E-ISSN: 2981-1686

Journal of Studies in Language Learning and Teaching
www.jsllt.yazd.ac.ir
JSLLT, Vol. 2, No. 1, 36–48.

Paper Type: Original Article

Comparative Effects of AI-Assisted vs. Teacher-Assisted Collaborative Listening on EFL Learners' Self-Efficacy and Metacognitive Awareness

Hassan Alizadehmahmoudalilo[1]

[1] Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran;

Received: 19 August, 2025 Revised: 29 October, 2025 Accepted: 28 November, 2025

Abstract

Listening comprehension remains one of the most challenging skills for EFL learners, often hindered by rapid speech and unfamiliar vocabulary. However, little is known about how learners' self-efficacy and metacognitive awareness respond to different modes of instructional mediation. To do this, the present research sought to determine the comparative impacts of AI-assisted collaborative listening (AI-CL) to teacher-assisted collaborative listening (T-CL) on listening self-efficacy and metacognitive awareness among Iranian EFL learners. The study employed a quasi-experimental pretest-posttest control group design. The participants were selected from three intact undergraduate (BA-level) TEFL classes at Tabriz Islamic Azad University, where Persian serves as the official language of instruction. After homogenization using the Cambridge Listening Key English Test, the final sample included 60 upper-intermediate EFL learners (aged 18-24), both male and female, who were native speakers of Azari. They were randomly assigned to three groups, including the AI-CL group, T-CL group, and control group. The Metacognitive Awareness Listening Questionnaire and Second Language Listening Self-Efficacy Questionnaire were administered to collect the data, which were analyzed using ANCOVA. The results indicated that AI-CL and T-CL had equal effect in increasing metacognitive awareness, and neither had a significant short-term effect on self-efficacy. This study highlighted that AI can be used as a helpful guidance tool, yet teachers are still necessary to hold motivational and affective aspects of language learning.

Keywords: Artificial intelligence, AI-assisted collaborative listening, Teacher-assisted collaborative listening, Listening self-efficacy, Listening metacognitive awareness.

I | Introduction

Listening comprehension is central to effective communication and academic success, yet it remains the most neglected component of language instruction (Mansouri & Graham, 2025). Many EFL learners struggle with listening due to rapid speech, unfamiliar vocabulary, and limited exposure to authentic input (Shamsi & Bozorgian, 2024). A prevailing misconception is that listening proficiency develops naturally without explicit strategy training (Abd Latip et al., 2022). However, this research underscores the necessity of deliberate strategy use, guided by two pivotal constructs including listening self-efficacy and metacognitive awareness.

Listening self-efficacy refers to learners' belief in their capacity to comprehend spoken language and perform listening tasks successfully (Xu et al., 2021). As posited by Bandura (1997), higher self-efficacy enhances motivation, persistence, and strategy use, ultimately leading to better comprehension outcomes. Metacognitive awareness, conversely, involves the ability to plan, monitor, and evaluate one's listening processes (Shamsi & Bozorgian, 2024). Metacognitive strategies, such as prediction,

Corresponding Author: hassan.alizadehmahmoudalilo@iau.ac.ir
10.22034/jsllt.2025.23543.1092

monitoring, and self-evaluation, have been shown to strengthen concentration and understanding (Mansouri & Graham, 2025). Learners who exhibit both confidence and strategic control are more adept at processing complex listening input (Rubin, 1994; Vandergrift & Goh, 2012).

To foster these skills, many educators employ collaborative listening, in which peers jointly interpret, discuss, and negotiate meaning from listening texts. Such interaction promotes metacognitive awareness and mutual scaffolding, allowing less proficient listeners to learn from more skilled peers (Cross, 2018; Shamsi & Bozorgian, 2024). When guided by instructors, known as T-CL, teachers model effective strategies, prompt reflection, and provide feedback that reinforces learners' self-efficacy within the framework of the social cognitive theory (Bandura, 1997). Yet, implementing such scaffolding can be time-consuming and difficult to sustain in large classes (Vandergrift & Goh, 2012).

Recent advances in AI offer new avenues for supporting listening development. AI-driven technologies such as automated speech recognition (ASR), chatbots, and intelligent tutoring systems deliver real-time feedback and adaptive scaffolding that can enhance comprehension and reduce anxiety (Bashori et al., 2021; Shadiev et al., 2024; Tsai, 2023). Evidence suggests that AI-assisted listening tools improve learners' self-confidence, strategy awareness, and motivation (Abdellatif et al., 2024; Xiao, 2025). Furthermore, AI-mediated collaboration, where learners interact with both peers and intelligent agents, has been found to increase engagement and satisfaction after the COVID-19 pandemic (Toboula & Martinien, 2023). Nonetheless, scholars caution that, while AI may effectively address surface-level linguistic feedback, it lacks the depth of human mediation required for affective and motivational support (Luo et al., 2025).

Despite these advances, empirical comparisons between AI-CL and T-CL remain scarce. Most AI-based studies focus on isolated skills, e.g., ASR tasks or chatbot dialogues, rather than collaborative listening contexts, leaving the impact of AI on learners' self-efficacy and metacognitive development largely unexplored (Abdellatif et al., 2024). This gap limits the understanding of how human and AI mediation differentially influence the cognitive and affective dimensions of listening.

Grounded in the social cognitive theory (Bandura, 1997), the present study addresses this gap by comparing the effects of AI-CL and T-CL on EFL learners' listening self-efficacy and metacognitive awareness. Consequently, the current research aims to clarify whether AI can emulate or complement the motivational and strategic functions traditionally fulfilled by teachers, thereby informing both theory and practice in technology-enhanced language learning and EFL listening pedagogy.

II. Review of the related literature

This study explores how AI-CL and T-CL influence EFL learners' listening self-efficacy and metacognitive awareness. The issue lies at the intersection of technology-enhanced language learning, pedagogical design, and learner psychology, grounded in Bandura's (1997) social cognitive theory. This theory posits that learning emerges through the reciprocal interaction of personal factors, behavior and environment with self-efficacy, and beliefs in one's capabilities, serving as a central determinant of motivation and performance (Zhang et al., 2025). In language learning, teachers and AI systems represent different environmental mediators. While teachers provide vicarious experiences and social persuasion, AI offers immediate feedback and autonomous interaction (Liu et al., 2024).

Listening self-efficacy and metacognitive awareness constitute the study's key variables. Self-efficacy drives learners' engagement and persistence in listening tasks (Calaguas & Consunji, 2022), whereas metacognitive awareness enables them to plan, monitor, and evaluate comprehension effectively (Vandergrift & Goh, 2012). According to Vandergrift's Metacognitive Pedagogical Cycle, activities such as prediction and reflection before, during, and after listening enhance understanding and focus (Pei et al., 2023). However, while metacognitive training often improves comprehension, it does not consistently increase self-efficacy, indicating that confidence may depend on social or affective reinforcement in addition to strategy awareness (Pei et al., 2023).

Collaborative listening unites these constructs by combining peer interaction with metacognitive strategy practice. Learners engaged in group listening tasks negotiate meaning, articulate strategies, and support each other's understanding, thereby promoting both awareness and self-regulation (Cross, 2018; Shamsi & Bozorgian, 2024). Empirical studies with Iranian EFL learners reveal that metacognitive instruction paired with dialogic collaboration enhances comprehension significantly more than isolated training (Fakhri &

Bozorgian, 2022). Such findings underscore that social interaction amplifies metacognitive development by making strategic reasoning explicit.

The integration of technology further expands collaborative learning opportunities. Multimedia tools, such as captioned videos and podcasts, provide authentic and multimodal input, improving comprehension, motivation, and vocabulary while lowering anxiety (Jones, 2006; Musayeva et al., 2025). Within this framework, AI-CL represents an advanced evolution of technology-mediated pedagogy, offering adaptive feedback and individualized scaffolding (Bashori et al., 2021).

Empirical evidence supports the affective and cognitive benefits of AI in listening contexts. Xiao (2025) found that AI-based speech recognition enhanced listening comprehension and reduced anxiety among EFL learners. Likewise, Elsayed et al. (2024) reported that combining an AI-assisted test-taking tool with teacher guidance decreased learner anxiety and demotivation while increasing academic achievement. This finding directly relates to the current study, as it illustrates the complementary relationship between AI and teacher mediation. AI provides individualized feedback and adaptive support, whereas teachers reinforce motivation and confidence, both essential elements of self-efficacy as conceptualized by the social cognitive theory.

Further research corroborates the broader impact of AI on learner psychology. Zhang et al. (2025) demonstrated that EFL learners' AI literacy indirectly influenced their willingness to communicate through enhanced self-efficacy. Similarly, Lai (2025) found that AI-assisted writing improved learners' confidence and resilience, even though listening was not directly examined. These studies collectively indicate that AI-mediated instruction can foster motivation and self-belief when pedagogically integrated with human facilitation.

Despite these promising findings, empirical evidence comparing AI and human facilitation in collaborative listening remains limited. The existing research has largely examined AI in isolated skills, such as ASR-based listening exercises or chatbot interactions, rather than in group-based listening contexts. Moreover, while many studies report comprehension gains, few have explored how these environments shape learners' self-efficacy and metacognitive awareness simultaneously. Addressing this gap, the present study compares AI-CL and T-CL to determine how distinct mediators, human and AI, differentially influence the cognitive and affective dimensions of listening within authentic EFL settings.

III. Aim of the study

This study explicitly addresses the unexplored comparison of AI-assisted versus teacher-assisted collaborative listening. By experimentally examining the effects of these two instructional approaches on EFL learners' listening self-efficacy and metacognitive awareness, we aim to highlight both the novelty and necessity of this research. To our knowledge, no prior study has investigated this specific juxtaposition in the listening domain. The findings will be relevant to both theories of language acquisition and practical decisions in curriculum development. The research seeks to determine whether classrooms should use AI to guide group listening activities or they expect teachers to be at the forefront of collaborative work. In either case, understanding which approach better builds learner confidence and strategy will have broad implications for contemporary EFL instruction. Therefore, this research study proposes the following research questions in order to guide the investigation:

Is there any statistically significant difference between the listening self-efficacy of EFL learners engaged in AI-assisted collaborative listening and teacher-assisted collaborative listening?

Is there any statistically significant difference between the listening metacognitive awareness of EFL learners engaged in AI-assisted collaborative listening and teacher-assisted collaborative listening?

IV. Methodology

This quasi-experimental research had a pre-test-post-test-control group design, so it required three groups, two experimental groups and a control group. The experimental groups were given treatment using innovative collaborative methodologies, and the control group was instructed conventionally. Thus, the dependent variables in the scope of this study are listening self-efficacy and listening metacognitive awareness, while AI-assisted collaborative listening and teacher-assisted collaborative listening serve as the independent variables.

4.1. Participants

The participants were drawn from three undergraduate (BA-level) classes in the TEFL program at Tabriz Islamic Azad University, each consisting of more than 35 students. The final sample was 60 upper-intermediate EFL learners, both male and female, with an age range of 18 to 24. They were native speakers of Azari and were recruited from the same curriculum. Although their first language was Azari, Persian served as the official language of instruction at the university. The standardization across the three instructional contexts, AI-CL, T-CL, and the control one, ensured equivalence in weekly contact hours, assessment procedures, and instructional materials, thereby maintaining methodological consistency and minimizing variability. This technique ensured that study subjects had similar academic backgrounds. The population was recruited using a convenience sampling procedure due to the feasibility constraints associated with accessibility and interest in being sampled. This method ensured that the participants would reflect the general population of upper-intermediate EFL students of the university since they were subjected to comparable educational backgrounds, standards of language mastery, and learning levels. To guarantee the integrity and uniformity of the participants, a listening proficiency test was administered before the commencement of the primary research. After homogenizing, the participants were then non-randomly assigned to three groups including two experimental groups (AI-assisted collaborative listening group AI-CLG and teacher-assisted collaborative listening (T-CLG)) and a control group (CG). Originally, after conducting the proficiency test, there were some extra homogeneous students in each class, but, to equalize their numbers, they were excluded from the main study. Those participants were not put into the treatment stage but received routine training by the university and were not subject to any experimental manipulations or tests. In order to remove the confounding variables of teachers, all the groups were instructed by the same teacher.

4.2. Instruments

To gather the data needed for the study, the researcher applied the following instruments at various stages of the study.

4.2.1. Cambridge Key English Listening Test

In order to achieve initial homogeneity among the participants in terms of English listening proficiency, the Listening component of the Cambridge Key English Test (KET), officially known as the A2 Key examination, was administered at the beginning of the research. As a standardized proficiency test, the KET has been internationally recognized as reliable and valid evidence, developed and validated by Cambridge Assessment English. Despite the entire assessment of KET in terms of reading, writing, listening, and speaking, only the listening part was used in this study since it is consistent with the main interest of the research, which is mainly on the listening performance. The KET Listening test aims to evaluate understanding of small spoken texts in real situations. In this study, the 2020 sample paper was used, which has a duration of about 35 minutes (including 6 minutes of transfer time) and comprises four sections and 25 questions. The test items are multiple-choice questions, gap-filling, and short-answer questions, which have one mark each. The validity of the test is that it is scored objectively and reliably based on the correct answers. It has been demonstrated that the KET Listening section possesses good psychometric properties. The reliability curves in the Cambridge English Technical Manual provide internal consistency coefficients of 0.87 or greater on the listening component, and the standard error of measurement is small enough that consistent classification decisions can be made at the A2 level. The construct validity was justified by means of profuse item trialing, so it shows association with the Common European Framework of Reference for Languages (CEFR) descriptions of A2 listening skills (e.g., understanding simple, everyday information, recognizing the key points in brief talks and announcements). The content validity was determined following a systematic test specification procedure, which specifies a representative sample of listening functions, topics, and text types applicable to daily use in English.

4.2.2. Second Language Listening Self-Efficacy Questionnaire

This study used the Second Language Listening Self-Efficacy Questionnaire (SLLSQ) developed by Kassem (2015) to assess the confidence of the participants in their capability to perform second language listening activities. The instrument is composed of 40 items that are rated on a 5-point Likert scale with the five dimensions of progress, observational comparison, physiological states, strategic awareness, and challenge. These dimensions reflect how learners evaluate their progress, compare their efforts with those of their peers, regulate their emotions during listening, apply strategies, and solve challenging tasks.

Psychometric evidence has already been shown on the SLLSQ, with Golchi (2012) recording a Cronbach's alpha of 0.92, demonstrating a high internal consistency. In this study, the overall Cronbach's alpha coefficient was found to be 0.82, further demonstrating the internal consistency of the SLLSQ. To apply in the Iranian EFL context, the questionnaire was translated into Persian and checked by a panel of experts in applied linguistics to ensure the linguistic correctness as well as the content validity. A pilot test was also done with a sample of EFL learners close to the main participants. The purpose of the pilot was to verify the clarity, linguistic accuracy, and contextual relevance of the Persian-translated items before administering the questionnaire to the main participants. Furthermore, such steps made sure that the construct validity of the original instrument was still presented in the adapted one, which was also applicable and understandable to the participants. To cover cognitive and affective performance of listening, SLLSQ was given a comprehensive picture of the self-beliefs of learners, which is why SLLSQ was useful in this research in evaluating the impact of various interventions of collaborative listening.

4.2.3. Metacognitive Awareness Listening Questionnaire

The present study utilized the 21-item Metacognitive Awareness Listening Questionnaire (MALQ) created by Vandergrift et al. (2006), which was used to quantify the five elements of metacognitive knowledge in listening planning and evaluation, problem solving, directed attention, mental translation, and person knowledge. These categories reflect on the strategies learners use prior to, during, and after completion of listening exercises, hence providing a global criterion of metacognitive involvement. The items were measured on a 6-point Likert-type scale that ranged from strongly disagree to strongly agree. No neutral midpoint was allowed in an attempt to draw a clear response on each point. The MALQ has been extensively used in second language listening and has shown close positive psychometric properties. According to Vandergrift et al. (2006), the reliability coefficients of the five factors were found satisfactory, which has since been replicated by other researchers in other EFL environments. Rahimi and Abedi (2014) also validated the Persian form of the questionnaire in Iran, with Cronbach's alpha coefficient of 0.76, suggesting a favorable internal consistency in line with research purposes. In this study, the overall Cronbach's alpha coefficient was 0.87. To guarantee content validity, the Persian version was checked by three experts in applied linguistics, and reliability was also tested during a pilot administration among a similar target group of learners, confirming its assumption via a reasonable alpha value being recorded in each of the subscales. Such steps made the instrument culturally tolerant and psychometrically applicable in the target group. Recording the strategy use and awareness of learners in a systematized and approved way, the MALQ offered critical insights into the effects of collaborative listening interventions on the metacognitive aspects of listening growth of the participants.

4.3. Data collection procedures

This study was carried out in some steps to obtain accuracy, reliability, and consistency in answering the research questions. The procedure commenced by administering the listening KET to three undergraduate (BA-level) classes in the TEFL program during the academic year of 2024-2025. This standardized test was employed to determine the participants' listening level prior to the intervention and to homogenize the final sample. The test was given in formal conditions, rated objectively, and statistically analyzed. The students whose scores were clustered near the mean of the classes were retained, but the students whose scores were clustered in the extreme were discarded. This process made the results across the groups comparable, and the final sample consisted of 60 upper-intermediate EFL learners, where 20 students were chosen from each classroom.

After the homogenization phase, the 60 participants were assigned to three groups in line with the quasi-experimental design of the study. The first experimental group was the AI-CLG, the second was the T-CLG, and the third was a CG that continued with conventional listening instruction. To reduce the possible confounding impact of teacher-related factors, all the three groups were taught by the same teacher. This was done to ensure that the variation observed in the outcomes could be attributed to the teaching method.

Before the main data collection, two pilot studies were carried out at Tabriz Payam-e Noor University to validate the Persian versions of the research instruments. The pilots involved 15 upper-intermediate EFL learners comparable to, but not included in, the main sample. The purpose was to ensure the clarity, linguistic accuracy, and contextual appropriateness of the items in both the SLLSQ and MALQ. The

results confirmed the reliability of the instruments, with acceptable Cronbach's alpha values across all the subscales.

Following the successful pilot phase, the finalized questionnaires were administered to the main participants under standardized conditions. The SLLSQ and MALQ were applied to compare the perception of learners regarding listening self-efficacy and metacognitive awareness, respectively. Both of the instruments had been translated into Persian and validated in the Iranian EFL context to provide clarity and accuracy, forming a baseline measurement that could be compared with the posttest results.

The treatment phase was carried out over a six-week instructional cycle during the spring semester of 2025. Each group received two 90-minute sessions per week, resulting in a total of 12 instructional sessions. The AI-CLG engaged in the collaborative listening tasks mediated by an AI tool, ASR, which provided scaffolding, automated feedback, and adaptive support as learners worked in pairs or groups. On the other hand, there were parallel collaborative listening activities upon which the T-CLG participated with a teacher guiding, modeling, and correcting. In contrast, the CG got regular listening instructions provided by the university, which focused on grand studies and individual understanding exercises, not on organized collaboration or AI-assisted learning. Notably, the groups were exposed to identical listening materials, and an equivalent time of instruction was applied, thus holding the groups to the comparison.

The MALQ and SLLSQ were re-administered to all the groups as posttests after the treatment period. These instruments were also performed under the same conditions as pretests, thus ensuring consistency in the data collection. This enabled measuring the change in listening self-efficacy and metacognitive awareness that could be ascribed to the varied instructional methods.

Lastly, all the data were loaded in SPSS version 30. Before performing the inferential statistics, accuracy, outliers, and missing items have been screened in the dataset. The Internal consistency of the questionnaires was analyzed through Cronbach's alpha, and the assumptions of homogeneity of variance and normality were tested on the basis of descriptive statistics and the Levene Test. After the data screening ensured that the dataset was suitable, ANCOVA was done to investigate the outcomes of AI-CL and T-CL on self-efficacy and metacognitive awareness in learners, directly answering the research questions.

V. Results

Prior to the main analysis, descriptive and homogeneity tests were conducted to ensure that the groups were comparable in their initial levels of listening self-efficacy and metacognitive awareness. Table 1 summarizes the descriptive statistics for the SLLSQ and MALQ across the three groups at both pretest and posttest stages.

Table 1. Descriptive statistics of the test scores across the groups

						95% confidence Interval for mean			
		N	Mean	SD	SE	Lower Bound	Upper Bound	Minimum	Maximum
Pre-SLLSQ	AI-CLG	20	73.65	3.923	.877	71.81	75.48	67.00	80.00
	T-CLG	20	74.80	2.930	.655	73.42	76.17	69.00	79.00
	CG	20	73.70	2.848	.637	72.36	75.03	69.00	79.00
Post-SLLSQ	AI-CLG	20	85.70	5.322	1.190	83.20	88.19	73.00	94.00
	T-CLG	20	81.95	7.472	1.670	78.45	85.44	69.00	96.00
	CG	20	82.10	8.077	1.806	78.31	85.88	69.00	96.00
Pre-MALQ	AI-CLG	20	50.80	3.318	.741	49.24	52.35	45.00	57.00
	T-CLG	20	48.25	3.823	.854	46.46	50.03	43.00	57.00
	CG	20	51.60	3.299	.737	50.05	53.14	44.00	57.00
Post-MALQ	AI-CLG	20	100.10	5.821	1.301	97.37	102.82	90.00	112.00
	T-CLG	20	102.65	5.733	1.282	99.96	105.33	92.00	114.00
	CG	20	54.00	3.684	.823	52.27	55.72	47.00	59.00

As displayed in Table 1, the three groups showed similar pretest means on both instruments, confirming initial homogeneity. The posttest results revealed noticeable improvements in both experimental groups, with the AI-CLG and T-CLG outperforming the control group on both SLLSQ and MALQ measures. To verify the assumption of equal variances, Levene's Test of Homogeneity of Variance was performed for all measures (Table 2).

Table 2. Test of homogeneity of variances for the test scores

	Levene statistic	df1	df2	Sig.
Pre-SLLSQ	1.844	2	57	.167
Post-SLLSQ	2.927	2	57	.062
Pre-MALQ	.559	2	57	.575
Post-MALQ	1.370	2	57	.262

As Table 2 demonstrates, none of the Levene's statistics were significant at the .05 level ($p > .05$), confirming the homogeneity of variances assumption for all the variables. Therefore, the data satisfied the preliminary conditions required for conducting the subsequent ANCOVA analyses.

5.1. Results for the first research question

An ANCOVA was performed to determine whether there was any statistically significant difference between the listening self-efficacy of the EFL learners engaged in AI-assisted collaborative listening and teacher-assisted collaborative listening. Table 3 shows the ANCOVA results for the SLLSQ scores in the three groups: AI-CLG, T-CLG, and the CG.

Table 3. ANCOVA results: Effects of AI-CL and T-CL on post-SLLSQ scores

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	286.656[a]	3	95.552	1.958	.131	.095
Intercept	1447.287	1	1447.287	29.660	.000	.346
Pre-SLLSQ	106.356	1	106.356	2.180	.145	.037
group	156.397	2	78.199	1.603	.210	.054
Error	2732.594	56	48.796			
Total	418853.000	60				
Corrected total	3019.250	59				

a. R Squared = .095 (Adjusted R Squared = .046)

As Table 3 represents, the overall model was not statistically significant, $F(3, 56) = 1.958$, $p = .131$, partial $\eta^2 = .095$, suggesting that only about 9.5% of the variance in post-test scores was explained by the predictors included in the model. The covariate, pre-SLLSQ, did not exert a significant influence on the post-test outcomes, $F(1, 56) = 2.180$, $p = .145$, partial $\eta^2 = .037$, which indicates that prior differences in self-efficacy did not meaningfully affect post-test performance. Similarly, the effect of group membership (AI-CL, T-CL, or control) on post-test self-efficacy was not significant, $F(2, 56) = 1.603$, $p = .210$, partial $\eta^2 = .054$, implying that neither AI-assisted nor teacher-assisted collaborative listening led to statistically higher self-efficacy gains compared to the control condition. Although the intercept was significant, $F(1, 56) = 29.660$, $p < .001$, partial $\eta^2 = .346$, this reflects the overall level of self-efficacy rather than the treatment effects. Taken together, these results suggest that the experimental interventions did not produce significant differences in listening self-efficacy across groups, despite a small-to-moderate proportion of variance being explained by the model.

5.2. Results for the second research question

An ANCOVA was performed to determine whether there was any statistically significant difference between the listening metacognitive awareness of the EFL learners engaged in AI-CL and T-CL. Table 4 outlines these analyses.

Table 4. ANCOVA results: Effects of AI-CL and T-CL on post-MALQ scores

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	30031.329[a]	3	10010.443	377.434	.000	.953
Intercept	2619.669	1	2619.669	98.772	.000	.638
Pre-MALQ	41.096	1	41.096	1.549	.218	.027
group	27108.288	2	13554.144	511.045	.000	.948
Error	1485.254	56	26.522			
Total	470987.000	60				
Corrected total	31516.583	59				

a. R Squared = .953 (Adjusted R Squared = .950)

According to Table 4, the overall model was highly significant, $F(3, 56) = 377.434$, $p < .001$, partial $\eta^2 = .953$, indicating that approximately 95.3% of the variance in post-MALQ scores was explained by the predictors. The covariate, pre-MALQ, did not significantly predict the post-test scores, $F(1, 56) = 1.549$, $p = .218$, partial $\eta^2 = .027$, suggesting that the initial differences in metacognitive awareness did not strongly influence the outcomes. However, group membership exerted a profound and statistically significant effect on the post-MALQ scores, $F(2, 56) = 511.045$, $p < .001$, partial $\eta^2 = .948$, reflecting an extremely large effect size. This result demonstrates that both AI-CL and T-CL interventions substantially improved the learners' metacognitive awareness compared to the control group, with the effect accounting for nearly all of the explained variance in the model. The significant intercept, $F(1, 56) = 98.772$, $p < .001$, partial $\eta^2 = .638$, reflects the overall high level of the post-treatment scores across conditions. Table 5 outlines the results of the Tukey HSD post hoc test, which compares the mean differences between groups.

Table 5. Tukey HSD pairwise comparisons for the post-MALQ scores

					95% confidence interval for difference[b]	
(I) group	(J) group	MD (I-J)	SE	Sig.[b]	Lower bound	Upper bound
AI-CLG	T-CLG	-1.929	1.703	.786	-6.133	2.274
	CG	45.905*	1.636	.000	41.867	49.943
T-CLG	AI-CLG	1.929	1.703	.786	-2.274	6.133
	CG	47.835*	1.755	.000	43.502	52.167
CG	AI-CLG	-45.905*	1.636	.000	-49.943	-41.867
	T-CLG	-47.835*	1.755	.000	-52.167	-43.502

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

As shown in Table 5, the analysis revealed no statistically significant difference between the AI-CLG and the T-CLG, $MD = -1.929$, $SE = 1.703$, $p = .786$, indicating that both interventions produced comparable improvements in metacognitive awareness. However, both experimental groups outperformed the control group by a substantial margin. Specifically, the AI-CLG scored significantly higher than the control group, $MD = 45.905$, $SE = 1.636$, $p < .001$, 95% CI [41.867, 49.943], and T-CLG also scored significantly higher than the control group, $MD = 47.835$, $SE = 1.755$, $p < .001$, 95% CI [43.502, 52.167]. These findings suggest that, while AI-CL and T-CL were equally effective in enhancing the learners' metacognitive awareness, both approaches provided dramatic advantages over conventional instruction.

VI. Discussion

This research evaluated the comparative impacts of AI-CL and T-CL on the listening self-efficacy and metacognitive awareness of Iranian EFL learners. The results indicated that neither AI-CL nor T-CL produced significant effects on listening self-efficacy, but both experimental groups showed substantial gains in metacognitive awareness, with no significant difference between them. These outcomes provide empirical insight into how collaborative learning mediated by humans or AI influences the key

psychological and strategic dimensions of listening within the framework of the social cognitive theory (Bandura, 1997).

The fact that the listening self-efficacy did not increase significantly suggests that learners' confidence in their listening ability may require longer-term interventions to develop. According to the social cognitive theory, self-efficacy evolves gradually through mastery experiences, vicarious observation, social persuasion, and emotional regulation (Bandura, 1997). In the present study, although the collaborative tasks were structured and strategy-focused, the relatively brief treatment period might not have provided enough sustained mastery or positive reinforcement to modify learners' deeply held efficacy beliefs. Moreover, as self-efficacy is also shaped by affective states and social comparison (Calaguas & Consunji, 2022), the emphasis on cognitive and strategic practice may have limited opportunities for the type of interpersonal feedback and affective support that strengthen confidence. This theoretical lens explains why metacognitive awareness increased substantially while self-efficacy remained stable.

Conversely, the results for metacognitive awareness were noteworthy. Both AI-CL and T-CL yielded significant improvement, demonstrating that collaborative listening can effectively enhance learners' awareness and control of listening strategies. This aligns with the emphasis of the social cognitive theory on self-regulation as a central component of motivated learning behavior, as well as with prior research indicating that strategy-based collaborative instruction fosters planning, monitoring, and evaluation skills (Vandergrift & Goh, 2012; Cross, 2018). In particular, AI mediation may reinforce the environmental dimension of reciprocal determinism, providing learners with continuous feedback that promotes self-observation and regulation (Liu et al., 2024). AI helps to achieve these benefits through instant feedback, adaptive task sequencing, and repeated exposure matched to learners' proficiency levels (Shadiev et al., 2024). AI systems can also facilitate metacognitive regulation by assessing learner performance in real time, by displaying areas of error, possible alternative strategies, and encouraging reflection. In that regard, AI does not just act as a content provider, but also provides a scaffold for the strategic activities of the learners in a responsive, consistent, and scalable manner.

The abovementioned finding that emphasis AI systems can facilitate metacognitive regulation derives from both the observed quantitative gains in the current study and previous empirical evidence. In this study, the learners in the AI-CL condition improved markedly in metacognitive awareness, suggesting that the system's adaptive feedback helped them to identify errors and refine strategies. Comparable findings were reported by Xiao (2025), who observed that AI-driven speech recognition tools enhanced comprehension and lowered anxiety, thereby promoting reflection and strategic adjustment. Similarly, Lai (2025) demonstrated that AI-assisted writing improved students' self-belief and resilience, though its effects on listening self-efficacy were limited, a pattern consistent with the current results, in which AI heightened strategy awareness rather than confidence. Thus, AI appears to operate as a metacognitive scaffold rather than an affective catalyst, supporting learners' monitoring and evaluation processes while leaving emotional and motivational dimensions less affected. In that regard, AI does not merely act as a content provider but functions as a responsive and scalable facilitator of strategic learning, consistent with the self-regulatory mechanisms outlined in the social cognitive theory.

Nevertheless, the social cognitive theory also clarifies why teacher-led collaboration remains vital. In contrast to AI, human instructors can provide social persuasion and vicarious experience, two of the most influential sources of efficacy development (Bandura, 1997). Teachers have the potential to build positive classroom environments, recognize the feelings of learners, and offer motivation that develops resilience to challenge (Calaguas & Consunji, 2022). Studies have demonstrated that constructive, empathetic, and dialogic feedback provided by the teacher raises confidence in and propensity to risk-taking of learners on the task of listening (Abdellatif et al., 2024; Zhang et al., 2025). In addition, teachers introduce cultural awareness, pedagogical decision-making, and the capacity to transfer strategy training into a genuine communicative situation, elements that the AI systems are unable to cover to a great extent. Such attributes justify the relevance of teacher-mediated interventions, specifically to facilitate self-efficacy, although AI can offer similar support to facilitate metacognitive awareness.

Like other empirical studies, the present research suffers from limitations. The short six-week duration of the treatment probably constrained opportunities for mastery experience and social persuasion, key mechanisms of the social cognitive theory responsible for long-term efficacy change. The recruited

sample was restricted to upper-intermediate Iranian EFL learners whose proficiency was confirmed via the Cambridge Key English Test at the A2 CEFR level in a single university, limiting generalizability. Reliance on self-report instruments, despite their strong psychometric properties, may have introduced bias. Finally, the study compared AI- and teacher-assisted collaborative listening separately, without exploring the blended models that may capitalize on both human and technological affordances. Future research should employ longer interventions, include more diverse populations, combine self-report with performance-based measures, and investigate hybrid teacher-AI approaches alongside affective factors such as anxiety and motivation.

Overall, the findings demonstrate that, while AI- and teacher-mediated collaboration both foster metacognitive awareness, the enhancement of self-efficacy requires sustained human interaction, affective scaffolding, and authentic social persuasion, as the core principles of the social cognitive theory. In line with the social cognitive theory, AI can strengthen self-regulation and strategy use, but the teacher remains central to developing confident, motivated, and autonomous listeners. The results, thus, affirm that AI may serve as a potential pedagogical complement, yet the human educator remains irreplaceable in shaping the holistic growth of EFL learners.

VII. Conclusion and implications

The findings of this study, which revealed a nuanced pattern of outcomes, contribute to a deeper understanding of how technological and human mediation shape the multidimensional nature of listening development. Specifically, the results illustrate that AI-based support enhances learners' metacognitive regulation through adaptive feedback and repeated exposure, while teacher mediation strengthens the motivational and affective dimensions that sustain self-efficacy. From the perspective of the social cognitive theory (Bandura, 1997), this distinction highlights that, while AI can effectively operationalize environmental scaffolding and promote self-regulatory behaviors, it cannot replicate the social persuasion and emotional reinforcement that teachers provide.

The study, thereby, advances current knowledge on the role of technology in listening pedagogy by empirically showing that AI functions as a powerful catalyst for strategic awareness and self-monitoring, but not as a substitute for the interpersonal processes that nurture confidence and resilience. This helps clarify ongoing debates in technology-enhanced language learning regarding the extent to which AI can replace or complement human instruction.

Self-efficacy, a construct grounded in long-term mastery experience, social comparison, and affective feedback (Bandura, 1997; Calaguas & Consunji, 2022), appears less responsive to short-term interventions such as the six-week treatment implemented in this study. In contrast, metacognitive awareness, which depends on explicit guidance and task-based practice, responded strongly to both AI- and teacher-mediated collaborative listening. These findings underscore that strategic listening skills can be effectively cultivated through structured, feedback-rich collaboration, while motivational and affective dimensions call for extended human interaction and reinforcement.

Thus, the study meaningfully contributes to our understanding of the pedagogical potential of technology in EFL listening instruction. AI can complement teachers by strengthening learners' metacognitive control and strategy use, but the development of enduring confidence and self-efficacy still relies on teacher's sustained presence and affective engagement. This balanced perspective justifies the broader implications of the study for designing integrated, hybrid models of AI-assisted and teacher-led listening pedagogy, aimed at fostering both cognitive and emotional dimensions of language learning success.

The pedagogical implications of such findings point to the necessity of a blended approach to incorporating human and technological mediation. Although AI can be used effectively to improve metacognitive awareness by using scalable and individualized strategy practice, teachers are still needed to assess the motivational, emotional, and situational factors that support the self-efficacy of learners. It means that AI cannot be positioned as a replacement but as an additional assistant with a role specifically tailored to the needs of learners and tightly intertwined with lessons administered by teachers. To illustrate, routine drills and personalized feedback might be handled by AI, allowing teachers to concentrate on building learner confidence, reducing listening anxiety, and linking strategies to useful communicative

practice. This combination embraces the fullest potential of both forms of support so that learners can acquire both strategic awareness and the confidence they need to apply strategies effectively in real-world listening.

In sum, this study highlights the power of collaborative listening to enhance metacognitive awareness and the continuous need of teachers to support self-efficacy. Although a valuable ally, AI-enhanced instruction should be perceived as supportive and flexible, not a primary partner of sorts but an essential auxiliary to the educator. An integrated model that gives AI a chance to strengthen the strategy training and the teacher to promote motivation and confidence as the means to achieve the objective is the most encouraging avenue to the growth of comprehensive, confident, and strategy-oriented EFL listeners.

Declaration of generative AI and AI-assisted technologies: During the preparation of this work, the author used https://chat.z.ai/ to report tables in the Results Section. After using this tool, the author reviewed and edited the content as needed. So, he takes full responsibility for the published content.

AUTHORS' BIOGRAPHIES

Dr. Hassan Alizadehmahmoudalilo earned his Ph.D. in TEFL from the Islamic Azad University of Tabriz. As an adjunct professor, he is presently engaged in teaching TEFL courses at the same university, while also serving as an appointed teacher in the Ministry of Education. His scholarly contributions include the publication of over 10 articles in various international and national journals, reflecting his research interests in alternative assessment and language teaching methodologies integrating innovative technologies.

REFERENCES

Abd Latip, N. A., Din, W. A., & Swanto, S. (2022). Exploring the relationship between listening self-efficacy and metacognitive awareness of listening strategies. *International Journal of Education, Psychology and Counselling, 7*(48), 492–500. https://doi.org/10.35631/IJEPC.748037

Abdellatif, M. S., Alshehri, M. A., Alshehri, H. A., Hafez, W. E., Gafar, M. G., & Lamouchi, A. (2024). I am all ears: Listening exams with AI and its traces on foreign language learners' mindsets, self-competence, resilience, and listening improvement. *Language Testing in Asia*, *14*(1), 1–54. https://doi.org/10.1186/s40468-024-00329-6

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

Bashori, M., Van Hout, R., Strik, H., & Cucchiarini, C. (2021). Effects of ASR-based websites on EFL learners' vocabulary, speaking anxiety, and language enjoyment. *System*, *99*(1), 102–116. https://doi.org/10.1016/j.system.2021.102496

Calaguas, N. P., & Consunji, P. M. P. (2022). A structural equation model predicting adults' online learning self-efficacy. *Education and Information Technologies*, *27*(5), 6233–6249. https://doi.org/10.1007/s10639-021-10871-y

Cross, J. (2018). Collaborative listening: Teaching listening pedagogical perspectives on listening. *The TESOL Encyclopedia of English Language Teaching, 3*(2), 1–6. https://doi.org/10.1002/9781118784235.eelt0598

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House Publishing Group.

Elsayed, A. M., Kholikov, A., Abdullayeva, I., Al-Farouni, M., & Wodajo, M. R. (2024). Teacher support in AI-assisted exams: An experimental study to inspect the effects on demotivation, anxiety management in exams, L2 learning experience, and academic success. *Language Testing in Asia*, *14*(1), 1–53. https://doi.org/10.1186/s40468-024-00328-7

Fakhri, A. E., & Bozorgian, H. (2022). Gender, metacognitive intervention, and dialogic interaction: EFL multimedia listening. *System, 104*(1), 1–13. https://doi.org/10.1016/j.system.2021.102709

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

Jones, L. C. (2006). Effects of collaboration and multimedia annotations on vocabulary learning and listening comprehension. *CALICO Journal, 24*(1), 33–58. https://doi.org/10.1558/cj.v24i1.33-58

Kassem, H. M. (2015). The relationship between listening strategies used by Egyptian EFL college sophomores and their listening comprehension and self-efficacy. *English Language Teaching*, *8*(2), 153-169. http://dx.doi.org/10.5539/elt.v8n2p153

Lai, Z. C. C. (2025). The impact of AI-assisted blended learning on writing efficacy and resilience. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, *15*(1), 1–21. https://doi.org/10.4018/IJCALLT.377174

Liu, H., Chen, B., Li, X., & Zhou, X. (2024). Exploring the predictive role of self-efficacy in engagement among EFL teachers in online teaching: The mediation of buoyancy. *The Asia-Pacific Education Researcher*, *33*(4), 879–888. https://doi.org/10.1007/s40299-024-00816-7

Liu, Q., Liu, J., Cai, J., & Zhang, Z. (2020). The relationship between domain-and task-specific self-efficacy and mathematical problem posing: A large-scale study of eighth-grade students in China. *Educational Studies in Mathematics*, *105*(3), 407–431. https://doi.org/10.1007/s10649-020-09977-w

Luo, M., Hu, X., & Zhong, C. (2025). The collaboration of AI and teachers in feedback provision and its impact on EFL learners' argumentative writing. *Education and Information Technologies*, *9*(2), 1–21. http://dx.doi.org/10.1007/s10639-025-13488-7

Mansouri, K., & Graham, S. (2025). Self-regulation in L2 listening: The role of teacher and learner self-efficacy and the mediating influence of metacognition. *System*, *129*(1), 1–16. https://doi.org/10.1016/j.system.2025.103598

Musayeva, N., Sabirova, G., Xadjikurbanova, G., Yuldasheva, S., Akhmadaliyev, A., & Djumabaeva, J. (2025). Developing listening comprehension using podcasts in second language education: A systematic overview of recent pedagogical innovations (2019–2025). *Forum for Linguistic Studies, 7*(8), 482–497. https://doi.org/10.30564/fls.v7i8.10287

Pei, T., Suwanthep, J., & Lu, H. (2023). The effect of self-directed online metacognitive listening practice on Chinese EFL learners' listening ability, metacognition, and self-efficacy. *Frontiers in Psychology*, *14*(1), 128–159. https://doi.org/10.3389/fpsyg.2023.1285059

Rahimi, M., & Abedi, S. (2014). The relationship between listening self-efficacy and metacognitive awareness of listening strategies. *Procedia-Social and Behavioral Sciences*, *98*(1), 1454–1460. http://dx.doi.org/10.1016/j.sbspro.2014.03.565

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, *78*(2), 199–221. https://doi.org/10.1111/j.1540-4781.1994.tb02034.x

Shadiev, R., Feng, Y., Zhussupova, R., & Altinay, F. (2024). Effects of speech-enabled corrective feedback technology on EFL speaking skills, anxiety, and confidence. *Computer Assisted Language Learning*, *13*(1) 1–37. https://doi.org/10.1080/09588221.2024.2430761

Shamsi, E., & Bozorgian, H. (2024). Collaborative listening using multimedia through metacognitive instruction: A case study with less-skilled and more-skilled EFL learners. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(1), 1–25. https://doi.org/10.1186/s40862-023-00248-8

Toboula, Z., & Martinien, C. (2023). Enhancing post-pandemic EFL education by leveraging immersive, NLP-driven, AI-based tools that promote collaboration and interactivity within an educational approach. *International Journal of Education (IJE)*, *11*(1), 63–80. Retrieved from: https://ssrn.com/abstract=4512111

Tsai, S. C. (2023). Learning with mobile augmented reality and automatic speech recognition-based materials for English listening and speaking skills: Effectiveness and perceptions of non-English major

English as a foreign language students. *Journal of Educational Computing Research*, *61*(2), 444–465. https://doi.org/10.1177/07356331221111203

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

Vandergrift, L., Goh, C., Mareschal, C., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire (MALQ): Development and validation. *Language Learning, 56*(1), 431–462. https://doi.org/10.1111/j.1467-9922.2006.00373.x

Xiao, Y. (2025). The impact of AI-driven speech recognition on EFL listening comprehension, flow experience, and anxiety: A randomized controlled trial. *Humanities and Social Sciences Communications*, *12*(1), 1–14. https://doi.org/10.1057/s41599-025-04672-8

Xu, J., Fan, J., & Luo, K. (2021). Exploring L2 listening instruction, self-efficacy, and strategy use: A mediation analysis. *Frontiers in Psychology*, *12*(1), 758–767. https://doi.org/10.3389/fpsyg.2021.758757

Zhang, Q., Nie, H., Fan, J., & Liu, H. (2025). Exploring the dynamics of artificial intelligence literacy on English as a foreign language learners' willingness to communicate: The critical mediating roles of artificial intelligence learning self-efficacy and classroom anxiety. *Behavioral Sciences (Basel, Switzerland)*, *15*(4), 523–548. https://doi.org/10.3390/bs15040523.