaluation strategies,

Quotations

Paul (less-successful-more-successful): "When I saw the questions of the listening test, I made a plan about the parts on which I needed to focus much more."

Taylor (less-successful-more-successful): "I have learned how to listen. So, when you write the title of the listening text on the board, I think that now something happens. I mean, there are some parts I should focus on, now focus on that part, then, during the second listening, focus on the details and so on... I make my plan like that."

Related codes:

– is associated with → • making ready to listen through applying planning, prediction, self-check and evaluation strategies,

Regarding being conscious of one's level of understanding while listening,

Quotations

Paul (less-successful-more-successful): "Now, it becomes more... I wasn't listening in English too often before. When I have started to listen in English very often, my self-confidence has been improving, I mean. Now, I am telling myself I am listening in English, and I am doing it."

Martin (more-successful-more-successful): "Yes, gradually, I consider that making such evaluation is necessary, because the problem is always not about the environment or other people, but sometimes it is me personally. Therefore, it is better to turn back and assess what the problem is and try to fix it in order to make healthy communication."

Taylor (less-successful-more-successful): "There, I say, you know, I thought that I understood it. Later, something happened. When I saw the transcript of the listening text, I told myself that I misunderstood that part. However, you know, as time goes by, I assess myself like that I listened to it as a whole, then I understood those words like that, but those words were different in the text, they are, they should be like that, maybe it is because of their pronunciation."



Figure 3. Negative changes from the first stimulated recall protocol

Taylor (less-successful-more-successful): "When I checked my notes (listening diaries) on before and during listening, I told myself that I took more notes on that part than those parts, and I asked myself why that happened, or that when I did not understand anything in the conversation, I asked myself why this happened. I asked myself if I did not concentrate on the listening text or, even so, if I did not understand anything during it. I compared my notes with earlier ones, you know, I know where I had difficulty understanding, so I could focus on those during second listening. I mean, as I said, it is related to recognize yourself. If I realize what I get and what I don't get, I focus much more on that part."

Related codes:

– is associated with → • being conscious of one's level of understanding while listening

Regarding bringing to mind one's knowledge or previous experiences on the subject of the listening text,

Quotations

Taylor (less-successful-more-successful): "At first, I looked at my notes (listening dairies), then I thought others. I said they were similar... I tried to accommodate them with each other. In the past, in high school, we had some tests or some readings; I started to compare the listening text with my earlier knowledge about it. Among the last listening texts, there was one about an aircraft crash. When I heard the subject of the listening text, I remembered that I read something about it; I said that I knew something about it, and therefore I could make some guesses very easily. So, knowing something about the text topic helped me during listening."

Related codes:

– is associated with → • bringing to mind one's knowledge or previous experiences on the subject of the listening text

Regarding constant reflection of the reasons of listening problems,

Quotations

Paul (less-successful-more-successful): "After listening, you evaluate yourself, I mean...you helped us do it, you showed how to evaluate yourself."

Taylor (less-successful-more-successful): "In the part of Reflection, we see what I would do for next listening and what I succeeded in. We wrote them on that part. So, I draw a roadmap about developing myself on it. I tell myself that 'Here was the speaker who talked too fast, that's why you did not understand it, so try to learn how to make it slower so as to understand it for next listening, at least, learn how to catch up the parts you understand from the parts you miss. I try to understand the listening texts with different accents by listening to and watching them much more. This is much more about educating me about it. I improve myself in that way, through the ways you have taught us.'"

Taylor (less-successful-more-successful): "In the part of Reflection, I wrote what I had trouble in and what I would do for next listening. Sometimes, I deduced

that I needed to focus on the accent, and so on. That's why I have started to listen to people with different accents much more. Therefore, I have improved myself on it."

Related codes:

– is associated with → • constant reflection of the reasons of listening problems

Considering Problem Solving, positive changes in correctly evaluating one's understanding while listening, and immediately reviewing one's interpretation when needed were explained.

Regarding correctly evaluating one's understanding while listening and immediately reviewing one's interpretation when needed,

Quotations

Alex (less-successful-less-successful): "Here, I mean, the word "quickly," immediately is important. I immediately adjust it, I mean, I correct my mistake... therefore, I would not misunderstand it or listen to it in the wrong way, I quickly adjust my mistake."

Related codes:

– is associated with → • correctly evaluating one's understanding while listening and immediately reviewing one's interpretation when needed

Considering Mental Translation, positive changes in making word-by-word translation consciously and on the dependence on translation were observed.

Regarding making word-by-word translation consciously,

Quotations

Paul (less-successful-more-successful): "Yes, I was making it to an extent, but now I'm not making it. I try to think in English, so I get it... at that time, I focused on it. If I didn't focus on it, it would be a problem."

Related codes:

– is associated with → • making word-by-word translation consciously

Regarding the dependence on translation,

Quotations

Taylor (less-successful-more-successful): "At the beginning, I mean, or in the middle of the process, I wrote some words that I did not understand and then tried to find their Turkish meanings or match them with something different. However, now, I guess the meanings from the sentence. Listening in English is enough for me. There is no need to translate to Turkish."

Related codes:

– is associated with → • the dependence on translation

Considering Directed Attention, positive changes in sustaining concentration while listening were observed.

Regarding sustaining concentration while listening,

Quotations

Alex (less-successful-less-successful): "Yes, difficulty in understanding... when my concentration was lost...now I have started to do, started to give up not doing it slowly, I could now concentrate on the rest of the listening text. I am telling myself that part is gone, now focus on the next sentences, and so I focus on listening to it much more."

Paul (less-successful-more-successful): "Even if I don't understand it, or I miss some parts, I am going on listening and try to connect it with the rest of the listening text."

Martin (more-successful-more-successful): "As I said before, I am trying to overcome my feature of perfectionism: 'I missed the beginning of the text, I didn't listen to the rest of it'; it is not like that anymore. Now, I am turning back to listen to it."

Alex (less-successful-less-successful): "Giving up listening is very nonsense. I need to keep understanding it."

Paul (less-successful-more-successful): "Actually, when you change even your posture while sitting and listening, it is effective, for example, when I sit up straight and try to listen, I could focus on listening much more, so it happens."

Martin (more-successful-more-successful): "I am trying to improve it. During the listening activities in the class, I have written that I need to focus on it much more; I need to collect myself, so on and so forth. I am changing my act in that way."

Taylor (less-successful-more-successful): "Here is a big change in my response. I've learned where and how to draw my attention because I recognize myself as well during that process. Therefore, later, you know, I've also found some ways to concentrate on and I can directly sustain my concentration... some ways, for example, I keep a part from the listening test in my mind, and I am asking myself that what I get about that part before, or if it is related to the sentence before it. I can find something to draw my attention to that part, so I could better focus on the text."

Related codes:

– is associated with → • sustaining concentration while listening

Considering Person Knowledge, positive changes in higher levels of difficulties in L2 listening, and feeling nervous while listening, perceiving L2 listening as a challenge, and learning how to listen were explained.

Regarding higher levels of difficulties in L2 listening,

Quotations

Taylor (less-successful-more-successful): "Umm, reading and writing are difficult; we are learning them in another language. I thought listening was the most difficult language skill for me after the listening proficiency test because I also had trouble with it while watching TV-series. However, now they all are about the same. I don't believe that one skill is easy; the other one is more difficult. They have the same levels of difficulty."

Taylor (less-successful-more-successful): "When I compare listening with other skills, also compare my current level of listening comprehension with my earlier level of listening comprehension, listening becomes easier. Now, I do not need subtitles while watching a TV-series if I don't do anything. I can go on only listening to it because I can focus on it in that way."

Related codes:

– is associated with → • higher levels of difficulties in L2 listening

Regarding feeling nervous while listening,

Quotations

Alex (less-successful-less-successful): "I was much more nervous in advance... because as you are gradually exposed to listen in English much, your reaction has been changed. So, you get relaxed, I mean, you know what you focus on and go on now."

Related codes:

– is associated with → • feeling nervous while listening

Regarding perceiving L2 listening as a challenge,

Quotations

Paul (less-successful-more-successful): "It is not a challenge as it was in the past anymore; I don't know, I answered more questions in the last listening test than the first one. According to the last test, it was better."

Related codes:

– is associated with → • feeling nervous while listening

Regarding learning how to listen,

Quotations

Taylor (less-successful-more-successful): "Initially, I thought that we would just listen, so what? Initially, I didn't know where to focus on while listening. I was listening to a sentence or a topic completely, then I guessed, but I wasn't good at it. Now, I am following it sentence by sentence or some given main points; then, if I miss some parts of the listening text, I guess the text's general idea. I mean, it is the right way to listen. I think we need to listen in that way. I felt it."

Taylor (less-successful-more-successful): "Now, at least, I know how to listen. After your study with us is done, in the future, I will focus on learning to listen by watching TV-series or, you know, by listening to songs. No vocabulary sheets were given to us anymore as it was done in high school to memorize them. Now, I need to get it from TV-series or songs. And while listening to a song, I need to focus not only on its melody but also on the song's vocabularies. Therefore, I improve myself on word structure, sentence structure, speaking. At least, because now I

know how to listen, it would be easier for me. At least, I could focus on the whole text or the part of the text, and try to understand it, match it with other texts, which is a success for me.”

Related codes:

- is associated with → • learning how to listen

A scheme of all the themes, the subdimensions, and the subtitles on positive changes from the second stimulated recall protocol is given in Figure 4.

Negative changes from the second stimulated recall protocol. After the analysis of positive changes from the second stimulated recall protocol, negative changes from the second stimulated recall in which the same four participants from the experimental group joined to discuss possible changes in their responses to the second and final MALQs and to make suggestions on having a better L2 listening comprehension skill was analyzed in terms of the subdimensions of Planning and Evaluation, Problem Solving, Mental Translation, and Directed Attention and related subtitles.

Related codes:

- is part of → • Metacognitive Awareness

Considering Planning and Evaluation, negative changes in following certain listening goals and making ready to listen through applying planning, prediction, self-check and evaluation strategies, being conscious of one's level of understanding while listening, bringing to mind one's knowledge or previous experiences on the subject of the listening text and constant reflection of the reasons of listening problems were found out.

Regarding following certain listening goals,

Quotations

Alex (less-successful-less-successful): “I see. If so, we should do it. I mean, we should have a goal to help you while listening. This goal may help you understand, deduce the meaning or guess the missing part and the rest of the text if you miss the topic. It is actually right, we should agree with it, but I don't agree with it.”

Related codes:

- is associated with → ● following certain listening goals

Regarding making ready to listen through applying planning, prediction, self-check, and evaluation strategies,

Quotations

Alex (less-successful-less-successful): “I don’t plan in my mind as I would focus on this; I would focus on that. If so, my concentration would be lost, I think. Instead of planning, I focus on just listening.”

Related codes:

- is associated with → ● making ready to listen through applying planning, prediction, self-check, and evaluation strategies

Regarding knowing one’s level of understanding while listening,

Quotations

Alex (less-successful-less-successful): “Assessing my level of comprehension at the moment of listening doesn’t come to my mind at that moment. If so, I think my concentration would be lost during listening; that’s why I changed my response in that way.”

Related codes:

- is associated with → ● knowing one’s level of understanding while listening

Regarding bringing to mind one’s knowledge or previous experiences on the subject of the listening text,

Quotations

Martin (more-successful-more-successful): “I mean, such a change happens: I had started to think that when I considered my previous listening experiences and topics, I fell under the influence of them, and my concentration headed for those experiences. So, I have changed my response as I have wanted to do something like I am learning something completely for the first time... I mean, my attention is easily lost. For example, we had an activity about Sully; I watched that movie before. Then, a movie named Flight came to my mind, which is similar in terms of topic. I got lost at that moment...”



Figure 4. Positive changes from the second stimulated recall protocol

Paul (less-successful-more-successful): "As I said before, when I focus on that part at that time, that part somehow conflicts with other parts of the listening text, and therefore focusing on it gets difficult for me."

Related codes:

– is associated with → • bringing to mind one's knowledge or previous experiences on the subject of the listening text

Regarding constant reflection of the reasons of listening problems,

Martin (more-successful-more-successful): "Here, I blame on the environment or, you know, on different factors; for example, when I listen to a song and if I do not hear the bass, I blame on my headphones because they are not quite of good quality, or also during the lessons, similarly, the classroom environment is completely not favorable, I think like that. I could not understand it for those reasons."

Related codes:

– is associated with → • constant reflection of the reasons of listening problems,

Considering Problem Solving, negative changes in correctly evaluating one's understanding while listening, and immediately reviewing one's interpretation when needed were explained.

Regarding correctly evaluating one's understanding while listening and immediately reviewing one's interpretation when needed,

Quotations

Paul (less-successful-more-successful): "It is a loss of time. I could not focus on the text at that moment if I am doing it. Because of this, I probably changed my response."

Related codes:

– is associated with → • correctly evaluating one's understanding while listening and immediately reviewing one's interpretation when needed

Considering Mental Translation, negative changes in making word-by-word translation consciously, making key words translation consciously, and in the dependence on translation were explained.

Regarding making word-by-word translation consciously,

Quotations

Martin (more-successful-more-successful): "Sometimes. For example, I give an example of Sully. There, the phrase 'brace for impact' was said. There, I needed to do word-by-word translations because it was a technique there. For this reason, I don't know, I don't get if I'm making or not, but I do word-by-word translations; it happens in some situations."

Related codes:

– is associated with → • making word-by-word translation consciously

Regarding making key words translation consciously,

Alex (less-successful-less-successful): "No, I don't. I cannot do a word-by-word translation, though. Here, what I mean is translating keywords which helps you understand the main ideas of the text, I mean those words."

Martin (more-successful-more-successful): "I do key words translation because of this reason: during the lessons or exams, we are doing something, we take notes... umm... or for example, especially in some documentaries, in some kinds of music, they are trying to tell something instead of music, yes, I am doing translation during listening to those kinds of music. Sometimes, I come across many vocabularies that I don't know, so I need to translate them."

Related codes:

– is associated with → • making key words translation consciously,

Regarding the dependence on translation,

Alex (less-successful-less-successful): "Actually, I have started to translate much more; I mean, because of trying to understand everything and integrating the parts of the text, I am doing it properly."

Martin (more-successful-more-successful): "I didn't understand the logic of translation. On the one hand, yes, I feel like I am doing translations; on the other hand, I consider I am not, it is not Turkish-English translation; it is to understand it in English directly... I feel like no translation was done. Therefore, I'm not sure if I'm doing translations or not."

Related codes:

– is associated with → • the dependence on translation

Considering Directed Attention, negative changes in going on listening over difficulties in understanding were explained.

Regarding going on listening over difficulties in understanding,

Quotations

Martin (more-successful-more-successful): "When I have difficulty in understanding because of some background music or sizzles while listening to the text or because the classroom environment is not silent enough, I generally miss the beginning of the text. I am such a perfectionist that if I miss the beginning of it, I don't care the rest of it, and I give up... It is not right, but that's just the way it is."

Related codes:

– is associated with → • going on listening over difficulties in understanding

Considering Person Knowledge, negative changes in higher levels of difficulties in L2 listening were explained.

Regarding higher levels of difficulties in L2 listening,

Quotations

Alex (less-successful-less-successful): "Yes, I have changed my response because everything in writing and speaking depends on you, but while listening, there is a situation depending on other parties as well, for example, speech rate or accent of the speaker, also other factors affect much more while listening. I understand that listening is more difficult as I listen."

Related codes:

– is associated with → • higher levels of difficulties in L2 listening

A scheme of all the themes, the subdimensions, and the subtitles on negative changes from the second stimulated recall protocol is given in Figure 5.

Suggestions on L2 listening comprehension success. After the analysis of positive and negative changes in the participants' responses to the first, mid and final MALQ, suggestions on having a better L2 listening comprehension skill in the future they made were analyzed under two subtitles: continuous L2 listening practice and communicating with others in L2, and listening to L2 speakers with different accents through authentic materials.

Related codes:

– is associated with → • Metacognitive Awareness

Regarding continuous L2 listening practice and communicating with others in L2,

Quotations

Alex (less-successful-less-successful): “Continuous listening practice and being involved in an interaction are necessary for a better L2 listening skill. Not only at school but also in my daily life, I need to do something special to improve it so that it will be effective for me ... I think that practicing L2 speaking and listening a lot, at the same time, interacting with people speaking in English make those negative factors eliminate very much. There is no other way.”

Paul (less-successful-more-successful): “Always practicing... my concentration and this accent thing... As I always said, listening practice is necessary.”

Related codes:

– is associated with → • continuous L2 listening practice and communicating with others in L2

Regarding listening to different L2 speakers with different accents through authentic materials,

Alex (less-successful-less-successful): “At first, I need to learn the correct pronunciations of many words. While listening, just because I don’t understand some words’ pronunciations, I could miss them or confuse them with other words. This is important for me. Secondly, I need to listen to something from different people. For example, an older man talking or a small kid talking, the intonation between them could be so different. Moreover, that can cause me not to understand some sentences. To the second, I need to do this, I mean, I should listen to people from different ages or, I don’t know, from different accents.”

Paul (less-successful-more-successful): “Always practicing... my concentration and this accent thing...”



Figure 5. Negative changes from the second stimulated recall protocol

Martin (more-successful-more-successful): "I think that we need to learn this by listening to natives. I mean, really, we need to listen to people speaking English as their mother tongue so that we could reach a universal level. Speaking English wherever we are in our country is a different thing, but speaking English all around the world is a different thing. That's why I think and suggest that heading towards more authentic sources is crucial."

Taylor (less-successful-more-successful): "Now, at least, I know how to listen. In the future, when your study with us is done, I will focus on learning to listen by watching TV-series or, you know, by listening to songs. No vocabulary sheets

were given to us anymore as it was done in high school to memorize them. Now, I need to get vocabularies from TV-series or songs. Furthermore, while listening to a song, I need to focus not only on its melody but also on the vocabularies of the song. Therefore, I improve myself on word structure, sentence structure, speaking. At least, because now I know how to listen, it would be easier for me. At least, I could focus on the whole text or the part of the text and try to understand it, match it with other texts; I mean; this is a success for me.”

Related codes:

– is associated with → • listening to L2 speakers with different accents and authentic materials

A scheme of suggestions on L2 listening comprehension success is presented in Figure 6.

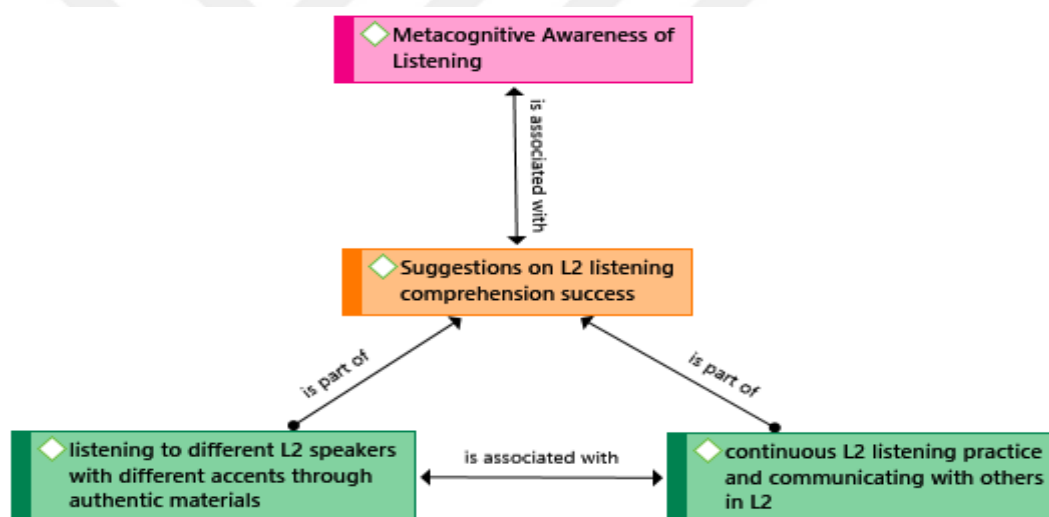


Figure 6. Suggestions on L2 listening comprehension success

Summary of Qualitative Findings

Briefly, according to the qualitative findings obtained from the stimulated recall protocols, positive and negative changes on metacognitive awareness of listening of four L2 listeners in the experimental group were explored in terms of the subdimensions of the MALQ.

Table 49 summarizes the qualitative findings under the themes and subthemes. In the next section, those findings will be discussed through interpreting the quotations and then elaborating them with quantitative findings.

Table 49

Summary of Qualitative Findings

Planning and Evaluation	Positive Changes	*following certain listening goals (1st and 2nd protocol) *making ready to listen through applying planning, prediction, self-check, and evaluation strategies (1st protocol), *being conscious of one's level of understanding while listening (2nd protocol), *bringing to mind one's knowledge or previous experiences on the subject of the listening text (2nd protocol), *constant reflection of the reasons of listening problems (2nd protocol)
	Negative Changes	*following certain listening goals (1st and 2nd protocol), *making ready to listen through applying planning, prediction, self-check, and evaluation strategies (1st and 2nd protocol), *being conscious of one's level of understanding while listening (1st and 2nd protocol), *bringing to mind one's knowledge or previous experiences on the subject of the listening text (1st and 2nd protocol), *constant reflection of the reasons of listening problems (1st and 2nd protocol)
Problem Solving	Positive Changes	*correctly evaluating one's understanding while listening and reviewing one's interpretation when needed (1st and 2nd protocol), *guessing the meaning of unknown words based on the meaning of known words (1st protocol), *regularly controlling the reasonability of one's interpretation based on background knowledge on the topic (1st protocol)
	Negative Changes	*correctly evaluating one's understanding while listening and reviewing one's interpretation when needed (1st and 2nd protocol), *regularly controlling the reasonability of one's interpretation based on background knowledge on the topic (1st protocol) *using one's background knowledge and experiences to interpret understanding while listening (1st protocol)
Mental Translation	Positive Changes	*the dependence on translation (1st and 2nd protocol) *making word-by-word translation consciously (2nd protocol)
	Negative Changes	*the dependence on translation (1st and 2nd protocol) *making key words translation consciously (1st and 2nd protocol) *making word-by-word translation consciously (2nd protocol)
Directed Attention	Positive Changes	*going on listening over difficulties in understanding (1st protocol) *sustaining concentration while listening (1st and 2nd protocol)
	Negative Changes	*going on listening over difficulties in understanding (2nd protocol) *sustaining concentration while listening (1st protocol)
Person Knowledge	Positive Changes	*higher levels of difficulties in L2 listening (2nd protocol), *feeling nervous while listening (2nd protocol), *perceiving L2 listening as a challenge (2nd protocol), *learning how to listen (2nd protocol)
	Negative Changes	- higher levels of difficulties in L2 listening (1st and 2nd protocol) - feeling nervous while listening (1st protocol)

Chapter 5

Conclusion, Discussion and Suggestions

The primary goal of the current research was to determine the effect of metacognitive instruction on EFL listeners' listening comprehension skills in Turkey. Principally, it aimed to explore to what extent metacognitive instruction influenced listening comprehension performance and metacognitive awareness of listening of L2 listeners. It also examined L2 listeners' listening comprehension performance and metacognitive awareness of listening of L2 showed any differences regarding the level of listening proficiency and preparatory school background over metacognitive instruction.

With the aim of achieving the study objectives, through the IELTS listening pretest and posttest, the MALQ, and the stimulated recall protocols, the data were quantitatively and qualitatively collected and analyzed, and the findings were explained in the previous chapter. In this part, the conclusion and discussion of the findings are presented and interpreted in parallel with previous studies in the field. Then, the implications and further research recommendations regarding the current study are given at the end of the report.

Conclusion and Discussion on Listening Comprehension Performance of L2 Listeners over Metacognitive Instruction

L2 listeners' listening comprehension performance was determined through the IELTS listening pretest and posttest after 8-week metacognitive instruction to determine the instruction's impact on their listening performance. The findings were analyzed by comparing the groups and the test scores of each group separately through the Wilcoxon signed ranks tests and the Mann-Whitney U tests. Also, less- and more-successful L2 listeners in the experimental group were compared in this sense, and the findings on the differences between them are discussed. Besides, since some of the participants studied at the preparatory school before while some of them did not study at the preparatory school before, this difference was accepted as a variable that may have an effect on the study results. Therefore, their listening comprehension performance was analyzed in terms of different preparatory school backgrounds and discussed in this chapter under the related sub-questions.

About it above, the findings revealed that the L2 listeners receiving metacognitive instruction showed a statistically significant and greater listening performance over 8-week metacognitive instruction, and this could be accounted for metacognitive instruction to a moderate extent. The result of this study could be supported by the findings of earlier studies (e.g., Vandergrift & Tafaghodtari, 2010, 2002, 2007; Alamdari & Maftoon, 2020; Bozorgian & Alamdari, 2017; Goh & Taib, 2006; Zeng, 2007, 2012; Bozorgian, 2012, 2014; Mareschal, 2007; Goh & Hu, 2013; Katrancı & Yangın, 2012; Graham & Macaro, 2008; Coşkun, 2010). These studies were focused on the listening process through a pedagogical cycle of metacognitive instruction rather than on the product of listening. They showed that metacognitive instruction facilitated L2 listening comprehension performance. Despite some differences in the implementation of those studies in terms of the length of instruction, participants, context, they were aligned with the idea that a process-based metacognitive instruction could enhance the listening skill.

Besides, the current study demonstrated that the L2 listeners who received metacognitive instruction showed a non-significant higher performance than L2 listeners who did not take place in metacognitive instruction to a moderate extent. Some previous studies demonstrated similar results. Rahimi and Katal (2013) investigated whether a process-based metacognitive instruction affected the upper-intermediate L2 listeners' listening and speaking performance. They found that the treatment group showed a non-significant higher level of listening and speaking proficiency at the end than the control group. Similarly, Çevikbaş (2016) studied metacognitive instruction's effects on English prep-school students' listening achievement in their L1, Turkish, explicitly for seven weeks. She found that the treatment group had a higher listening performance, but it was not statistically significant. Cross (2009) intended to determine the impacts of explicit strategy treatment using BBC news video texts including four metacognitive strategies together with seven cognitive and three social-affective strategies for 12 hours on EFL listeners. The results showed that strategy instruction facilitated learners' listening performance, but no significant differences were indicated in the final performances of both. The length of instruction, the study's setting and participants, and the listening proficiency could explain the lack of significance (Alamdari & Maftoon, 2020). However, it is important to state that metacognitive instruction was

explicitly given in those studies rather than a guidance practice of metacognitive strategies.

On the other hand, in this study, a process-based approach was adopted, including the implicit practice of the whole of metacognitive strategies (problem-solving, planning and evaluation, monitoring, person knowledge, directed attention, mental translation) in the listening process by authentic listening texts through instructor- and peer-guidance. L2 listeners did not listen for the comprehension questions related to the text in the study; they listened to the text by attending, guessing, checking their comprehension, finding solutions for problems, and assessing their comprehension on their own in pairs. It was assumed that they implicitly found out the knowledge of L2 listening. Nonetheless, this kind of approach did not correspond to some studies related to L2 listening strategy use and teaching in the literature (e.g., Thompson & Rubin, 1996; Cross, 2009; Coşkun, 2010; O'Malley & Chamot, 1990; Alamdari & Maftoon, 2020). Those highlighted direct strategy instruction based on explicit teaching, only one or two cognitive or metacognitive listening strategies for each lesson, and consciously practicing the selected strategies. Considering the literature, Cross (2009), Coşkun (2010), Çevikbaş (2016), and Alamdari & Maftoon (2020) adopted explicit strategy instruction and found that more outstanding but non-significant performance in listening comprehension over strategy training except for Alamdari & Maftoon (2020). They conducted 10-week metacognitive instruction in which the five metacognitive strategies were given separately by explaining its function and practicing in the classroom by encouraging the learners each week. This was a different explicit implementation from the current study and the seminal study of Vandergrift and Tafaghodtari (2010). However, Alamdari and Maftoon found that this kind of metacognitive instruction could significantly affect listening comprehension performance.

Nevertheless, Vandergrift and Tafaghodtari (2010) claimed that with this guidance in listening comprehension practice, L2 listeners could automatize the cycle of metacognitive and cognitive processes during listening, construct a mental representation on the listening text and internalize the listening process, which was advocated by Field (2008), and Graham and Macaro (2008). Also, Vandergrift (2003) asserted that guiding L2 listeners to use their metacognitive knowledge

within a pedagogical cycle through leading them to make themselves ready for, check, and assess their listening process could cause a better listening comprehension success and a self-regulated learner. Goh and Taib (2006), in their study on a process-based metacognitive instruction to young L2 listeners, found that younger listeners showed a more remarkable development in comprehending the listening nature and processes, feeling confident while engaging in listening comprehension tasks and using their strategic knowledge while dealing with challenges while listening as much adult L2 listeners. However, the researcher suggested that explicit strategy training on person, task, and strategy knowledge could be better for those young learners because of their limited but developing knowledge and learning capacity.

Regarding the level of listening proficiency. In addition to limited metacognitive knowledge capacity and implicit/explicit type of strategy training affecting L2 listeners' comprehension performance through a process-based metacognitive instruction, L2 listeners' listening proficiency could be another factor that impacted their listening success. In this paper, metacognitive instruction could significantly have a moderate level impact on enhancing less-successful L2 listeners' listening comprehension performance and a large effect on more-successful L2 listeners' listening success on the final test of the treatment group. In the literature, this was not supported by the seminal study of Vandergrift and Tafaghodtari (2010) highlighted that less-successful L2 listeners who received metacognitive instruction showed significantly greater development in their listening performance than more-successful L2 listeners in the experimental group. Besides, Mareschal (2007), in the study on a self-regulatory approach and a process-based listening treatment to adult French listeners, found that both high-skilled and low-skilled listeners increased their listening performance, but the low-skilled listeners were the ones who most benefited from the approach. Goh and Taib (2006) concluded that young L2 listeners took advantage of a process-based metacognitive instruction on listening, especially the less successful listeners at most: they learned to listen by dealing with the mental processes of listening actively rather than only answering the comprehension questions. Also, in the study of Bozorgian (2012), more-skilled listeners showed a slight increase in their final listening performance; less-skilled listeners most benefited from metacognitive instruction. This could be

because less-successful listeners did not use L1 listening process transfer to the L2 mostly and were affected more by metacognitive instruction (Goh & Taib, 2006). It may be because less-successful listener did not know the way of listening, and through metacognitive instruction, they learned how to use their memory for recall (Bozorgian, 2012); because less-skilled listeners develop and use their growing metacognitive knowledge with the help of the instructor and more-skilled listeners in the classroom (Vandergrift & Tafaghodtari, 2010; Goh, 2008).

Unlike the literature, the reasons why metacognitive instruction had a large and greater effect on more-skilled L2 listeners' listening comprehension performance might be that their prior metacognitive knowledge was higher than less-skilled L2 listeners' prior metacognitive knowledge, which was supported by Goh and Hu (2013) who found that almost 13%-15% of the difference in listening performance may be accounted for L2 listeners' metacognitive awareness. That is to say, because more-effective L2 listeners were much more aware of the listening process and more effectively used metacognitive listening strategies listening (Goh, 2002), they could show a higher listening performance on the final test than less-skilled L2 listeners even though both received metacognitive instruction.

When two groups were compared, less-successful L2 listeners who received metacognitive instruction showed a greater but non-significant listening performance than less-successful L2 listeners who did not receive metacognitive instruction to a medium extent. This result was slightly in line with Vandergrift and Tafaghodtari (2010), in which the result, however, was significant. Nonetheless, the present research showed that more-successful L2 listeners in the treatment group had a nonsignificant higher listening performance than the control group to a small extent. This wholly corresponded to the study of Vandergrift and Tafaghodtari (2010), in which more-successful L2 listeners who received metacognitive instruction demonstrated a higher listening performance than in the control group, but it was a non-significant slight difference.

Besides, considering that higher-proficient L2 listeners showed a better listening performance than the lower ones in the final test, it may be concluded that the IELTS listening test tested what it intended to test with the idea that the higher-proficient ones were expected to show a better listening performance than the lower ones. This proved that the IELTS test was valid in this study context.

Briefly, it was concluded that although less-successful L2 listeners who received metacognitive instruction showed a statistically significant and greater development on listening comprehension performance over the 8-week metacognitive instruction to a moderate extent, more-successful L2 listeners who received metacognitive instruction had a higher level of listening comprehension performance than less-successful L2 listeners who received metacognitive instruction as well as than less-successful and more-successful L2 listeners who did not receive metacognitive instruction over the process of the study. According to this, listening proficiency could cause a difference in L2 listeners' listening performance over metacognitive instruction. If L2 listeners have a higher level of listening proficiency, they could show a better listening performance over metacognitive instruction to a large extent.

Regarding different preparatory school backgrounds. In the literature, no studies in which preparatory school background was regarded as a factor to affect listening comprehension performance over metacognitive instruction. For this reason, the findings under this title were dependent on this study context.

In this study context, L2 listeners who took a preparatory school exam at the beginning of the fall semester and did not pass the proficiency exam had to have a one-year preparatory school education, including exposure to English language skills and knowledge, till getting a B2 level (intermediate level) of English. After passing the exam with a B2 level of English, they started to study at the ELT department. When L2 listeners took and passed the proficiency test for the first try, they could directly transfer to the department to study there with a B2 level of English. In the study, even if they studied or did not study at the preparatory school before, it was assumed that they had a B2 level of English, which was a must to start to major in the department. However, despite this, one-year preparatory school education before metacognitive instruction was considered a factor to affect the study results and therefore analyzed.

According to the findings, L2 listeners who did not have any preparatory school backgrounds and received metacognitive instruction showed a significant increase in the listening comprehension performance at the end of the process, which could be explained by metacognitive instruction, to a large extent. However, L2 listeners who studied at the preparatory school before and received the

instruction showed a non-significant increase on the final test to a moderate extent. Also, in the experimental group, L2 listeners who studied at the preparatory school had a higher level of listening performance than L2 listeners who did not study at the preparatory school at the end of the study to a medium extent. It could be deduced that L2 listeners with one-year preparatory school background and receiving metacognitive instruction had a higher level of listening performance.

Furthermore, L2 listeners in the experimental group who studied at the preparatory school non-significantly outnumbered the control group's L2 listeners who studied at the preparatory school in the posttest to a medium extent. On the other hand, L2 listeners in the treatment group who had no preparatory school background had a higher level of listening performance than L2 listeners in the control group who had no preparatory school background to a small extent. This could be explained by a moderate effect of metacognitive instruction after one-year preparatory school education on L2 listening performance.

Concerning these results, one-year exposure to English may cause L2 listeners to be more-proficient in L2 listening; therefore, have a higher listening performance at the end of metacognitive instruction. That could be attributed to the relationship of listening proficiency and prior metacognitive awareness, which may be supported with that higher-proficient L2 listeners were more aware of the listening process and more effectively use metacognitive strategies (Goh, 2002). Hence, they could show a higher listening performance on the final test than less-skilled L2 listeners even though both received metacognitive instruction. However, metacognitive instruction affected the listening performance of L2 listeners who did not study at the prep-school to a more considerable extent. In the same vein, it may be attributed the relationship of listening proficiency and prior metacognitive awareness, and it might be advocated by the view that because less-skilled listeners develop and use their growing metacognitive knowledge with the help of the instructor and more-skilled listeners in the classroom (Vandergrift & Tafaghodtari, 2010; Goh, 2008). Thus, L2 listeners who did not have a preparatory school background could be less proficient in L2 listening. Therefore, they could be affected to a more considerable extent by metacognitive instruction and showed a more considerable development than the listeners who had a preparatory school background.

Conclusion and Discussion on Metacognitive Awareness of Listening of L2 Listeners over Metacognitive Instruction

The findings on metacognitive awareness of listening of L2 listeners are discussed here by answering the related research questions under the subtitles of overall metacognitive awareness and of the subdimensions of MALQ, which are Problem Solving, Planning, and Evaluation, Mental Translation, Directed Attention, and Person Knowledge with integrating them with the results of qualitative findings. Firstly, the differences in overall metacognitive awareness of listening of L2 listeners are discussed.

Conclusion and discussion on the differences of overall metacognitive awareness of L2 listeners over metacognitive instruction. The study's findings showed that no statistically significant differences were determined in the overall metacognitive awareness of listening of L2 listeners who received metacognitive instruction between the beginning, the midpoint, and the endpoint of the study. However, according to the findings, the L2 listeners showed an increase from the beginning through the midpoint, but a slight decrease from the midpoint to the end of the study; however, an increase from the beginning to the end of the treatment process in their overall metacognitive awareness of listening.

Also, a statistically significant difference in overall metacognitive awareness between the groups with a medium effect size was concluded over the study, but not at the beginning and the midpoint of the study between the groups to a small extent. It showed that overall metacognitive awareness of listening of L2 listeners who received metacognitive instruction outnumbered L2 listeners who did not receive metacognitive instruction throughout the process.

It could also be concluded that metacognitive instruction statistically significantly affected and increased L2 listeners' metacognitive awareness of listening to a medium extent over the instruction. Several previous studies supported this finding (e.g., Vandergrift & Tafaghodtari, 2010; Goh & Hu, 2013; Mareschal, 2007; Zeng, 2012; Rahimi & Katal, 2012; Vandergrift, 2002, 2007; Bozorgian & Alamdari, 2017, Alamdari & Maftoon, 2020). Goh and Taib (2006) found that this kind of treatment caused students to show more remarkable development in their metacognitive awareness regarding understanding the process and nature of

listening, feeling confident while engaging in listening comprehension tasks, and using their strategic knowledge while dealing with challenges during listening. Besides, Rahimi and Katal (2012a) found that students' metacognitive awareness of listening was generally high. However, high-school learners showed a higher metacognitive awareness of listening, comparing with university-level learners, and deduced that this could be due to their level of motivation, self-efficacy, or listening ability (Vandergrift, 2003, 2005).

Some researchers investigated the metacognitive awareness of L2 listeners in different strategy instruction implementations. According to Cross (2010), dialogues emphasizing the sociocultural theory with peer-peer interaction and learner diaries could raise the metacognitive awareness of advanced-level adult Japanese EFL listeners through a pedagogical cycle. Rahimi and Katal (2012b) searched the metacognitive awareness of EFL learners with their readiness to use podcasts in their English learning process. Their metacognitive awareness of listening was found as moderate, and that the more metacognitive awareness of listening they had, the more they felt ready for the podcast-use to develop their listening ability. Goh and Hu (2013) concluded that 22% of listening performance differences were explained by metacognitive awareness of listening.

These studies above indicated that a process-based metacognitive instruction facilitated L2 listeners' overall metacognitive awareness of listening. However, Bozorgian (2014) differed from them concerning this finding of the study and concluded that overall metacognitive awareness of high-intermediate EFL listeners after metacognitive instruction for eight weeks did statistically not differ over the process. According to the researcher, this could be due to the lack of familiarity with metacognitive knowledge and strategies, exposure to and understanding practical instruction.

Regarding the level of listening proficiency. It was revealed that overall metacognitive awareness of listening of less-successful L2 listeners in the experimental group increased from the beginning through the midpoint, a slight decrease from the midpoint to the end of the study, and an increase from the beginning to the end of the treatment process but non-significantly. Unlike less-successful L2 listeners, the overall metacognitive awareness of listening of more-successful L2 listeners in the experimental group decreased from the beginning

through the midpoint, but a slight increase from the midpoint to the end of the study; however, a decrease from the beginning to the end of the process although it was not statistically significant. Besides, it was concluded that more-successful L2 listeners in the treatment group had a non-significant greater level of metacognitive awareness of listening than less-successful L2 listeners in the treatment group at the beginning and the midpoint of the process to a small extent and at the end of metacognitive instruction to a medium extent. Even though the results were not statistically significant, less-successful L2 learners in the treatment group demonstrated higher metacognitive awareness of listening than less-successful L2 listeners in the control group over the instruction to a small extent. Also, more-successful L2 learners outperformed the control group's more-successful L2 listeners in terms of metacognitive awareness at the midpoint of the study process to a small extent and at the end of the process to a medium extent.

Briefly, even the results were not statistically significant; they showed that even though less-successful L2 listeners who received metacognitive instruction pointed a greater development on metacognitive awareness of listening over the 8-week instruction process, more-successful L2 listeners who received metacognitive instruction had a higher level of overall metacognitive awareness to a moderate extent at the end of the study than the former as well as than less-successful and more-successful L2 listeners who did not receive metacognitive instruction over the process of the study. However, more-successful L2 listeners who received metacognitive instruction showed a non-significant slight decrease in their overall metacognitive awareness over the process.

According to this, listening proficiency could cause a difference in L2 listeners' metacognitive awareness of listening. If L2 listeners had a higher level of listening proficiency, they could have much more metacognitive awareness of listening together with metacognitive instruction. Some studies investigated metacognitive awareness regarding the level of proficiency in the literature and did not support this finding. In Vandergrift and Tafaghodtari's (2010) seminal research, the metacognitive awareness of less-successful L2 listeners showed more growth than more-successful listeners in the experimental group, which was contradicted with the current study. Similarly, Mareschal (2007) examined a self-regulatory approach and a process-based listening treatment to adult French listeners with

different proficiency levels for nine weeks within a pedagogical cycle, including reporting their responses and discussing them with each other. According to the results, both high-skilled and low-skilled listeners increased their metacognitive awareness, self-confidence, strategy use, and interest in listening. Nevertheless, the low-skilled listeners were the ones who most benefited from the approach, which was also conflicted with the present study findings. Also, Goh and Taib (2006) found that especially the less successful listeners took advantage of a process-based metacognitive instruction on listening at most: they learned to listen by dealing with the mental processes of listening actively rather than only answering the comprehension questions, which also differed from the current study.

Nonetheless, the frequency of metacognitive strategy use raised when the proficiency level increased (Vandergrift, 1997). Besides, Zhang and Goh (2006) highlighted that the metacognitive strategy knowledge of L2 learners positively correlated with their perceived strategy use. Goh (1998) determined the categories of strategies and tactics practiced by L2 listeners through retrospective protocols and diaries and found out that the higher proficient ones used more these strategies and tactics. In other words, in the current study, more-successful listeners' higher metacognitive awareness could be explained by their higher level of proficiency, metacognitive knowledge, and strategy use in light of these studies.

Furthermore, several studies examined metacognitive awareness regarding different listening strategies used by less- and more-successful L2 listeners. Murphy (1989) emphasized that level of listening proficiency made a difference regarding using listening strategies; for that reason, it was necessary to guide ESL listeners to experience those different strategies while focusing on the listening process by peer-supported and teacher-supported techniques rather than only answering the questions following listening to the text. According to Murphy (1985)'s study, more-successful listeners tended to actively use different strategies much more such as inferring, predicting, and dealt with top-down processes rather than bottom-up processes, while less-successful listeners preferred to focus on the text and related textual processes so much that it took a long time to elaborate what they heard. Henner Stanchina (1987) highlighted that successful listeners tended to listen and consistently monitor themselves by using their world knowledge while listening.

Regarding preparatory school background. In the present study, the overall metacognitive awareness of listening of L2 listeners who had a preparatory school background and received the treatment showed a decrease from the beginning through the midpoint, but an increase from the midpoint to the end of the study and from the beginning to the end of the treatment process, although it was not statistically significant. Similarly, in the experimental group, L2 listeners who did not study at the preparatory school in terms of their overall metacognitive awareness of listening showed an increase from the beginning through the midpoint, but a slight decrease from the midpoint to the end of the study; a slight increase from the beginning to the end of the treatment process even though it was not statistically significant. In addition, it was concluded that L2 listeners who studied at the preparatory school and received the instruction had much more metacognitive awareness of listening than L2 listeners neither who studied at the preparatory school nor who received the instruction before and after metacognitive instruction to a small extent. In contrast, the former had less metacognitive awareness of listening than the latter at the midpoint of the process to a small extent.

Also, the overall metacognitive awareness of L2 listeners having a preparatory school background between the groups demonstrated a statistically significant difference before the treatment process to a large effect size. The experimental group with preparatory school backgrounds had greater metacognitive awareness to a large extent before the instruction, which could be interpreted that L2 listeners who had a one-year preparatory school background could have some previous metacognitive knowledge and experience of listening and that the background knowledge and experience on it could affect the findings of the study. Also, before giving metacognitive instruction, testing the prior metacognitive awareness of L2 listeners and grouping them according to this could provide a more in-depth insight into the study before implementing it. On the other hand, the training group non-significantly outnumbered the control group at the midpoint of the study to a small extent and the end of the treatment process to a medium extent. Furthermore, the treatment group with no preparatory school background had a non-significantly lower level of metacognitive awareness of listening than the control group with no preparatory school background only at the beginning of the process.

In contrast, the former had a greater metacognitive awareness of listening during and at the end of the study to a small extent.

Briefly, although metacognitive instruction did not cause any statistically significant difference between L2 listeners with different preparatory school backgrounds at the end of the study, L2 listeners who studied at the preparatory school and received metacognitive instruction had a higher level of metacognitive awareness than L2 listeners who did not study at the preparatory school and received metacognitive instruction over the 8-week treatment process. This indicated that when 8-week metacognitive instruction after one-year preparatory school background and exposure to English was received, the metacognitive awareness of listening of L2 listeners could be enhanced over the process to a small extent, but non-significantly.

Furthermore, at the end of the process, L2 listeners with different preparatory school backgrounds who received metacognitive instruction had a higher level of metacognitive awareness than L2 listeners with different preparatory school backgrounds who did not receive metacognitive instruction. Above all, it could be deduced that preparatory school background could cause a difference in L2 listeners' metacognitive awareness of listening. When L2 listeners received metacognitive instruction following a one-year preparatory school background, they could have much more metacognitive awareness of listening together with metacognitive instruction.

In the literature, no studies have been found in this context related to these findings. This study could fill the gap in the literature regarding metacognitive awareness of listening. As mentioned above, one-year exposure to English may cause L2 listeners to be more-proficient in L2 listening and, therefore, have a higher listening performance at the end of metacognitive instruction. The relationship between prior language exposure and prior metacognitive awareness cannot be controlled in the study. Therefore, their growing language proficiency and metacognitive awareness could be related to these findings, showing that L2 listeners who had one-year preparatory school background had a higher level of metacognitive awareness, and metacognitive instruction accounted for it to a small extent. Goh and Hu (2013) concluded that 22% of the difference in listening performance was attributed to metacognitive awareness of listening and that there

existed a significant correlation between listening comprehension performance and metacognitive strategy use. Namely, there could be an important connection of one-year prep-school experience with increasing listening proficiency even though L2 listeners finish the preparatory school education and start to major in the department with others with a B2 level of English. Therefore, there could be seen some differences in language performance and skills between learners who studied one-year preparatory school education and who did not. Further research would be on the relationship between these factors.

Conclusion and discussion on Problem Solving over metacognitive instruction. Problem Solving was a set of metacognitive listening strategies related to making inferences and guessing what L2 listeners did not understand and checking what they made inferences. It included the strategies which were deducing the unknown meaning by the known words; deducing the unknown meaning by using the main idea of the text; interpreting the text by using L2 listeners' background knowledge or prior experiences; correctly evaluating one's understanding while listening; reviewing the incorrect interpretations immediately; controlling the reasonability of L2 listeners' interpretation based on background knowledge on the topic. It was composed of the processes of problem-solving, retrieval, and monitoring used by L2 listeners to overcome the difficulties in listening. (Vandergrift et al., 2006).

Henner Stanchina (1987) showed that successful listeners used problem-solving strategies during listening, such as applying their world knowledge to make guesses related to the text, interpreting the meaning using what they know and what they understand, assessing and reviewing their interpretations when needed. According to the researcher, successful listeners tended to be involved in listening and consistently monitoring themselves by using their world knowledge during listening. In the current research, the findings showed metacognitive awareness of Problem Solving of L2 listeners in the experimental group demonstrated an increase from the beginning through the midpoint, but a slight decrease from the midpoint to the end of the study; a slight increase from the beginning to the end of the treatment process even though all this was not statistically significant. It could be concluded that metacognitive instruction did not cause any statistically significant change in Problem Solving in the treatment group. However, while this group non-significantly

outnumbered the control group in their metacognitive awareness of Problem Solving at the beginning and the midpoint of the process to a small extent, the treatment group significantly outnumbered the control group in their metacognitive awareness of Problem Solving at the end of the research. This significant growth in Problem Solving could be explained by metacognitive instruction to a medium extent.

In the literature, there exist several studies in line with these findings (e.g., Vandergrift & Tafaghodtari, 2010; Rahimi & Katal, 2012a; Goh & Hu, 2013; Bozorgian, 2014; Altuwairesh, 2016). Vandergrift and Tafaghodtari (2010) explained this significant growth in Problem Solving by the guided implicit metacognitive instruction. Goh and Hu (2013) found a considerable relationship between listening strategy use, especially Problem Solving, and listening comprehension performance. According to Rahimi and Katal (2012a), guessing the unknown vocabularies by using their prior experience and world knowledge, assessing their inferences, comparing what they know about the topic, and revising them when needed were among the most used strategies by the listeners. Altuwairesh (2016) advocated the previous study and added that in terms of Problem Solving, the participants mostly and significantly used the strategies on guessing the unknown words with the known ones, on deducing the unknown words by using the main idea of the text, and on using prior knowledge or experience to help to understand. However, the study's result was contradicted with the study of Alamdari and Maftoon (2020). Using the strategies of Problem Solving was significantly not found in their study, and it could be attributed to the listeners' failure in making inferences.

Regarding the level of listening proficiency. Less- and more-successful L2 listeners in the treatment group demonstrated growth from the beginning through the midpoint, but a slight decrease from the midpoint to the end of the study, but an increase from the beginning to the end of the treatment process although these results were not statistically significant. More-successful L2 listeners in the treatment group had a non-significantly higher level than their counterparts in the same group at the beginning of the study to a small extent, and at the midpoint and the endpoint of the study to a medium extent.

Through the midpoint of the study, the experimental group's less-successful listeners non-significantly outnumbered in the control group to a small extent, and both groups had the same metacognitive awareness level of listening on Problem

Solving in the end to a moderate extent. Besides, more-successful L2 listeners in the treatment group non-significantly outnumbered in the control group throughout the study to a small extent.

Briefly, although both received metacognitive instruction, more-proficient listeners had a moderately greater metacognitive awareness regarding Problem Solving. Less-proficient listeners between the groups did not show any changes at the end of the study; however, more-proficient listeners showed small growth in Problem Solving between the groups. According to these results, several studies advocate that the level of listening proficiency causes a change in using the strategies of Problem Solving (e.g., Bacon, 1992; Vandergrift, 1997, 1998; Goh, 1998; Vandergrift, 2003; Chamot & Küpper, 1989; O'Malley, Chamot, & Küpper, 1989). The metacognitive strategy use increased when the level of proficiency was raised, according to Vandergrift (1997). Significantly, the use of monitoring, problem identification, and selective attention strategies differed from listeners with different listening proficiency levels. The earlier studies indicated that higher-efficient L2 listeners applied much more metacognitive strategies such as problem identification and comprehension monitoring, and integrated them with inferencing and elaboration mostly among cognitive strategies (Chamot & Küpper, 1989); monitoring, using their world knowledge, and inferencing more than their counterparts (O'Malley, Chamot, & Küpper, 1989); especially comprehension monitoring, less translation, more questioning elaboration, which was related to being flexible and open to different potentials while interpreting the meaning (Vandergrift, 2003). However, in the study of Bacon (1992), it was claimed that the level of proficiency did not cause any differences in monitoring, but effective listeners assessed their level of comprehension more realistically. Considering those studies, it is likely to deduce that a higher level of listening proficiency is positively correlated to the higher use of Problem Solving.

Validating and completing with the qualitative findings of Problem Solving. Vandergrift (1998) emphasized the importance of Problem Solving strategies in listening by indicating that monitoring the listening process was the primary metacognitive strategies leading to other metacognitive strategies such as guessing and attending selectively together with cognitive strategies such as using prior knowledge and making an inference. In other words, successful listeners

monitored their comprehension by applying successful prediction strategies following successful elaboration and inferencing practices. Less-successful listeners tended to depend more on using their background knowledge to interpret the unknown rather than on the cues such as contextual or extralinguistic cues directing themselves the meaning to compose a framework helping comprehension on their mind. Also, they spent their time mostly on bottom-up strategies as well. However, instead of lower-level processes, successful listeners deal with the meaning of using top-down processes together with metacognitive strategies.

When analyzed the stimulated recall protocols, it was found that there were some changes in the experimental group's metacognitive awareness regarding Problem Solving's strategies over the study. In the use of the strategies on *correctly evaluating one's understanding while listening and immediately reviewing one's interpretation when needed*, Paul (less-successful-more-successful) thought evaluating his understanding while listening caused him to not focus on the text at that moment. Also, he thought of it as a loss of time activity at the end of the study, which is in line with the small non-significant difference in Problem Solving between the more-skilled L2 learners of both groups. Paul (less-successful-more-successful) was initially a less-successful listener and became a more successful L2 listener at the end of the study. However, it cannot be anticipated that Paul (less-successful-more-successful) could use this strategy very well because he was in the process of acquiring and internalizing monitoring. This was supported by Alex (less-successful-less-successful), who thought it could cause him to lose track of listening if he would do it immediately while listening at the midpoint of the study. However, later Alex (less-successful-less-successful) changed his ideas positively at the end of the process, saying that he started to adjust his mistake at the moment of listening, which is in line with the non-significant decrease at the midpoint and non-significant increase at the endpoint of the research.

In terms of *guessing the meaning of unknown words based on the meaning of known words*, Paul (less-successful-more-successful) stated that he started to make guesses based on known words in the texts saying that he was aware of some words he knew and thought what others could be, which is aligned with the non-significant increase in Problem Solving at the end of the study and an earlier study (e.g., Altuwairesh, 2016).

In terms of *regularly controlling the reasonability of one's interpretation based on background knowledge on the topic*, Paul (less-successful-more-successful) stated that before listening, he thought some ideas based on the title of the text. During listening, his predictions based on his background knowledge about the topic came to his mind, and he started to check their reasonability much more. While in terms of *using one's background knowledge and experiences to interpret understanding while listening*, only Taylor (less-successful-more-successful) stated she did sometimes not have any background knowledge on the topic. This supported in the literature by Goh (2000), who highlighted that high-skilled listeners experienced problems in understanding the message meant despite understanding the words because of the lack of contextual or background knowledge; L2 listeners who received metacognitive instruction and showed an increase in listening proficiency over the study could have trouble having background knowledge and therefore making inferences while this was contradicted with the earlier studies (see Altuwaresh, 2016; Rahimi and Katal, 2012a).

In sum, the qualitative findings mostly completed and validated the results of the study's quantitative results even though some were not in line with the literature.

Regarding preparatory school background. The results highlighted that both L2 listeners who had a preparatory school background and who did not study at the preparatory school showed an increase in Problem Solving in the treatment group, from the beginning through the midpoint, but a decrease from the midpoint to the end of the study; however, no changes for L2 listeners who had a preparatory school background, but an increase for L2 listeners who had no preparatory school background although these results were not statistically significant.

Besides, L2 listeners who studied at the preparatory school and received the training had a higher-level metacognitive awareness of Problem Solving than L2 listeners who did not study at the preparatory school in the experimental group at the beginning, at the midpoint, and the endpoint of the study to a small extent, despite no statistically significant differences in them. Similarly, L2 listeners who studied at the preparatory school in the treatment group outnumbered L2 listeners who studied at the preparatory school in the control group throughout the study to a moderate extent. Besides, the experimental group's L2 listeners who did not study

at the preparatory school had a greater metacognitive awareness of Problem Solving during and after the instruction to a small extent than in the control group.

These findings were aligned with this study's results on Problem Solving regarding listening proficiency levels. As mentioned, a considerable relationship between metacognitive strategy use and listening performance was mentioned (Goh & Hu, 2013), and one-year exposure to English may increase the learner's listening performance. Therefore, the higher use of Problem Solving by the L2 listeners who studied at the preparatory school and received metacognitive instruction could be explained with this possible growing listening proficiency. Further research would be on the relationship between these factors.

Conclusion and discussion on Planning and Evaluation over metacognitive instruction. Planning and Evaluation is another metacognitive listening strategy used to make L2 listeners ready for listening and evaluate their listening comprehension performance. According to Vandergrift et al. (2006), it consisted of the strategies which were having and following certain goals in listening, preparing through applying planning, prediction, self-check and evaluation strategies, being conscious of one's level of understanding while listening, and bringing to mind one's knowledge or previous experiences on the subject of the text and constant reflection of the reasons of listening difficulties.

The findings revealed that L2 listeners in the treatment group showed an increase in Planning and Evaluation use from the beginning through the midpoint, but a slight decrease from the midpoint to the end of the study and from the beginning to the end of the treatment process despite no statistically significant differences. Similarly, no statistically significant difference was found between the groups before, during, and after the study. However, through the midpoint of the study, the control group outnumbered the experimental group, and both groups' levels of metacognitive awareness on Planning and Evaluation were the same at the endpoint of the study to a small extent.

In the literature, Vandergrift and Tafaghodtari (2010) and Rahimi and Katal (2012a) found similarly no significant increase in Planning and Evaluation. In contrast, Bozorgian (2014) and Bozorgian and Alamdari (2017) concluded a

significant increase in Planning and Evaluation strategy use and indicated that metacognitive instruction helped L2 listeners plan and assess their understanding.

Concerning the results of the study, metacognitive instruction did not show any significant change in Planning and Evaluation; even the treatment group showed a non-significant slight decrease throughout the study. In the literature, previous studies indicated that planning and evaluation were a facilitator for language learning (Wenden, 1998) and L2 listening (Vandergrift, 2003). According to this, although it was expected that metacognitive instruction could increase the metacognitive awareness of listeners, the findings on Planning and Evaluation were contradicted with the literature, which could explain the low-reliability values of the construct of Planning and Evaluation in the study. Schmitt (1994) stated that, in some circumstances, finding out low levels of reliability values could be useful. This may show an undesirable interaction in Planning and Evaluation in terms of its first and second implementation for this study. As it was seen, its low values could be explained meaningfully due to the differences in Planning and Evaluation (Peterson, 1994).

Regarding the level of listening proficiency. Although the findings were not statistically significant, less-successful L2 listeners in the experimental group demonstrated an increase from the beginning through the midpoint, but a decrease from the midpoint to the end of the study and from the beginning to the end of the treatment process. Also, more-successful L2 listeners in the treatment group pointed to a decrease from the beginning through the midpoint but increased from the midpoint to the end of the study; no changes from the beginning to the end of the treatment on Planning and Evaluation. In addition, less-successful L2 listeners in the experimental group non-significantly outnumbered at the beginning and the midpoint of the study to a small extent. However, they did not outnumber more-successful L2 listeners in the treatment group regarding Planning and Evaluation at the end of the study to a small extent.

Less-successful L2 listeners in the treatment group did not outnumber in the control group at the end of the process to a small extent, whereas more-successful L2 learners who received the instruction non-significantly outnumbered in the control group at the endpoint of the process to a small extent, but did not outnumber them

at the midpoint of the process to a small extent in metacognitive awareness of listening on Planning and Evaluation.

Briefly, the level of listening proficiency caused no consistent significant change throughout the study in Planning and Evaluation. In the literature, similarly, not many studies were found to be in line with this finding. Murphy (1985) and Chamot and Küpper (1989) asserted that more-efficient listeners had used more planning, predicting, and evaluation strategies. Goh (1998) asserted that more-effective listeners practiced comprehension evaluation more. However, these supported the study's finding that more-successful L2 listeners in the experimental group regarding Planning and Evaluation non-significantly outnumbered at the end of the study to a small extent.

Nonetheless, the impact of metacognitive instruction was relatively low on Planning and Evaluation in terms of the level of listening proficiency. This could be attributed that the L2 listeners, both less-successful and more-successful L2 listeners, were not familiar with the planning and especially evaluation strategy use. After listening, evaluating their listening performance and making reflections about it are not common in language learning classroom (Goh, 2010) because the listening was based on the product, not the process of listening; after they answered the questions, they did not talk over their performance and next listening activities. For this reason, the 8-week metacognitive instruction could not provide them with the growth in Planning and Evaluation usage in this study context.

Validating and completing with the qualitative findings of Planning and Evaluation. In terms of *following certain goals*, throughout the study, the participants showed positive and negative changes. At the beginning of the research, Taylor (less-successful-more-successful) stated she did not have a goal during listening and listened pointlessly at the beginning; she started to focus more on listening and understanding at the midpoint of the study, which is not line with the non-significant decrease in Planning and Evaluation at the midpoint of the study while Alex (less-successful-less-successful) and Paul (less-successful-more-successful) claimed that having a goal and focusing on that goal while listening could cause to lose the track of listening and their concentration, and for this reason they did not set and follow any listening goals during listening at the midpoint and at the end of the study, Alex (less-successful-less-successful) noticed that following

certain listening goals was necessary, but he still did not do it, which is line with the non-significant decrease in Planning and Evaluation for the experimental group, especially for less-successful listeners at the midpoint of the study. However, Martin (more-successful-more-successful) indicated at the end of the research that type of text and type of input affected to have a goal and follow that set goal and illustrated it by saying while listening to some music or songs, there was no need for a listening goal, but while listening to an academic text or a podcast depending on its topic, having a goal and following this goal was necessary.

In terms of *making ready to listen through applying planning, prediction, self-check, and evaluation strategies*, Alex (less-successful-less-successful) stated that at the beginning of the study, he just listened to texts inattentively, through the midpoint of the study, he prepared himself making a plan on focusing on keywords. At the end of the study, Alex (less-successful-less-successful), as he thought before, considered when he made a listening plan, he would lose his concentration; that is why he started to listen to the texts aimlessly, which is in line with the non-significant decrease in Planning and Evaluation throughout the study for less-successful listeners. Taylor (less-successful-more-successful) indicated that she would know where to concentrate during listening when she made a plan and herself ready to listen. At the end of the study, Taylor (less-successful-more-successful) became aware of herself during listening: she stated that she knew how to listen, focus, and know the listening process, and therefore she prepared herself to listen to the text. Paul (less-successful-more-successful) realized that his predictions on the text's title came to his mind and started to think much more on predictions and write down his guesses before listening to in My Prediction part. Similarly, Paul (less-successful-more-successful) made himself ready to listen by checking the questions related to the text before and therefore knew where to focus, which could be supported by the increase at the midpoint and then no change in Planning and Evaluation at the end of the study for more-successful L2 listeners.

In terms of *being conscious of one's level of understanding while listening*, Paul (less-successful-more-successful) interpreted his negative change in assessing his understanding level while listening by his increased self-confidence on L2 listening. Martin (more-successful-more-successful), at the end of the research, asserted such assessment was necessary to find out the source of the

problems that emerged out during listening. Taylor (less-successful-more-successful) stated that she became conscious of her understanding level when she saw the text's transcript and the misunderstood words because of pronunciation differences. Also, she had the same experience when she compared her notes about before and during listening on her listening diary with each other and assessed her understanding level. On the other hand, Alex (less-successful-less-successful) negatively changed his ideas and pointed that he did not assess his level of understanding during listening because of fear of losing concentration and missing out on some parts of listening texts. Those findings were associated with the study's quantitative findings, which stated the non-significant decrease along with the study for less-successful L2 listeners while the non-significant increase at the midpoint and the same level to the end of the study for more-skilled L2 listeners.

In terms of *bringing to mind one's knowledge or previous experiences on the subject of the listening text*, at the midpoint of the study, according to Paul (less-successful-more-successful), it was not useful because his previous experiences and knowledge on the topic were different. However, it helped Taylor (less-successful-more-successful) make guesses on the listening text topic based on her previous experiences at the end of the study. On the other hand, Martin (more-successful-more-successful) and Paul (less-successful-more-successful) claimed that the subjects were different from their background knowledge on the topic, their concentration headed for those experiences or the topic conflicted with each other, and therefore their attention was lost. According to them, not having previous knowledge and experiences similar to the text's topic could cause them to lose their concentration during listening. Goh (2000) claimed that high-skilled listeners also experienced problems in understanding the message meant despite understanding the words because of the lack of contextual or background information, which is supported by the present report.

In terms of *constant reflection of the reasons of listening problems*, at the beginning of the study, Taylor (less-successful-more-successful) stated that she reflected her ideas on listening problems on the Reflection part of her listening diary and tried different listening methods for those listening problems. However, through the midpoint of the study, she did not make constant reflections on those problems because she found her method to deal with them. On the other hand, at the end of

the study, Paul (less-successful-more-successful) stated that the instructor (the researcher) showed how to evaluate themselves after listening; it meant that metacognitive instruction helped the participants to be aware of this reflection process and evaluate themselves after listening. Similarly, Taylor (less-successful-more-successful) stated that evaluating themselves after listening and writing their ideas on the Reflection part made herself aware of what they had trouble in and would do in an unlike way during the next listening. For example, she started to listen to different people with different accents because she realized that she needed to develop herself. However, at the end of the study, Martin (more-successful-more-successful) changed his ideas negatively, yet added that while he tried to find out the source of listening problems, he blamed on classroom environment, the quality of audio text instead of himself.

Regarding preparatory school background. The findings reported that in the treatment group, both L2 listeners who had preparatory school background and who had no preparatory school background showed a decrease from the beginning through the midpoint, but an increase from the midpoint to the end of the study; and when analyzed from the beginning to the end of the instruction process, the former showed a slight increase while the latter showed a decrease although these results were not statistically significant. Also, it was found that L2 listeners who did not study at the preparatory school in the experimental group had a significantly growing metacognitive awareness of Planning and Evaluation than L2 listeners who studied at the preparatory school in the treatment group at the midpoint to a large extent while before and after the treatment process to a small extent but non-significantly.

It was concluded that the experimental group's L2 listeners who studied at the preparatory school had a non-significant higher level of Planning and Evaluation use at the end of the study to a small extent than in the control group, while the control group had a significantly greater level of Planning and Evaluation usage at the midpoint of the study to a large extent. It was also revealed that the experimental group's L2 listeners who did not study at the preparatory school non-significantly outnumbered L2 listeners who did not study at the preparatory school in the control group at the beginning, at the midpoint, and the endpoint of the study to a small extent.

Conclusion and discussion on Mental Translation over metacognitive instruction. Mental Translation is the other metacognitive strategy that is expected to be used less when listeners become more proficient in listening. It included the strategies doing word-by-word translation, key-word translation, and the dependence on translation, which need to be overcome to be more-successful L2 listeners (Vandergrift et al., 2006).

The results pointed that regarding Mental Translation, the experimental group showed no changes from the beginning through the midpoint, but a non-significantly slight decrease from the midpoint to the end of the study and from the beginning to the end of the treatment process. Besides, the control group non-significantly outnumbered the experimental group on Mental Translation at the beginning and the end of the process to a small extent; the treatment group outnumbered the control group through the midpoint of the research.

Accordingly, the treatment group demonstrated a non-significantly decrease in Mental Translation to a small extent. In other words, metacognitive instruction could slightly cause to decrease in the use of translation in L2 listening. In contrast, Bozorgian and Alamdari (2017) and Vandergrift and Tafaghodtari (2010) found a significant increase in using the translation by L2 listeners who received metacognitive instruction, a paradoxical result. On the other hand, Altuwairesh (2016) found that Mental Translation was the least used strategy by L2 listeners, which was aligned with the current study because it showed a non-significant decrease in use.

Regarding the level of listening proficiency. The findings showed that both less-successful L2 listeners in the treatment group showed a non-significant increase while more-successful L2 listeners in the treatment group reported a decrease from the beginning through the midpoint and to the end of the treatment process. More-successful L2 listeners in the treatment group had a non-significant higher level of metacognitive awareness on Mental Translation at the beginning and the midpoint of the study to a small extent, but a lower-level strategy use of Mental Translation than less-successful L2 listeners in the treatment group over the research to a small extent.

Also, less-successful L2 listeners in the control group outnumbered in the treatment group at the beginning of the process to a small extent; however, at the midpoint and the end of the process, less-successful L2 listeners in the treatment group outnumbered in the control group to a small extent on Mental Translation. On the other hand, more-successful L2 learners in the experimental group outnumbered in the control group at the beginning and the midpoint of the study to a small extent. However, in the end, the control group's more-successful L2 listeners outnumbered their counterparts in the experimental group on Mental Translation to a small extent. However, these results presented no statistically significant differences between each other.

The literature indicated that less-successful L2 listeners use Mental Translation due to the lack of listening competence in recognizing the words and vocabulary knowledge (Vandergrift, 2003; Goh, 1998; Alamdari & Maftoon, 2020). Vandergrift (2003) asserted that more-successful L2 listeners mostly practiced metacognitive strategies, especially comprehension monitoring, more questioning elaboration, which was related to being flexible and open to different potentials while interpreting the meaning, and less translation. Less-successful listeners mostly used translation as a bottom-up process, which caused them not to keep in mind what they heard and create their mental representations because of the limited memory capacity and, therefore, led not to understand it. According to the literature, it was expected that less-successful L2 listeners were supposed to use more translation, while more-successful L2 listeners used not much translation. The literature supports this finding. More-successful listeners showed a slight decrease in using translation at the end of the study, while less-successful listeners who received metacognitive instruction increased translation use; even they used much more translation than less-successful L2 listeners who did not receive metacognitive instruction. That was supported by Bozorgian (2014)'s study, which claimed that metacognitive instruction did not facilitate to overcome the use of mental translation, and by the studies of Vandergrift and Tafaghodtari (2010) and Bozorgian and Alamdari (2017), in which the significant increase in the use of translation by L2 listeners who received metacognitive instruction was considered as a paradoxical result.

Validating and completing with the qualitative findings of Mental Translation. In terms of *the dependence on translation*, Taylor (less-successful-more-successful) stated that she did not translate anymore at the midpoint of the study. Again, Taylor (less-successful- more-successful) had a positive change over the study and indicated that she used to try to find Turkish meanings when she did not understand some parts. However, later she made guesses on them and had no need to translate and that listening in L2 was enough to understand, which is aligned with the present research and some earlier studies regarding more-successful L2 listeners in the experimental group (see Vandergrift, 2003; Goh, 1998; Alamdari & Maftoon, 2020). On the other hand, at the end of the study, Alex (less-successful-less-successful) started to do it much more because of the desire to understand everything and integrate the parts of the text, which was supported by the present study in terms of less-successful L2 listeners and also Goh (2000) emphasizing that less-skilled listeners used more because they could not keep the input on the working memory and recall the necessary knowledge, and therefore forget the understood information. Also, Martin (more-successful-more-successful) was completely not aware of if he was doing it or not.

In terms of *making key words translation*, at the midpoint of the study, Alex (less-successful-less-successful) stated that he chose and translated some words to deduce the text's central idea. At the end of the study, he started to do it much more and negatively changed his ideas, which was aligned with the current study and the previous studies stating that when the level of proficiency decreased, the use of translation increased. Martin (more-successful-more-successful) indicated that sometimes he needed to translate some keywords to understand them while listening to texts such as documentaries or songs. The type of input and type of text affected his situation. Also, sometimes he met some unknown words and needed to translate them as required. In terms of *making word-by-word translation*, Paul (less-successful-more-successful) stated that he did not do word-by-word translations anymore and tried to think in English when focusing on listening. Martin (more-successful-more-successful) asserted that sometimes word-by-word translation was necessary to understand the important part of the text; however, still, he did not realize if he was doing translations or not during listening. In the current study, considering more-successful L2 listeners who received metacognitive instruction,

the use of key-word or word-by-word translation could sometimes be necessary based on the text type and type of input for those who needed to overcome to become more component in L2 listening.

Regarding preparatory school background. Regarding Mental Translation, in the experimental group, the metacognitive awareness of L2 listeners who had a preparatory school background showed a non-significant increase from the beginning through the midpoint and did not change from the midpoint to the end of the treatment process and a slight increase from the beginning to the end of the treatment process. In contrast, L2 listeners who did not study at the preparatory school but received the treatment reported a non-significant decrease in Mental Translation from the beginning through the midpoint and to the end of the treatment. In the treatment group, L2 listeners who studied at the preparatory school had a non-significantly higher level of metacognitive awareness of listening on Mental Translation than L2 listeners who did not study at the preparatory school at the beginning, at the midpoint, and the end of the study to a small extent.

Regarding Mental Translation, the experimental group's L2 listeners who studied at the preparatory school non-significantly outnumbered L2 listeners who studied at the preparatory school in the control group throughout the research to a small extent, whereas the control group's L2 listeners who did not study at the preparatory school non-significantly outnumbered in the experimental group at the beginning, at the midpoint and the endpoint of the study to a small extent.

Briefly, in the experimental group, L2 listeners who did not have one-year preparatory school background showed a lower use of translation than L2 listeners who studied at the prep school in the experimental group as well as L2 listeners who did not study at the preparatory school in the control group. L2 listeners who had a one-year preparatory school background showed higher use of translation than their counterparts in the experimental and control groups. This could be contrasted with the idea advocated above that an increasing listening proficiency could be related to having one-year preparatory school background, which meant L2 listeners who had a preparatory school background could have a growing level of listening proficiency. That could be why they showed an increase in their overall listening performance and metacognitive awareness; however, it could be paradoxical in this context. Therefore, the higher use of mental translation by L2 listeners with one-

year of preparatory school background could be attributed to their growing listening competence and the need to use L1 sometimes to become aware of the process and facilitate their listening comprehension (Alamdari & Maftoon, 2020).

Conclusion and discussion on Directed Attention over metacognitive instruction. Directed Attention is the fourth metacognitive strategy for focusing and staying on the listening text. It included going on listening and giving up over difficulties in understanding and concentrating harder on the text over difficulties, sustaining concentration while listening, and trying to recover one's concentration when lost (Vandergrift et al., 2006). According to Ross (2002), successful listeners could sustain their concentration while controlling the other processes and dealing with the problem during listening.

The results highlighted that the experimental group demonstrated no changes in Directed Attention throughout the study, although they were not statistically significant. However, the experimental group showed a significantly higher level of Directed Attention only after the instruction with a moderate effect size. At the midpoint, the treatment group outnumbered the control group on Directed Attention to a small extent but non-significantly. Although the experimental group showed no change in Directed Attention throughout the study, the significant increase could only be explained by the decrease in metacognitive awareness on Directed Attention at the end by the control group. Therefore, it could not be attributed to metacognitive instruction, which had no impact on the treatment group regarding Directed Attention. In the literature, most studies supported this finding of the study (e.g., Vandergrift & Tafaghodtari, 2010; Rahimi & Katal, 2012a; Bozorgian & Alamdari, 2017; Bozorgian; 2014), which emphasized that significant differences were not determined in Directed Attention strategy use of L2 listeners who received metacognitive awareness. Goh and Hu (2013) found an important correlation between metacognitive strategy use, primarily Directed Attention and Problem Solving, and listening comprehension performance. Altuwairesh (2016) found that Directed Attention was one of the most used metacognitive strategies by L2 listeners over metacognitive instruction. It was concluded that they mostly used the strategies on concentration harder on the text over difficulties in understanding, secondly going on listening when the concentration lost, and thirdly on sustaining concentration

while they did not commonly use the strategy on giving up when they lost track of listening, which were supported by the qualitative results of the study though.

Regarding the level of listening proficiency. The findings showed that while more-successful L2 listeners in the experimental group showed no changes from the beginning through the midpoint to the end of the study, less-successful L2 listeners in the experimental group showed non-significantly no changes from the beginning through the midpoint, but a slight decrease from the midpoint to the end of the study and from the beginning to the end of the treatment process. No changes were found between less- and more-successful L2 learners who received the instruction before and during the study to a small extent; however, the latter had a non-significantly higher level than the former at the endpoint of the study to a small extent.

Through the midpoint of the study, less-successful L2 listeners in the treatment group had no significant increase in Directed Attention than the control group to a small extent, and less-successful L2 listeners of both groups had the same level of Directed Attention at the endpoint of the study to a small extent. More-successful L2 learners who received the instruction non-significantly outnumbered in the control group at the midpoint to a small extent. On the contrary, it was found that more-successful L2 listeners who received the instruction significantly outnumbered in the control group regarding Directed Attention at the end of the process. Similarly, the large difference could not be explained by metacognitive instruction because more-successful L2 listeners who received the instruction showed no change over the instruction; the control group showed a large decrease in their awareness of Directed Attention at the end of the study.

In the literature, Chamot and Küpper (1989) found greater differences in selective attention and evaluation between less- and more-efficient high-school L2 listeners. Goh (1998) supported the previous study emphasizing that higher-proficient listeners used “selective attention, directed attention, comprehension monitoring, real-time assessment of input, comprehension evaluation” while the less-proficient ones used these first three metacognitive strategies (p.13). Vandergrift (1997) showed that monitoring, problem identification, and selective attention differed from effective and less-effective listeners. These studies did not support the current study because more-successful L2 listeners did not change

using directed attention over metacognitive instruction. However, O'Malley, Chamot, and Küpper (1989) asserted that more-efficient ones did not mostly prefer directed and selected attention in their study's quantitative findings, which is in line with the study's quantitative findings.

Validating and completing with the qualitative findings of Directed Attention. In terms of *going on listening over difficulties in understanding*, at the beginning of the study, Paul (less-successful-more-successful) had given up listening when he had difficulty understanding because of different accents, at the midpoint of the study; he started to keep going on listening despite those difficulties. However, Martin (more-successful-more-successful) changed his ideas negatively at the end of the study because when he missed the beginning of the text due to the classroom environment, background music, sizzles, he lost his concentration and gave up listening. According to this, such difficulties because of the environmental factors caused to loss of L2 listeners' attention during listening.

In terms of *sustaining concentration while listening*, at the beginning of the study, Paul (less-successful-more-successful) used not to sustain his concentration when he lost track of listening; at the midpoint of the study, even if it was still difficult for him, he was trying to recover his concentration while listening. Furthermore, he indicated that at the end of the study, even though he missed some parts of the text while listening, he preferred to connect those parts with the rest of the text to understand it wholly. Also, he stated that changing the posture like sitting up straight was useful for concentration in listening. Besides, at the midpoint of the study, Alex (less-successful-less-successful) changed his ideas negatively because he thought that if he had a goal in his mind, he would lose his focus while listening, whereas at the end of the study, he started to sustain listening even if a part of the text was missed and thought that giving up was unnecessary, he needed to keep going for comprehending the text. As for Martin (more-successful-more-successful), at the midpoint of the study since he said that he could be bored very quickly and gave up while listening, that he did not like being unsuccessful because of his personality and gave up listening when he did not understand a part and lost his concentration, and that having an opportunity of second listening and knowing it was a chance to recover the missing parts for him; nonetheless, at the end of the study, he positively changed his ideas as saying that he would not give up listening anymore and tried

to improve it. Similarly, Taylor (less-successful-more-successful) did not sustain listening because of accent or personal problems at the moment of listening through the midpoint, but she stated that she found out how to do it, where and how to draw her attention, and some methods to concentrate on listening asking herself the parts she understood now and the parts she did not understand, which is about Problem Solving, assessing herself regularly while listening at the end of the research.

Those findings are not in line with the quantitative findings for more-successful and less-successful L2 learners who received the instruction because the former showed no change over the study, and the latter showed a slight decrease at the endpoint of the research. However, the L2 listeners above showed a growing metacognitive awareness of Directed Attention. Nonetheless, the quantitative findings were non-significant with no or small effect. Thus, the difference could not be revealed with the statistics, but the participants' perceptions. As in Planning and Evaluation, according to Schmitt (1994), in some circumstances, finding out low levels of reliability values could be useful. This may show an undesirable interaction in Directed Attention in terms of its first and second implementation for this study. As it was seen, as a result of the study, its low values could be explained meaningfully due to the variances in Directed Attention (Peterson, 1994). Nonetheless, thanks to the mixed-methods research design, the perceptions of L2 listeners who received metacognitive instruction could be shown, and when analyzed, the four's perceptions of Directed Attention are consistent with each other.

O'Malley, Chamot, and Küpper (1989) asserted that more-successful learners tended to prepare themselves for what to focus on during listening, stay focused, keep going on listening, and immediately recover concentration when lost. Besides, they preferred to interact with the text through much more top-down or utilization processes such as using their world or personal knowledge, asking themselves questions about the text they listened to or using contextual information to help their comprehension. Less-effective listeners tended to give up listening when they met unknown words or phrases, interact with the text using the lower processes, such as recognizing vocabularies, and did not much integrate what they knew about the text with what they understood. This is associated with the current study. Goh (1997) found five main factors influencing listening comprehension:

listening task, listening text, listeners, environment, and speakers. The top five factors concluded were speech rate, input type, speaker accent, vocabulary, and background knowledge. Goh (1999) also found that listeners' point of interest for listening, their psychological and physical situation including anxiety, tiredness, their knowledge about the topic of the text, their ability to stay focused, visual-supported listening activities, the speaking ability of the speaker were among other factors affecting listening comprehension. According to Goh (1999), for high-skilled listeners, listening was an interactive process in which the meaning was constructed together with listeners and speakers while listening was mainly regarded as a text-oriented activity and textual factors such as vocabulary, rate of speech, and input type rather than environment and speaker, which were considered influencing listening comprehension for the lower-skilled listener. Goh (2000) found that low-skilled listeners had difficulty in understanding the rest of the text as a result of focusing too much on the input they just heard because of the lack of vocabulary knowledge, while the higher ones did not experience this problem because they kept going on listening using the strategy of directed attention.

Considering the problems L2 listeners could meet in the process of listening, L2 listeners in the current study, in the beginning, especially less-successful L2 listeners in the experimental group, tended to give up listening when they lost track of their concentration because of the speaker accent, background noises, personal situations. However, at the end of the study, they tended to try not to give up, keep listening, and recover their concentration even when lost. As a result, it could be deduced that metacognitive instruction could slightly increase metacognitive awareness in Directed Attention of the experimental group even though the quantitative results did not support it.

Regarding preparatory school background. The findings showed that L2 listeners who had a preparatory school background in the treatment group indicated a non-significant decrease from the beginning through the midpoint, but an increase from the midpoint to the end of the study, a slight decrease from the beginning to the end of the treatment process. L2 listeners who did not have preparatory school backgrounds in the experimental group showed no changes from the beginning through the midpoint, but non-significantly increased from the midpoint to the end of the study and from the beginning to the end of the treatment process.

It also demonstrated that in the experimental group, L2 listeners who did not study at the preparatory school had a significantly greater metacognitive awareness of listening on Directed Attention at the midpoint, which could be explained by different preparatory school backgrounds on metacognitive awareness of listening on Directed Attention to a large extent. Also, they had a non-significantly higher level of metacognitive awareness on Directed Attention both at the beginning and the end of the treatment process to a medium extent than those who studied at the prep-school. According to the results, the experimental group's L2 listeners who studied at the preparatory school outnumbered L2 listeners who studied at the preparatory school but did not receive the instruction over the study to a medium extent but non-significantly. Besides, the experimental group's L2 listeners who did not study at the preparatory school outnumbered L2 listeners who did not study at the preparatory school and receive the instruction on Directed Attention to a moderate extent over the research. No changes between them in their metacognitive awareness on Directed Attention at the beginning and the midpoint of the study to a small extent were seen, although they were not statistically significant.

Briefly, L2 listeners who did not study at the preparatory school showed non-significant growth throughout the research and a higher level of Directed Attention use than L2 listeners who studied at the preparatory school to a moderate extent in both groups. However, the experimental group did not show any change over the study, and the findings were not significant and had no considerable effect. That is why these results could be attributed to the L2 listeners' concentration efforts, which may show differences according to the text type, psychological situation, accent, background noises (Goh, 2000).

Conclusion and discussion on Person Knowledge over metacognitive instruction. Person Knowledge is the fifth metacognitive strategy related to perceptions of L2 listeners on listening difficulties and their self-efficacy beliefs in L2 listening, which is about their capability to learn and achieve a specific goal in language learning (Wenden, 1998). It consisted of listening perceived as a challenge and difficulty level of listening regarding other skills and feeling anxious during listening (Vandergrift et al., 2006). Person knowledge makes learners aware of themselves as a listener with the listening process (Vandergrift, 2002), and a

close association is available between metacognitive knowledge and self-efficacy beliefs with each other (Vandergrift, 2005).

Along with this, the findings showed that the experimental group's Person Knowledge showed a significant decrease from the beginning, through the midpoint, and to the end of metacognitive instruction to a small extent. There were no statistically significant variances between any two of the three-time points to a small extent. Also, the treatment group outnumbered the control one in Person Knowledge throughout the research to a small extent and non-significantly.

The findings are supported by previous research (e.g., Vandergrift & Tafaghodtari, 2010; Rahimi & Katal, 2012a, 2012b; Bozorgian; 2014; Altuwaresh, 2016; Bozorgian & Alamdari, 2017) in which metacognitive instruction had no effect on significant growth in Person Knowledge to a small extent. On the contrary, L2 listeners who received metacognitive instruction showed a significant decrease from the beginning to the end of the study. In other words, although it was expected that listening instruction would develop the proficiency of listening and self-efficacy of the students, and therefore their Person Knowledge (Graham & Macaro, 2008; Vandergrift, 2005), the significant decrease over metacognitive instruction is paradoxical in the literature. This could be associated with the lower-level of self-efficacy beliefs of the L2 listeners, their affective status, their perceptions of listening as a challenge or a reason for anxiety, which are the factors changing depended on the study context.

Regarding the level of proficiency. The findings revealed that less-successful L2 listeners in the experimental group pointed to development from the beginning through the midpoint, but no changes from the midpoint to the end of the study, and a slight increase from the beginning to the end of the treatment process despite not being statistically significant. More-successful L2 listeners who received the instruction demonstrated non-significant growth from the beginning through the midpoint, but a decrease from the midpoint to the end of the study, and a decrease from the beginning to the end of the treatment process.

It was found that more-successful L2 listeners had a non-significantly higher level of Person Knowledge use than the less-successful listeners who received the instruction at the end of the process to a medium extent. In comparison, the former

outnumbered the latter significantly at the midpoint of the study, which could be explained by different listening proficiency levels in Person Knowledge use to a large extent. Less-successful L2 listeners in the control group outnumbered less-successful L2 listeners in the treatment group at the beginning of the process to a small extent, but both had the same level of metacognitive awareness in Person Knowledge at the midpoint of the study to a medium extent and the endpoint of the study to a small extent. Besides, more-successful L2 listeners in the experimental group outnumbered more-successful L2 listeners in the control group at the beginning, at the midpoint, and at the end of the process to a small extent even though they were not statistically significant.

In the literature, no specific findings on Person Knowledge regarding the level of proficiency with metacognitive instruction were available. However, the earlier studies highlighted the positive relationship between metacognitive knowledge and listening comprehension performance (Goh & Hu, 2013); between the metacognitive strategy knowledge of L2 listeners and their perceived metacognitive strategy use (Zhang & Goh, 2006); and between the frequency of using metacognitive strategies with the increasing level of listening proficiency increased (Vandergrift, 1997). Concerning this, in the current study, more-successful L2 listeners had a higher level of Person Knowledge than less-successful L2 listeners over metacognitive instruction, which was explained by metacognitive instruction to a medium extent and supported by the studies in the literature. In other words, it could be interpreted that when listeners have a higher level of listening proficiency, they use more Person Knowledge strategies.

Validating and completing with the qualitative findings of Person Knowledge. In terms of *higher levels of difficulties in L2 listening*, according to Taylor (less-successful-more-successful), at the beginning of the study, she was not aware of the difference in the level of difficulty between the four skills, and they had the same level of difficulty for her, but at the midpoint of the study, listening was more difficult than other skills for her. Also, due to the fact that she did not get every detail in listening and the type of text was a factor to affect the difficulty level of listening, she changed her ideas negatively. Nevertheless, at the end of the study, she thought that even though listening was still difficult for her, the four skills had the same levels of difficulty at that moment and that she had no need for subtitles

to understand the listening text. On the other hand, at the end of the study, Alex (less-successful-less-successful) changed his ideas negatively and considered that because while in writing and reading, everything depended on you, in listening, there were other parties he could not interfere with such as speech rate, accent, listening was more difficult than other skills. These are supported by the quantitative findings of the current research in which the experimental group showed a decrease in Person Knowledge throughout the study.

In terms of *feeling nervous while listening*, Martin (more-successful-more-successful) stated that since there were no chances to ask a question while listening to an audio text on the tape, unlike face-to-face communication and sometimes because of accent, he felt nervous during listening. Also, Taylor (less-successful-more-successful) said that she was afraid of making incorrect interpretations and misunderstandings while listening, and she thought that she was not good at listening at the midpoint of the study that she would never be good at it; she changed her ideas negatively. However, she stated that she learned how to listen and focus on listening and positively changed her ideas at the end of the study. In addition, Alex (less-successful-less-successful), at first, was much more nervous, but, at the end of the study, he was more relaxed while listening and knew what to focus on because he was much more exposed to listening during the process.

In terms of *perceiving L2 listening as a challenge*, only Paul (less-successful-more-successful) showed a change positively and stated that it was not a challenge anymore because he realized that he answered many more questions of the second listening test.

On *learning how to listen*, only Taylor (less-successful-more-successful) expressed her ideas at the end of the study. She said she found out how to listen and improved her own guessing methods, concentrating, and understanding while listening. According to this, she was aware of the process of listening, which was a success for her.

Regarding the changes in Person Knowledge above, metacognitive instruction could provide the L2 listeners with becoming aware of L2 listening, its process, and its difficulties they could meet. Even less-successful L2 listeners made themselves aware of their own efficiency in L2 listening and compared listening with

other skills and therefore noticed the nature of listening. Together with metacognitive awareness, their self-efficacy beliefs towards L2 listening positively changed even its effect was not larger, which is in line with the literature in which the L2 listeners who are aware of themselves and their listening process with difficulties thanks to their metacognitive awareness have growing self-efficacy and motivation, and develop the oral language and involve in oral interaction effectively, as well as become aware of their learning and after all self-regulated learners (Wenden, 1998; Vandergrift, 2003; Graham & Macaro, 2008).

Regarding preparatory school background. The findings showed that L2 listeners who had a preparatory school background in the treatment group reported a non-significant increase from the beginning through the midpoint, but no changes from the midpoint to the end of the study, and a decrease from the beginning to the end of the treatment process. L2 listeners who did not study at the preparatory school in the experimental group showed no changes from the beginning through the midpoint, but a non-significant decrease from the midpoint to the end of the study and from the beginning to the end of the treatment process. Regarding the experimental group, L2 listeners who studied at the preparatory school had a non-significantly higher level of metacognitive awareness of listening on Person Knowledge than L2 listeners who did not study at the preparatory school, in the beginning, at the midpoint, and the end of the study to a small extent.

The experimental group's L2 listeners who studied at the preparatory school non-significantly outnumbered L2 listeners who studied at the preparatory but did not receive the instruction throughout the process to a medium extent. The experimental group's L2 listeners who did not study at the preparatory school outnumbered in the control group before and during the study to a small extent. In contrast, the control group outnumbered the treatment group after the research to a small extent on Person Knowledge.

According to this, L2 listeners who studied at the preparatory school showed a higher level of Person Knowledge than their counterparts in the experimental group to a medium extent even though both indicated a decrease at the instruction's endpoint. It could be attributed that one-year preparatory school education and one-year exposure to English before majoring at the faculty may increase L2 listeners' self-efficacy and decrease their negative perceptions on L2 listening to a moderate

extent. On the other hand, L2 listeners who just directly started to study at the faculty may have some anxiety problems related to studying at the university for the first time because it was a new environment. They were not familiar with this kind of education after high-school education, and therefore showed a lower level of self-efficacy and the struggle with the difficulties they met. Together with one-year preparatory school education, L2 listeners had more opportunities to practice English at the university, and this provided them with being aware of themselves in terms of their competencies in L2.

Overall Conclusion of the Study

This paper intended to explore the impact of metacognitive instruction on the listening comprehension performance of L2 listeners and reported that L2 listeners showed a statistically significant and greater development in their final listening performance over 8-week metacognitive instruction; this could be explained by metacognitive instruction to a moderate extent. Besides, L2 listeners who received metacognitive instruction showed a non-significantly higher performance than L2 listeners who did not receive metacognitive instruction to a moderate extent. In other words, it could be concluded that metacognitive instruction could moderately facilitate listening comprehension performance. Regarding listening proficiency levels, metacognitive instruction could have a moderately significant impact on enhancing less-successful L2 listeners' listening performance while a large effect on more-successful L2 listeners' listening success on the final test of the treatment group. Also, more-successful L2 listeners who received metacognitive instruction showed a higher level of listening comprehension performance than less-successful L2 listeners who received metacognitive instruction. If L2 listeners had a higher level of listening proficiency, they could show better-listening performance over metacognitive instruction to a large extent in this study context. In terms of preparatory school background, L2 listeners who did not have any preparatory school backgrounds and received metacognitive instruction significantly increased their listening performance at the end of the process, which could be explained by metacognitive instruction to a large extent. Also, in the experimental group, L2 listeners who studied at the preparatory school had a higher level of listening performance than L2 listeners who did not study at the preparatory school at the

end of the study to a medium extent. If L2 listeners had a preparatory school background, they could show better-listening performance over metacognitive instruction to a large extent in this study context.

The study secondly intended to question the impact of metacognitive instruction on the overall metacognitive awareness with metacognitive strategy use, including Problem Solving, Planning and Evaluation, Mental Translation, Directed Attention, and Person Knowledge of L2 listeners. The research concluded that metacognitive instruction significantly affected and increased L2 listeners' overall metacognitive awareness of listening to a medium extent over the research. Regarding listening proficiency levels, more-successful L2 listeners who received metacognitive instruction had a higher level of overall metacognitive awareness to a moderate extent at the endpoint than their counterparts in both groups. In terms of preparatory school background, when L2 listeners received metacognitive instruction with one-year preparatory school background, it could affect non-significantly and slightly their metacognitive awareness of listening.

About Problem-Solving, even though L2 listeners who received metacognitive instruction did not show significant growth in the use of Problem Solving, they significantly outnumbered the control group after the treatment, and the considerable increase in Problem Solving could be explained by metacognitive instruction to a medium extent. Regarding the level of listening proficiency, more-proficient listeners had a moderately greater metacognitive awareness regarding Problem Solving at the endpoint, while less-proficient listeners between the groups did not show any change after the treatment; however, more-proficient listeners showed small growth in Problem Solving between the groups. In the stimulated recall protocols, they stated that guessing the unknown words based on the known words and background knowledge would be helpful; however, when they did not have any background knowledge, they had difficulty in that, and that is why they could not understand the text, which could be related to lower-level proficiency of English. Regarding preparatory school background, one-year preparatory school experience with metacognitive instruction non-significantly and slightly caused a higher level of Problem- Solving from their counterparts in the treatment group, but they outnumbered their counterparts in the control group throughout the study to a moderate extent.

About Planning and Evaluation, metacognitive instruction did not show any significant change in Planning and Evaluation; even the treatment group showed a non-significant slight decrease throughout the study. Regarding the level of listening proficiency, more-successful L2 listeners in the treatment group non-significantly outnumbered more-successful L2 listeners in both groups at the endpoint to a small extent; the level of listening proficiency made no consistent significant change throughout the study in Planning and Evaluation. Similarly, one-year preparatory school background with metacognitive instruction did not have considerable effects on the listeners' metacognitive awareness of listening and strategy use. In the stimulated recall protocols, it was highlighted that having a goal in mind, and following a certain goal caused L2 listeners to lose track of listening because they tried to focus on that goal much more. Also, input type could be an essential factor for having a goal and following it because some text types do not need to have a goal in their mind, such as songs but academic texts or podcasts. Besides, predicting before listening was the strategy, they started to use more, especially My Prediction part they benefited more, but still, less-successful L2 listeners could have a problem in that part because of the lack of contextual or background knowledge. Regarding assessing their understanding, they, especially more-successful L2 listeners, stated that, at the end of the study, reflective exercises using the diaries and assessing themselves in terms of the difficulties they met would be useful for L2 listening; therefore, they would be aware of their level of understanding. After listening, giving transcripts of the text made them see their misunderstood words and pronunciation mistakes, which could contribute to their development in word recognition and auditory discrimination skills.

In terms of Mental Translation, metacognitive instruction could non-significantly and slightly cause to decrease in the use of translation in L2 listening. Regarding the level of listening proficiency, more-successful listeners showed a slight decrease in using translation while less-successful listeners who received metacognitive instruction increased the use of translation; even they used much more translation than less-successful L2 listeners who did not receive metacognitive instruction. All this claimed that metacognitive instruction did not significantly and more moderately facilitate to overcome the use of mental translation. Regarding preparatory school background, the higher use of mental translation by L2 listeners

with one-year of preparatory school backgrounds followed by metacognitive instruction was observed, but its effect was found non-significant and small. In the stimulated recall protocols, more-successful L2 listeners who received metacognitive instruction stated that the use of key-word or word-by-word translation could sometimes be necessary based on the text type and type of input while less-successful L2 listeners tended to use it much more because of the desire to understand everything.

In terms of Directed Attention, the experimental group showed no change in Directed Attention throughout the research; the significant variance found between the groups might only be explained through the decrease in metacognitive awareness on Directed Attention of the control group at the end. In other words, metacognitive instruction had no growing effect on Directed Attention use. Regarding the level of listening proficiency over metacognitive instruction did not cause any significant or growing effect on the use of Directed Attention strategies. Regarding preparatory school background, L2 listeners who did not study at the preparatory school showed a non-significant higher level of metacognitive awareness than L2 listeners who studied at the preparatory school in both groups, to a moderate extent. However, the experimental group did not show any change over the study; the findings were not significant and explained by metacognitive instruction. In the stimulated recall protocols, the L2 listeners stated that in the beginning, especially the less-successful L2 listeners in the experimental group tended to give up listening when they lost track of their concentration because of speaker accent, background noises, personal situations. However, at the end of the study, they tended to try not to give up, keep listening, and recover their concentration even when lost. As a result, it could be deduced that metacognitive instruction could slightly increase metacognitive awareness in Directed Attention of the experimental group even though the quantitative results did not support it.

In terms of Person Knowledge, it could be concluded that metacognitive instruction had no effect on the use of Person Knowledge strategies because the experimental group's Person Knowledge showed a significant decrease over the study to a small extent even though they had slightly greater than the comparison group. Regarding the level of listening proficiency, more-successful L2 listeners had a non-significant higher level of Person Knowledge than less-successful L2 listeners

over metacognitive instruction, which was explained by metacognitive instruction to a medium extent. Regarding preparatory school background, L2 listeners who studied at the preparatory school showed a higher level of Person Knowledge than their counterparts in the experimental group to a medium extent even though both showed a decrease at the end of the study. That is to say, that one-year preparatory school education with metacognitive instruction had no larger effects on their self-efficacy beliefs in listening. In the stimulated recall protocols, the L2 listeners stated that metacognitive instruction could help them become aware of L2 listening, its process, and its difficulties they could meet. Even less-successful L2 listeners made themselves aware of their own efficiency in L2 listening and compared listening with other skills; therefore, noticed L2 listening features. Their self-efficacy beliefs towards L2 listening positively changed with increasing metacognitive awareness even though its effect was not more extensive.

Briefly, metacognitive instruction could have a considerable impact on enhancing the listening performance, overall metacognitive awareness of listening, and the use of Problem Solving; a small effect on the higher use of Planning and Evaluation with the lower use of Mental Translation; no effect on Directed Attention strategy use while a decreasing effect of the use of Person Knowledge strategies.

Implications of the Study

The current research provides fruitful implications for teaching and learning L2 listening regarding pedagogical and methodological aspects. First of all, a process-based metacognitive instruction with pedagogical sequences could facilitate the listening comprehension performance of L2 listeners and their overall metacognitive awareness of listening with metacognitive strategy use. In L2 listening classrooms, as mentioned at the beginning of the study, the instructor first gives learners the text topic and makes them think it over regarding what they know about it. At that part, learners' background knowledge gets activated, and in some classes, brainstorming and prediction activities take place in the classroom before the first listening. It could be considered that here a metacognitive aspect of listening has been reflected in the classroom. This kind of prediction and brainstorming activities are helpful; however, the objectives of these tasks are, again, product-based. How could they be product-based? Learners carry out these activities to answer the

questions related to the text. They listen to the listening text to see if their predictions are correct enough to answer the questions. Nonetheless, the researchers (especially see Vandergrift and Goh's studies) have emphasized throughout this study that a process-based guided metacognitive instruction with pedagogical procedures makes L2 listeners aware of the process of listening, the complex nature of listening, the difficulties they could meet during listening, and the strategies they can use to deal with those difficulties.

In terms of the process of listening, metacognitive instruction could provide L2 listeners with becoming aware of and dealing with the bottom-up processes, including attending and keeping in their mind what they hear, considering possible vocabularies, and create a mental representation of those in their working memory as well as the top-down processes including using their background, contextual knowledge to interpret the intended message by guessing and assessing their understanding and reaching an ultimate meaning after listening. Thanks to process-based metacognitive instruction, L2 listeners actively take place in these processes with their peers and their guided instructor as an active processor. Regarding the difficulties they could meet, they could have the knowledge to struggle with them, called strategy knowledge, by planning, monitoring, making inferences, assessing their understanding and themselves, and finding solutions to those task problems using the metacognitive strategies which are implicitly taught during the instruction. Considering all this, adopting classroom-based metacognitive instruction in English language classrooms helps L2 learners enhance their metacognitive awareness of the listening processes, show better-listening comprehension performances, and eventually become self-regulated autonomous L2 listeners. However, how could this kind of metacognitive instruction be adopted by the current English language learning and teaching system in this study context?

The answers are in our minds as pre-service and/or in-service English language teachers, ELT professors, teacher educators, educational policymakers, curriculum and educational program designers, material and coursebook developers, and/or a person who is responsible for teaching and/or learning English effectively, efficiently, and fluently. A process-based metacognitive instruction with pedagogical procedures should first be adopted by the educational policymakers to bring a process-based listening approach to the classroom rather than product-

based listening. Therefore, the curriculum designers and material and coursebook developers then reflect it to the education and teaching programs of primary schools to high-schools, even to the English classrooms in colleges. The materials could include the sample listening diaries, including the pedagogical procedures and authentic listening tasks.

Following this, in-service ELT teachers in Turkey, especially the ones working for two and more decades, could not be aware of metacognition and metacognitive instruction since the concepts have become so familiar in the ELT classroom for almost two decades. For this reason, in-service training on contemporary teaching methodologies, including a process-based metacognitive approach to English listening, could be provided to the in-service ELT teachers in order to make them familiar with it and practice it in the language classroom. Furthermore, as for pre-service English teachers, also teacher educators need to adopt and practice a process-based metacognitive instruction in the classes such as listening, pronunciation, and speaking classes in the first-year ELT undergraduate program as well as in the classes of teaching listening and speaking skills, teaching integrated skills to teach pre-service ELT teachers this kind of listening instruction. In this way, a process-based metacognitive instruction with pedagogical procedures will be reflected in our lives in every related part.

While adopting it to a real classroom environment, the aspect of pedagogical procedures besides process-based listening should also be prioritized. This type of listening lesson plan, including the stages of prediction, the first listening, the second listening, the third listening with a transcript of the text, and reflection exercises after listening, should be employed. In the current study, L2 listeners stated that the prediction stage would help them meet familiar topics and become aware of their background knowledge about the listening text. In this sense, it becomes vital to decide the listening text in terms of listener needs and levels and to activate their schemata by brainstorming and increasing their vocabulary knowledge. Failing this could result in a lost track of listening; at that point, teaching the use of the problem-solving strategies by guessing the unknown words based on the contextual knowledge or discussing with their peers after the first listening and the second listening through assessing their understanding and deciding the points they need to more focus on could become prominent. Along with this, discussion with their

peers or the whole class discussion also enhances their communication and cooperation skills with their developing self-efficacy beliefs and self-confidence in practicing language skills.

Furthermore, while adopting metacognitive instruction into the class, giving the transcript of the listening test during the third listening could make them realize the sentence structures in the speech with their written forms and the syntactic structures of the text; this is beneficial to their development of word recognition and auditory discrimination skills (Vandergrift & Tafaghodtari, 2010). In the current study, the L2 listeners stated that they noticed the parts they misunderstood during listening and checked the vocabularies on the transcript, and realized their mistakes because of some misunderstood pronunciations. Therefore, they could become recognize the words and their syntactic forms. Considering this, a transcript should be given to L2 listeners for its positive effects on a better listening performance. Also, reflective exercises after listening are a strategy that L2 listeners are not familiar with in the traditional language classrooms in Turkey. Evaluating their understanding, assessing the points they had trouble in and their possible reasons, and deciding the future goals for next listening activities by discussing with their peers and classmates after they write their lesson diaries enhance their metacognitive awareness, self-efficacy beliefs and make them become more-autonomous learners (Vandergrift, 2002). In the current study, L2 listeners indicated that they started to evaluate themselves with this instruction, and by using their reflection notes in their listening diaries, they changed their behaviors. For example, one of them started to listen to different people with different accents because she realized that she needed to develop herself in terms of accent. Regarding all this, this kind of lesson plan should be adopted to develop metacognitive awareness and listening performance as well as to boost self-assessment and learner autonomy, which are the concepts Turkish students are not familiar with and need to learn how to take responsibility of and assess their own learning (Şentürk & Mirici, 2020).

On the one hand, this report showed that L2 listeners could have problems in listening because of speaker accent, classroom environment, the quality of the audio, the input type of the text, background noises, concentration issues, and some personal problems such as anxiety and lower self-efficacy beliefs, as stated in the literature (Goh, 2000). With the aim of preventing or dealing with these problems,

the instructor/the teacher or the educational program developer should be careful in deciding the listening text in the coursebook or during the lesson. They should provide L2 listeners with opportunities for practicing different accents inside and outside the classroom. Listening to different kinds of texts from different people could help L2 listeners become aware of and get used to these differences. Again, much practice of authentic listening texts such as songs, podcasts, and academic texts could be helpful for concentration issues. With this, their problems related to themselves and their self-efficacy beliefs could be decreased.

Secondly, the study's methodological implications could be evaluated regarding the adopted research design and data collection tools. The mixed-methods research design provided the study with completing, validating, and comparing the numerical data on listening performance and metacognitive awareness of listening with the L2 listeners' perceptions on their growth in metacognitive awareness. It helped reveal the changes in their awareness and their reasons and assess them from different perspectives. In addition, the MALQ ensured their metacognitive awareness level of listening; furthermore, it could be used in the classroom context: it can be a consciousness-raising tool for students; therefore, they can be aware of the process of listening, and consequently, they raise their metacognitive awareness and learn how to listen better and become self-regulated learners; it can be a diagnostic and teaching tool; hence, teachers determine students' current metacognitive awareness and evaluate their preferences of using metacognitive strategies; and it can be a research tool, researchers could conduct it as a pretest-posttest to determine metacognitive awareness of listening of the sample group (Vandergrift et al., 2006). In addition, in the study, the researcher was the instructor of both groups; this helped control the variable of the instructor. Moreover, L2 listeners with different listening proficiency levels and preparatory school backgrounds, a new concept regarded as a variable in this study area, were involved in this study. Along with this, it presented a new original study based on a mixed methodology of 8-week process-based implicit metacognitive instruction, which provided a transcript during third listening, to the literature of L2 listening in Turkey and different contexts. Therefore, the present study tried to put its remarkable findings into the gap in the ELT literature and showed a different perspective in English language learning and teaching.

Suggestions for Future Research

Regarding the gap needing more remarkable studies on this perspective, the study calls for more research into L2 listening learning and teaching. This research focused on exploring the effect of 8-week process-based metacognitive instruction on 37 first-grade ELT students' listening performance and metacognitive awareness of listening in Turkey.

First, further research could be done with a larger study group to take a broader view of the results. Also, it could be investigated for a longer-term or as a longitudinal study. Therefore, the impact of a longer-term metacognitive instruction could be observed and evaluated in a broader period on the growth in listening success and metacognitive awareness.

Besides, the data of the study were normally not distributed; non-parametric methods were applied, and the participants were voluntarily chosen and assigned to the group; a quasi-experimental study was adopted. For future studies, the sampling groups may be organized by different sampling methods and variables. Gender, age groups (adults or younger listeners), different grade levels (high-school or primary school students), or different levels of proficiency (beginner, elementary, or advanced learners) could be regarded as variables that could affect the study results, and the effect of metacognitive instruction could be examined based on one or more of these different factors.

Also, for future research, the European Language Portfolio, a tool proposed by the Language Policy Division of the Council of Europe providing learners with recording and reflecting their learning experiences and enhancing motivation, self-assessment and learner autonomy in language learning (Council of Europe, 2001) could be used to record and monitor the development of listening performance and metacognitive awareness of listeners in L2 listening.

Also, in the literature, studies based on the mixed-methods research design are quite limited. Future research could be done by employing thinking-aloud procedures, interviews, with surveys. Furthermore, each lesson in which metacognitive instruction is practiced may be video-recorded. Thus, other possible points of view of metacognitive instruction in the classroom could be observed and considered a basis for further studies.

References

- Agoos, S. (2016, March). Transcript of "5 tips to improve your critical thinking". Retrieved from https://www.ted.com/talks/samantha_agoos_5_tips_to_improve_your_critical_thinking/transcript#t-41221
- Altuwairesh, N. (2016). EFL Saudi undergraduate students' use of metacognitive listening strategies. *Arab World English Journal*, 7(1), 78-93.
- Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). New York: Freeman.
- Anderson, A., & Lynch, T. (1988). *Listening*. London: Oxford University Press.
- Aron, A., & Aron, E. (1999). *Statistics for psychology* (2 ed.). Upper Saddle River, NJ: Prentice Hall.
- Bacon, S. M. (1992). Phases of listening to authentic input in spanish: a descriptive study. *Foreign Language Annals*, 25(4), 317-334.
- Birjandi, P., & Rahimi, A. H. (2012). The effect of metacognitive strategy instruction on the listening performance of EFL students. *International Journal of Linguistics*, 4(2), 495-517.
- Bozorgian, H. (2012). Metacognitive instruction does improve listening comprehension. *ISRN Education*, 1–6. doi:10.5402/2012/734085
- Bozorgian, H. (2014). The role of metacognition in the development of EFL learners' listening skill. *International Journal of Listening*, 28(3), 149–161. <https://doi.org/10.1080/10904018.2013.861303>
- Bozorgian, H. (2015). *Metacognitive strategy instruction: Empowering listening skill*. Germany: Scholar's Press.
- Bozorgian, H., & Alamdari, E. F. (2017). Multimedia listening comprehension: Metacognitive instruction or metacognitive instruction through dialogic interaction. *ReCALL*, 30(1), 131–152. <https://doi.org/10.1017/S0958344016000240>
- Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbooks in*

psychology®. *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. <https://doi.org/10.1037/13620-004>

- Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54(1), 106–148. <https://doi.org/10.1111/j.1467-6494.1986.tb00391.x>
- Buck, G. (2001). *Assessing listening*. Cambridge, UK: Cambridge University Press.
- Chamot, A. U., & L., Küpper. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals*, 22(I), 13-24.
- Churchill, G. A. Jr., & Peter, J. P. (1984). Research design effects on the reliability of rating scales: a meta-analysis. *Journal of Marketing Research*, 21 (November), 360-375.
- Clark, L., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Coşkun, A. (2010). The effect of metacognitive strategy training on the listening performance of beginner students. *Novitas-ROYAL (Research on Youth and Language)*, 4(1), 35-50.
- Council of Europe (2001). *Common European Framework of References for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and Conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Cross, J. (2009). Effects of listening strategy instruction on news videotext comprehension. *Language Teaching Research*, 13(2), 151–176. doi:10.1177/ 1362168809103446

- Cross, J. (2010). Raising L2 listeners' metacognitive awareness: a sociocultural theory perspective, *Language Awareness*, 19(4), 281-297, DOI: 10.1080/09658416.2010.519033
- Cross, J. (2011). Metacognitive instruction can help less-skilled listeners. *ELT Journal*, 65(4), 408–416. doi:10.1093/elt/ccq073
- Çevikbaş, H. (2016). *The effect of instructional activities designed with metacognitive strategies on developing listening skills of students in university prep class*. Unpublished master dissertation, University of Akdeniz, Turkey.
- Dall'Oglio A. M., Rossiello B., Coletti M. F., Caselli M. C. , Ravà L., di Ciommo V., Orzalesi M., Giannantoni P., Pasqualetti P. (2010). Developmental evaluation at age 4: Validity of an Italian parental questionnaire. *J Paediatr Child Health*, 46(7-8), 419-26. doi: 10.1111/j.1440-1754.2010.01748.x. Epub 2010 Jun 8. PMID: 20546104.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Dunning, D. (2017, November). Transcript of "Why incompetent people think they're amazing. Retrieved from https://www.ted.com/talks/david_dunning_why_incompetent_people_think_they_re_amazing/transcript
- Ehrich, J. F., & Henderson, D. B. (2019). Rasch analysis of the metacognitive awareness listening questionnaire (MALQ), *International Journal of Listening*, 33(2), 101-113, DOI: 10.1080/10904018.2017.1418350
- Field, A. (2009). *Discovering statistics using SPSS*. (3rd ed.). Sage Publications Ltd., London.
- Field, J. (2008). *Listening in the language classroom*. Cambridge: Cambridge University Press.
- Field, J. (2019). Second language listening: current ideas, current issues. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge Handbook of Language Learning* (pp. 283-319). Cambridge: Cambridge University Press.

- Flavell, J. H. (1976). Metacognitive aspects of Problem Solving. In Resnick, L. B. (Ed.), *The nature of intelligence* (pp. 231–236). Hillsdale, NJ: Erlbaum.
- Flavell, J. H. (1979) Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906–911.
<https://doi.org/10.1037/0003-066X.34.10.906>
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Goh, C. (1997). Metacognitive awareness and second language students. *ELT Journal*, 51, 361-369. <https://doi.org/10.1093/elt/51.4.361>
- Goh, C. (1998). How ESL students with different listening abilities use comprehension strategies and tactics. *Language Teaching Research*, 2(2), 124-147. <http://dx.doi.org/10.1177/136216889800200203>
- Goh, C. (1999). How much do learners know about the factors that influence their listening comprehension? *Hong Kong Journal of Applied Linguistics* 4(1), 17-40.
- Goh, C. (2000). A cognitive perspective on language students' listening comprehension problems. *System*, 28(1), 55-75.
- Goh, C. (2002). *Teaching listening in the language classroom*. Singapore: SEAMEO Regional Language Centre.
- Goh, C. (2005). Second language listening expertise. In K. Johnson (Ed.), *Expertise in second language learning and teaching* (pp. 64–84). Palgrave Macmillan.
- Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELJ Journal: A journal of language teaching and research in Southeast Asia*, 39(2), 188-213.
- Goh, C. (2010). Listening as process: Learning activities for self-appraisal and self-regulation. In N. Harwood (Ed.), *Materials in ELT: Theory and practice* (pp. 179–206). Cambridge: Cambridge University Press.

- Goh, C. (2014). Reconceptualising second language oracy instruction: Metacognitive engagement and direct teaching in listening and speaking. *The Asian Journal of English Language & Pedagogy*, 2, 1-31.
- Goh, C., & Hu, G. (2013). Exploring the relationship between metacognitive awareness and listening performance with questionnaire data. *Language Awareness*, 23(3), 255–274. <https://doi.org/10.1080/09658416.2013.769558>
- Goh, C., & Taib, Y. (2006). Metacognitive instruction in listening for young students. *ELT Journal*, 60, 222-232.
- Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34(2), 165–182. doi:10.1080/09571730385200221
- Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate students of French. *Language Learning*, 58(4), 747-783.
- Greene, J., Caracelli, V., & Graham, W. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255-274.
- Henner Stanchina, C. (1987). Autonomy as metacognitive awareness: suggestions for training self-monitoring of listening comprehension. *Mélanges Pédagogiques*, 69-84.
- Johnson, B., & Christensen, L. (2012). *Educational research: Quantitative, qualitative, and mixed approaches*. Los Angeles, CA: SAGE publication.
- Johnson, R.B., & Onwuegbuzie, A.J. (2004) Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33, 14-26. <http://dx.doi.org/10.3102/0013189X033007014>
- Katrancı, M. & Yangın, B. (2013). Üstbiliş stratejileri öğretiminin dinlediğini anlama becerisine ve dinlemeye yönelik tutuma etkisi. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(11), 733-771.
- Lewis-Beck, M. S., Bryman, A., & Futing Liao, T. (2004). *The SAGE encyclopedia of social science research methods*, (1-0). Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412950589

- Listening practice tests. (2019). Retrieved June 20, 2019, from <https://takeielts.britishcouncil.org/take-ielts/prepare/free-ielts-practice-tests/listening>
- Liu, X. L., & Goh, C. (2006). Improving second language listening: Awareness and involvement. In T. S. C. Farrell (Ed.), *Language teacher research in Asia* (pp. 91-106). Alexandria, VA: TESOL.
- Maftoon, P., & Alamdari, E. F. (2020). Exploring the effect of metacognitive strategy instruction on metacognitive awareness and listening performance through a process-based approach, *International Journal of Listening*, 34(1), 1-20, DOI: 10.1080/10904018.2016.1250632
- Mareschal, C. (2007). *Student perceptions of a self-regulatory approach to second language listening comprehension development*. Unpublished doctoral dissertation, University of Ottawa, Canada.
- Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice*, 13, 522-525.
- Mendelsohn, D. (1994). *Learning to listen: A strategy-based approach for the second-language learner*. San Diego, CA: Dominic Press.
- Murphy, J. (1985). *An investigation into the listening strategies of ESL college students*. (ERIC Document Reproduction Service No. ED278275)
- Nacamulli, M. (2015, June). The benefits of a bilingual brain. Retrieved from https://www.ted.com/talks/mia_nacamulli_the_benefits_of_a_bilingual_brain?utm_campaign=tedsread&utm_medium=referral&utm_source=tedcomshare
- Nunan, D. (1998). Strategy training in the classroom: an empirical investigation. *A Journal of Language Teaching and Research*, 28(2), 56-81.
- Nunan, D. (2002). Listening in language learning. In J. C. Richards, & W. A. Renandya (Eds.), *Methodology in language teaching* (pp. 238-41). Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511667190.032>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9781139524490>
- O'Malley, J. M., Chamot, A. U., & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, 10(4), 418-437. <http://dx.doi.org/10.1093/applin/10.4.418>
- Oxford, R. L. (1993). Research on second language learning strategies. *Annual Review of Applied Linguistics*, 13, 175-187.
- Oz, H. (2005). Metacognition in foreign/second language learning and teaching. *Hacettepe University Faculty of Education Journal*, 29(29), 147-156.
- Paris, S. G., Winograd, P., 1990. How metacognition can promote academic learning and instruction. In: Jones, B.F., Idol, L. (Eds.), *Dimensions of Thinking and Cognitive Instruction*. Lawrence Erlbaum Associates, New Jersey.
- Pallant, J. (2011). *SPSS survival manual: A step by step guide to data analysis using the SPSS program*. (4rd ed.). Allen & Unwin, Berkshire.
- Peat, J. K., Mellis, C., Williams, K., & Xuan, W. (2001). *Health Science Research: A Handbook of Quantitative Methods*. London: Allen & Unwin.
- Peterson, R. A. (1994). A meta-analysis of cronbach's coefficient alpha. *Journal of Consumer Research*, 21(2), 381-391.
- Phakiti, A. (2014). *Experimental Research Methods in Language Learning*. Bloomsberry, London.
- Purdy, M. (1997). What is listening? In M. Purdy, & D. Borisoff (Eds.), *Listening in Everyday Life. A Personal and Professional Approach*. Lanham, MD: University Press of America.
- Rahimi, M., & Katal, M. (2012a). Metacognitive listening strategies awareness in learning English as a foreign language: A comparison between university and high-school students. *Procedia - Social and Behavioral Sciences*, 31, 82–89.
- Rahimi, M., & Katal, M. (2012b). The role of metacognitive listening strategies awareness and podcast-use readiness in using podcasting for learning

- English as a foreign language. *Computers in Human Behaviour*, 28, 1153-1161. <http://dx.doi.org/10.1016/j.chb.2012.01.025>
- Rahimi, M., & Katal, M. (2013). The impact of metacognitive instruction on EFL learners' listening comprehension and oral language proficiency. *The Journal of Teaching Language Skills (JTLS)*, 5 (2), 69-90.
- Rosenthal R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- Rost, M. (2001). Listening. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 7-13). Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511667206.002>
- Rost, M. (2002). *Teaching and researching listening*. London: Longman.
- Scarcella, R. C., & Oxford. R. L. (1992). *The tapestry of language learning: The individual in the communicative classroom*. Boston, MA: Heinle & Heinle.
- Schmitt, N. (1994), Method bias: The importance of theory and measurement. *Journal of Organizational Behaviour*, (15), 393-398. <https://doi.org/10.1002/job.4030150504>
- Seferoğlu, G. & Uzakgören, S. (2004). Equipping learners with listening strategies in English language classes. *Hacettepe University Faculty of Education Journal*, 27, 223-231.
- 6 Minute English / Being a sedentary life. (2019, July 1). Retrieved from <http://www.bbc.co.uk/learningenglish/english/features/6-minute-english/ep-180906>
- 6 Minute English / Is talking on the phone embarrassing?. (2019, July 1). Retrieved from <http://www.bbc.co.uk/learningenglish/english/features/6-minute-english/ep-1906138>
- Şentürk, B. & Mirici, I. H. (2020). The effect of different self-assessment tools on students' attitudes towards learning English. *Elementary Education Online*, 19(1), 42-53. doi:10.17051/ilkonline.2020.643689
- Test Statistics. (2020). Retrieved from <https://www.ielts.org/research/test-statistics>

- Thompson, I., & Rubin, J (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals*, 29, 331–342.
<http://dx.doi.org/10.1111/j.1944-9720.1996.tb01246.x>
- Underwood, M. (1989). *Teaching listening*. New York: Longman.
- Vandergrift, L. (1997). The comprehension strategies of second language (French) listeners: A descriptive study. *Foreign Language Annals*, 30, 387-409.
- Vandergrift, L. (1998). Successful and less successful listeners in French: What are the strategy differences? *The French Review*, 71(3), 370–395.
- Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal*, 54, 168-176.
- Vandergrift, L. (2002). “It was nice to see that our predictions were right”: Developing metacognition in L2 listening comprehension. *Canadian Modern Language Review*, 58, 555–75.
- Vandergrift, L. (2003a). Orchestrating strategy use: Towards a model of the skilled L2 listener. *Language Learning*, 53, 461–494.
- Vandergrift, L. (2003b). From prediction to reflection: Guiding students through the process of L2 listening. *Canadian Modern Language Review*, 59, 425–440.
- Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics*, 24, 3-25.
- Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied Linguistics*, 26, 70-89.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40, 191–210.
- Vandergrift, L., & Goh, C. (2012). *Teaching and learning second language listening: Metacognition in action*. New York, NY: Routledge.
- Vandergrift, L., Goh, C., Mareschal, C., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire (MALQ): Development and validation. *Language Learning*, 56, 431–462.
doi:10.1111/j.1467-9922.2006.00373.x

- Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 students how to listen does make a difference: An empirical study. *Language Learning*, 60, 470–497. doi:10.1111/j.1467-9922.2009.00559.x
- Veenman, M. V. J., Van Hout-wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, 1, 3–14.
- What the IELTS? (n.d.). Retrieved from <https://www.ielts.org/what-is-ielts/ielts-introduction>
- Wenden, A. (1987). Metacognition: An expanded view of the cognitive abilities of L2 learners. *Language Learning*, 37(4), 573-94.
- Wenden, A. (1991). *Learner strategies for learner autonomy*. London, England: Prentice Hall.
- Wenden, A. (1998). Metacognitive knowledge and language learning. *Applied Linguistics*, 19, 515–537. doi:10.1093/applin/19.4.515
- Young, M.Y.C. (1997). A serial ordering of listening comprehension strategies used by advanced ESL learners in Hong Kong. *Asian Journal of English language Teaching*, 7, 35-53.
- Zeng, Y. (2007). Metacognitive instruction in listening: A study of Chinese non-English major undergraduates (Unpublished master's thesis). National Institute of Education, Nanyang Technological University, Singapore.
- Zeng, Y. (2012). Metacognition and self-regulated learning (SRL) for Chinese EFL listening development (Unpublished doctoral dissertation). Nanyang Technological University, Singapore.
- Zhang, D.L., & Goh, C. (2006). Strategy knowledge and perceived strategy use: Singaporean students' awareness of listening and speaking strategies. *Language Awareness*, 15(3), 199-219.

APPENDIX-A: Consent Form

Gönüllü Katılım Formu

..../..../.....

Merhaba,

Yapacak olduğum araştırma için ayırdığın zaman ve ilgili tavrın için şimdiden çok teşekkür ederim. Gönüllü katılım formu ile ne amaçladığımı ve çalışmaya katılman durumunda çalışma boyunca neler yapacağımızı anlatmayı amaçladım.

Bu araştırma için Hacettepe Üniversitesi Etik Komisyonundan izin alınmıştır. Araştırma, üstbilişsel eğitimin yabancı dil öğrenenlerin dinleme anlama becerisi üzerine etkisini incelemek için, Dr. Öğr. Üyesi İsmail Fırat ALTAY danışmanlığında hazırlanacak bir yüksek lisans tezidir. Bu sebeple de, sınıfta yapılacak olan aktivitelere katılman ve bu aktivitelerle ilgili sunacağın görüşlerin çok önemli.

Araştırmaya gönüllü olarak katılım esastır. Gönüllü olduğun takdirde bu süreçte dinleme başarı testi yapmak, ders etkinliklerini takip etmek, anket uygulamak ve uyarılmış geri çağırma görüşmeleri yapmak istiyorum. Bu görüşme sırasında ses kaydı almak da istiyorum. Bunun yanı sıra, yapılan aktivitelerde de fotoğraf çekmek ve gerekirse tezime eklemek istiyorum. Bütün bu kayıtlar ve görüşmeler sadece bilimsel bir amaç için kullanılacak ve bunun haricinde hiçbir şekilde kullanılmayacaktır. İstediğin takdirde görüşmeler ve kayıtlar silinebilecek ya da sana geri verilecektir. Araştırma süresince gerçek adın yerine takma bir ad kullanılacaktır. İstediğin zaman görüşmeyi kesebilir ya da çalışmadan ayrılabilirsin. Böyle bir durumda elde edilen görüşmeler ve kayıtlar çalışma için kullanılmayacaktır.

Bu bilgiler ışığında araştırmaya gönüllü katılımını ve sağladığın güvenceye dayanarak bu formu imzalamanı rica ediyorum. Sormak istediğin herhangi bir durumda ya da araştırma sonucu hakkında bilgi almak istediğinde benimle her zaman iletişime geçebilirsin. Formu okuyarak imzaladığın için çok teşekkür ederim.

*** Bu araştırma için Hacettepe Üniversitesi Etik Komisyonundan izin alınmıştır.**

Katılımcı Öğrenci

Adı-Soyadı:

Adres:

İmza:

Sorumlu Araştırmacı

Dr. Öğr. Üyesi

İsmail Fırat ALTAY

HÜ Eğitim Bilimleri Enstitüsü

İmza:

Araştırmacı:

Kübra YETİŞ

Düzce Üniversitesi

Eğitim Fakültesi

İmza:

APPENDIX-B: Weekly Listening Texts

Week (45-minute lessons)	Topic of Listening Text	Source
1st week	Is talking on the phone embarrassing?	BBC 6 Minute English - General
2nd week	The Benefits of a Bilingual Brain	TED Education
3rd week	Beating a Sedentary Lifestyle	BBC 6 Minute English - General
4th week	5 Tips to Improve to Your Critical Thinking	TED Education
5th week	Why incompetent people think they're amazing?	TED Education
6th week	Are you human?	TED Talk
7th week	3 things I learned while my plane crashed.	TED Talk
8th week	8 secrets of success	TED Talk

APPENDIX-C: Sample Page of Listening Diary

For the experimental group:

Topic:

Date:

My predictions:

(Note down main ideas that you think will be mentioned in the text)

-	_____	_____
-	_____	_____
-	_____	_____
-	_____	_____

First Listening:

(Listen to the text. Put a tick next to the ideas that you predicted and that were also mentioned in the text, and note down any other ideas that you could not predict but were mentioned. Then, discuss your listening results with your peer.)

-	_____	_____
-	_____	_____
-	_____	_____
-	_____	_____

Second Listening:

(Listen to the text again to check your predictions and to try to solve problems in understanding between you and your partner. Add any potential points and significant details that you may not have comprehended during the first listen.)

-	_____	_____
-	_____	_____
-	_____	_____

Third Listening: *Listen to the text for the third time to confirm your understanding after a whole-class discussion while reading the transcript of the text.*

Reflection:

What I succeeded:

What I had trouble in:

What I will do for future listening:

For the control group:

Topic:

Date:

First Listening:

- _____
- _____
- _____
- _____

Second Listening:

- _____
- _____
- _____
- _____

Third Listening with the transcript of the text:

APPENDIX-D: The IELTS Listening Test

The IELTS Listening Test

Nickname: _____

Time 40 minutes

Instructions and Information

Listen to the instructions for each part of the paper carefully.

Answer all the questions.

While you are listening, write your answers on the question paper.

You will have 10 minutes at the end of the test to copy your answers onto the separate answer sheet. Use a pencil.

At the end of the test, hand in this question paper.

There are **four** parts to the test.

You will hear each part **once** only.

There are **40** questions.

Each question carries one mark.

For each part of the test, there will be time for you to look through the questions and time for you to check your answers.

The IELTS Listening Test

Section 1 Questions 1–10

Questions 1–5

Complete the notes below.

Write no more than two words and/or a number for each answer.

Transport from Bayswater

Example	Answer
Destination	Harbour City

- Express train leaves at **1**
- Nearest station is **2**
- Number 706 bus goes to **3**
- Number **4** bus goes to station
- Earlier bus leaves at **5**

Questions 6–10

Complete the table below.

Write no more than one word and/or a number for each answer.

Transport	Cash fare	Card fare
Bus	6 \$	\$1.50
Train (peak)	\$10	\$10
Train (off-peak) – before 5 pm or after 7 pm)	\$10	8 \$
9 ferry	\$4.50	\$3.55
Tourist ferry (10)	\$35	–
Tourist ferry (whole day)	\$65	–

Section 2

Questions 11–20

Questions 11–14

Which counsellor should you see?

Write the correct letter, **A**, **B** or **C**, next to questions 11–14.

A Louise Bagshaw

B Tony Denby

C Naomi Flynn

11 if it is your first time seeing a counsellor

12 if you are unable to see a counsellor during normal office hours

13 if you do not have an appointment

14 if your concerns are related to anxiety

Questions 15–20

Complete the table below.

Write no more than two words for each answer.

Workshop	Content	
Adjusting	what you need to succeed academically	15 students
Getting Organised	use time effectively, find 16 between study and leisure	all students
Communicating	talking with staff, communicating across cultures	all students, especially 17
Anxiety	18, breathing techniques, meditation, etc.	students about to sit exams
19	staying on track for long periods	20 students only

Section 3 Questions 21–30

Questions 21–30

Complete the notes below.

Write no more than three words for each answer.

Novel: **21**

Protagonists: Mary Lennox; Colin Craven

Time period: Early in **22**

Plot: Mary → UK – meets Colin who thinks he'll never be able to
23 They become friends.

Point of view: "Omniscient" – narrator knows all about characters' feelings,
opinions and **24**

Audience: Good for children – story simple to follow

Symbols (physical items that represent **25**):

- the robin redbreast
- **26**
- the portrait of Mistress Craven

Motifs (patterns in the story):

- the Garden of Eden
- secrecy – metaphorical and literal transition from **27**

Themes: Connections between

- **28** and outlook
- **29** and well-being
- individuals and the need for **30**

Section 4 Questions 31–40

Questions 31–35

Complete the table below.

Write one word only for each answer.

Time Perspectives		
Time Zone	Outlook	Features & Consequences
Past	Positive	Remember good times, e.g. birthdays. Keep family records, photo albums, etc.
	31	Focus on disappointments, failures, bad decisions.
Present	Hedonistic	Live for 32 ; seek sensation; avoid pain.
	Fatalistic	Life is governed by 33 , religious beliefs social conditions. Life's path can't be changed.
Future	34	Prefer work to play. Don't give in to temptation
	Fatalistic	Have a strong belief in life after death and importance of 35 in life.

Questions 31–35

Choose the correct letter, **A**, **B** or **C**.

36. We are all present hedonists

A at school

B at birth

C while eating and drinking

37. American boys drop out of school at a higher rate than girls because

A they need to be in control of the way they learn

B they play video games instead of doing school work

C they are not as intelligent as girls

38. Present-orientated children

A do not realise present actions can have negative future effects

B are unable to learn lessons from past mistakes

C know what could happen if they do something bad, but do it anyway

39. If Americans had an extra day per week, they would spend it

A working harder

B building relationships

C sharing family meals

40. Understanding how people think about time can help us

A become more virtuous

B work together better

C identify careless or ambitious people

Listening Proficiency Test Answer Sheet

Nickname:

1.	2.
3.	4.
5.	6.
7.	8.
9.	10.
11.	12.
13.	14.
15.	16.
17.	18.
19.	20.
21.	22.
23.	24.
25.	26.
27.	28.
29.	30.
31.	32.
33.	34.
35.	36.
37.	38.
39.	40.

APPENDIX-E: Metacognitive Awareness Listening Questionnaire (MALQ)

Dear Participant,

The following survey is about some strategies for listening comprehension and how you feel about listening in English. I would like you to state whether you agree with them or not. By responding to these statements, you can help me diagnose and assess your awareness of processes underlying successful listening in English, which will contribute to my master's thesis. Also, you can help yourself and understand your progress in learning to listen.

This is not a test, so there is no "right" or "wrong" answers. You don't even have to write your real name on it, I want you to write a nickname on it instead. It is very important that you give your answers sincerely, as only this will guarantee the success of the investigation. Thank you very much for your help.

Kübra YETİŞ

Hacettepe University- ELT (M.A.)

Part A: Student Demographics

1. Nickname: _____

2. Year of Birth: _____

3. Gender: () Female () Male Other: _____

Part B: Metacognitive Awareness Listening Questionnaire

Please indicate your opinion after each statement. Circle the number which best shows your level of agreement with the statement. For example:

Items	Strongly disagree Kesinlikle katılmıyorum	Disagree Katılmıyorum	Slightly disagree Nispeten katılmıyorum	Partly agree Kısmen katılıyorum	Agree Katılıyorum	Strongly agree Kesinlikle katılıyorum
I like learning another language. <i>İkinci bir dil öğrenmeyi seviyorum.</i>	1	2	3	4	5	6

Please circle only **ONE** number for each statement.

Items	Strongly disagree Kesinlikle katılmıyorum	Disagree Katılmıyorum	Slightly disagree Nispeten katılmıyorum	Partly agree Kısmen katılıyorum	Agree Katılıyorum	Strongly agree Kesinlikle katılıyorum
1. Before I start to listen, I have a plan in my head for how I am going to listen. <i>Dinlemeye başlamadan önce, kafamda nasıl işleyeceğime dair plan yaparım.</i>	1	2	3	4	5	6

Items	Strongly disagree Kesinlikle katılmıyorum	Disagree Katılmıyorum	Slightly disagree Nispeten katılmıyorum	Partly agree Kısmen katılıyorum	Agree Katılıyorum	Strongly agree Kesinlikle katılıyorum
2. I focus harder on the text when I have trouble understanding. <i>Anlamada sıkıntı yaşadığım zaman dinleme parçasına daha çok odaklanırım.</i>	1	2	3	4	5	6
3. I find that listening in English is more difficult than reading, speaking, or writing in English. <i>İngilizce dinlemenin, İngilizce okumaktan, konuşmaktan veya yazmaktan daha zor olduğunu düşünürüm.</i>	1	2	3	4	5	6
4. I translate in my head as I listen. <i>Dinleme sırasında kafamın içinde çeviri yaparım.</i>	1	2	3	4	5	6
5. I use the words I understand to guess the meaning of the words I don't understand. <i>Anlamadığım kelimelerin anlamlarını tahmin etmek için anladığım kelimeleri kullanırım.</i>	1	2	3	4	5	6
6. When my mind wanders, I recover my concentration right away. <i>Başka şeyler düşünmeye başladığım zaman konsantrasyonumu derhal toplarım.</i>	1	2	3	4	5	6
7. As I listen, I compare what I understand with what I know about the topic. <i>Dinleme sırasında anladığım şeyler ile konu hakkında bildiğim şeyleri karşılaştırırım.</i>	1	2	3	4	5	6
8. I feel that listening comprehension in English is a challenge for me. <i>İngilizcede dinlediğini anlamanın benim için bir zorluk olduğunu hissederim.</i>	1	2	3	4	5	6
9. I use my experience and knowledge to help me understand. <i>Anlamama yardımcı olması için önceki deneyimlerimden ve bilgilerimden yararlanırım.</i>	1	2	3	4	5	6
10. Before listening, I think of similar texts that I may have listened to. <i>Dinlemeden önce, daha önce dinlemiş olabileceğim benzer dinleme parçalarını düşünürüm.</i>	1	2	3	4	5	6
11. I translate key words as I listen. <i>Dinleme sırasında anahtar kelimelerin çevirisini yaparım.</i>	1	2	3	4	5	6
12. I try to get back on track when I lose concentration. <i>Konsantrasyonumu kaybettiğim zaman dinlemeye devam etmeye çalışırım.</i>	1	2	3	4	5	6
13. As I listen, I quickly adjust my interpretation if I realize that it is not correct. <i>Dinleme sırasında çıkardığım anlamların yanlış olduğunu fark edersem onları çabucak düzeltirim.</i>	1	2	3	4	5	6

Items	Strongly disagree Kesinlikle katılmıyorum	Disagree Katılmıyorum	Slightly disagree Nispeten katılmıyorum	Partly agree Kısmen katılıyorum	Agree Katılıyorum	Strongly agree Kesinlikle katılıyorum
14. After listening, I think back to how I listened, and about what I might do differently next time. <i>Dinleme sonrasında nasıl dinlediğim hakkında ve gelecek zaman daha farklı neler yapabilirim konusunda düşünürüm.</i>	1	2	3	4	5	6
15. I don't feel nervous when I listen to English. <i>İngilizce dinleme sırasında gergin hissetmem.</i>	1	2	3	4	5	6
16. When I have difficulty understanding what I hear, I give up and stop listening. <i>Duyduğumu anlama sırasında sıkıntı yaşadığım zaman, pes ederim ve dinlemeyi bırakırım.</i>	1	2	3	4	5	6
17. I use the general idea of the text to help me guess the meaning of the words that I don't understand. <i>Anlamadığım kelimelerin anlamlarını tahmin etmeme yardımcı olması için dinleme parçasının ana fikrini kullanırım.</i>	1	2	3	4	5	6
18. I translate word-by-word, as I listen. <i>Dinleme sırasında kelimesi kelimesine çeviri yaparım.</i>	1	2	3	4	5	6
19. When I guess the meaning of a word, I think back to everything else that I have heard, to see if my guess makes sense. <i>Bir kelimenin anlamını tahmin ettiğim zaman, tahminimin mantıklı olup olmadığını kontrol etmek için duyduğum diğer şeyleri tekrar düşünürüm.</i>	1	2	3	4	5	6
20. As I listen, I periodically ask myself if I am satisfied with my level of comprehension. <i>Dinleme sırasında anlama seviyemden memnun olup olmadığımı belirli aralıklarla kendime sorarım.</i>	1	2	3	4	5	6
21. I have a goal in mind as I listen. <i>Dinleme sırasında aklımda bir amaç vardır.</i>	1	2	3	4	5	6

APPENDIX-F: Sample Questions of Stimulated Recall Protocols

The Turkish Version of Questions of Stimulated Recall Protocols

1. Madde X'deki cevabınızda bir farklılık olmuş. Birinci / İkinci / Üçüncü ölçekte, siz kesinlikle katılıyorum cevabını vermişsiniz, ancak Birinci / İkinci / Üçüncü ölçekte siz bu cevabınızı kısmen katılıyorum olarak değiştirmişsiniz. Burada bir değişiklik var. Bu ölçekte bu cevabı verirken ne düşünüyordunuz ve şu ölçekte bu cevabı verirken ne düşünüyordunuz? Bu noktalar üzerinde biraz düşünelim.
2. Bu değişikliğin sebepleri neler olabilir? Bu seçeneği işaretliyorken ne düşündüğünüzü hatırlayabilir misiniz?
3. Bu maddede bu seçeneği işaretliyorken kendi dinleme performansınızda ne fark etmişsiniz? Neydi o? Neden onu seçtiniz?
4. (İkinci ve üçüncü ölçeklerin bütün maddeleri karşılaştırıldıktan sonra, ikinci protokol sonunda) Daha iyi bir yabancı dilde dinleme performansına sahip olmak için neler önerebilirsiniz? Yabancı dilde dinleme anlama becerinizi nasıl geliştirebilirsiniz? Daha iyi bir yabancı dil dinleyicisi olmak için neler yapılabilir?

The English Version of Questions of Stimulated Recall Protocols

1. There is a difference on your response on the statement X. In the first/second/third MALQ, you chose, for example, "strongly agree"; however, in the first/ second/third MALQ, you changed it with "slightly agree". There is a change here. What were you thinking here? What were you thinking there? Let's think over these points.
2. What could the reasons for this change be? Can you remember what you were thinking while you were choosing that statement?
3. What did you realize on your listening performance when choosing this option? What could they be? Why did you choose it?
4. (At the end of the second stimulated-recall protocol after comparing all responses of the second and third MALQ) What can you suggest for a better L2 listening performance? How can you improve your L2 listening comprehension skill? What could be done to be a more successful L2 listener?

APPENDIX-G: Ethics Committee Approval



T.C.
HACETTEPE ÜNİVERSİTESİ
Rektörlük

Tarih: 29/11/2019

Sayı:

E-35853172-101.02.02-00000881066



0000881066

Sayı : 35853172-101.02.02
Konu : Kübra YETİŞ (Etik Komisyon İzni)

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 05.11.2019 tarihli ve 51944218-101.02.02/00000851495 sayılı yazı.

Enstitünüz Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Bilim Dalı yüksek lisans öğrencilerinden **Kübra YETİŞ**'in **Dr. Öğr. Üyesi İsmail Fırat ALTAY** danışmanlığında yürüttüğü “**Üstbilişsel Eğitimin Yabancı Dil Olarak İngilizce Öğrenenlerin Dinleme Anlama Becerisi Üzerine Etkisi/The Impact of Metacognitive Instruction on EFL Learners’ Listening Comprehension Skill**” başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun **19 Kasım 2019** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini saygılarımla rica ederim.

e-imzalıdır

Prof. Dr. Rahime Meral NOHUTCU
Rektör Yardımcısı

Evrakın elektronik imzalı suretine <https://www.turkiye.gov.tr/hu-ebys> adresinden 5f19e825-f0f2-46c0-ac9e-a4082b2417d1 kodu ile erişebilirsiniz.
Bu belge 5070 sayılı Elektronik İmza Kanunu'na uygun olarak Güvenli Elektronik İmza ile imzalanmıştır.

Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara
Telefon:0 (312) 305 3001-3002 Faks:0 (312) 311 9992 E-posta:yazimd@hacettepe.edu.tr İnternet
Adresi: www.hacettepe.edu.tr

Sevda TOPA



APPENDIX H: Declaration of Ethical Conduct

I hereby declare that...

- I have prepared this thesis in accordance with the thesis writing guidelines of the Graduate School of Educational Sciences of Hacettepe University;
- all information and documents in the thesis/dissertation have been obtained in accordance with academic regulations;
- all audio visual and written information and results have been presented in compliance with scientific and ethical standards;
- in case of using other people's work, related studies have been cited in accordance with scientific and ethical standards;
- all cited studies have been fully and decently referenced and included in the list of References;
- I did not do any distortion and/or manipulation on the data set,
- and **NO** part of this work was presented as a part of any other thesis study at this or any other university.

22/01/2021
[Signature]
Kübra YETİŞ

APPENDIX-I: Thesis/Dissertation Originality Report

04/01/2021

HACETTEPE UNIVERSITY
Graduate School of Educational Sciences
To The Department of Foreign Language Education

Thesis Title: The Impact of Metacognitive Instruction on EFL Learners' Listening Comprehension Skill

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defence	Similarity Index	Submission ID
03/01/2021	202	378483	22/01/2021	10%	1481328772

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

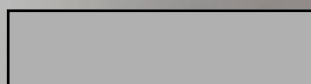
Name Lastname: Kübra YETİŞ
Student No.: N17123106
Department: Department of Foreign Language Education
Program: English Language Teaching MA Thesis Program
Status: ☒ Masters ☐ Ph.D. ☐ Integrated Ph.D.



Signature

Kübra YETİŞ

ADVISOR APPROVAL



(Assist. Prof. Dr. Tsmail Firat ALTAY)

APPENDIX-J: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına ilişkin Yönerge" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- o Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- o Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren ... ay ertelenmiştir. ⁽²⁾
- o Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

22.01.2021

11.02.2021

Kübra YETİŞ

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına ilişkin Yönerge"

- (1) Madde 6. 1. Lisansüstü teze ilişkin patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tez erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metodların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3 şahıslara veya kurumlara haksız kazanç, imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tez erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

* Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

