

Metacognitive Instruction for Metaphor Learning: Effects on Metacognitive Awareness, Self-Efficacy, and Learning Performance among Chinese EFL Learners

Baiwen Niu

Universiti Sains Malaysia

Malini Ganapathy

`malinik@usm.my`

Universiti Sains Malaysia

Dongxue Han

Universiti Sains Malaysia

Jianping Chen

Ningbo University

Research Article

Keywords: Teaching Metaphor, Metacognitive Awareness, Self-efficacy, Metaphor learning Performance, Chinese EFL Learners

Posted Date: September 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10727616/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

This study examines the effects of metacognitive instruction (MI) on Chinese EFL learners' metacognitive awareness of metaphor learning (MAML), self-efficacy in metaphor learning (SEML), and metaphor learning performance (MLP) among Chinese EFL undergraduates. It aims to extend current understanding of metaphor pedagogy by investigating how metacognitive support facilitates learners' regulation of cognitive awareness, self-efficacy, and metaphor learning outcomes. A quasi-experimental design was adopted with 70 Chinese EFL undergraduates assigned to an experimental group (EG) receiving metacognitive instruction and a control group (CG) following conventional Presentation–Practice–Production (PPP) instruction. Pretest and post-test data were collected through MAIML, QMSEML, and MLPT. Independent-samples t-tests and one-way ANCOVA were conducted to examine post-intervention differences between the two groups. The results reveal that the EG significantly outperformed the CG in MAML, SEML, and MLP after the intervention. The metacognitive intervention was particularly effective in promoting learners' metacognitive knowledge and regulation, with notable gains in conditional knowledge and debugging skills. Learners in the EG also demonstrated stronger SEML, especially academic self-efficacy. Regarding MLP, the EG showed significant improvement in complex metaphor construction. These findings suggest that effective metaphor learning involves not only acquiring figurative expressions but also developing the ability to understand and regulate the cognitive processes involved in interpreting and using metaphorical language. This study contributes to metaphor pedagogy by demonstrating the value of metacognitive instruction for supporting learners' metacognitive awareness, self-efficacy, and metaphor learning performance. It extends research on metacognitive instruction beyond conventional EFL skills by highlighting the roles of metacognitive awareness and self-efficacy in metaphor learning. The findings provide theoretical and pedagogical implications for designing more effective EFL metaphor instruction.

Introduction

Metaphor is an essential element of language, acting both as a linguistic device and a cognitive mechanism that shapes individuals' conceptual understanding and communication (Bolognesi & Horvat, 2022; Low, 1988; Lakoff & Johnson, 1980). In EFL domain, however, metaphors learning can be challenging because learners need to move beyond literal meanings, recognise underlying conceptual relationships, and interpret expressions in context. Conceptual metaphor theory (CMT) has consequently been explored as a pedagogical foundation for helping learners understand and use metaphor more effectively (Wang & Yuan, 2024; Zhou et al., 2022; Chen & Lai, 2015). Yet, much of this work has centred on the conceptual development of metaphorical competence, with comparatively less empirical attention to learners' actual performance in acquiring, applying, and evaluating metaphors. This limitation is particularly relevant to Chinese EFL learners, for whom metaphor learning may require not only access to conceptual knowledge but also the ability to regulate how that knowledge is understood and applied. As Su et al. (2023) noted, planning, monitoring, and evaluation are central metacognitive processes in learners' linguistic and conceptual development. From this perspective, successful

metaphor learning may depend on learners' capacity to reflect on and regulate their own learning rather than on exposure to metaphorical concepts alone (Seidenberg & Bernstein, 1988).

MI has attracted growing attention in EFL research because it encourages learners to become more aware of how they learn and to regulate their learning processes more deliberately (Flavell, 1979; Schraw, 1998). Through planning, monitoring, and evaluation, MI can support strategic regulation and has been associated with improved learning across different EFL domains (Kim & Whitehead, 2025; Ruan, 2014; Teng, 2020, 2021). Its potential relevance to metaphor learning lies in the close fit between these metacognitive processes and the demands of working with metaphorical language. At the same time, metacognitive guidance may extend beyond cognitive regulation to learners' motivational beliefs and perceptions of their own capabilities (Sasson & Tifferet, 2025). Research in EFL contexts has linked MI with stronger self-efficacy through the development of strategic knowledge, self-regulatory awareness, and perceptions of successful learning experiences (Milliner & Dimoski, 2024; Bai et al., 2022; Rahimirad & Zare-ee, 2015; Taghani & Razavi, 2022). Whether these benefits extend to self-efficacy in metaphor learning, however, is less clear, particularly given the conceptual and interpretive demands involved in understanding and producing metaphorical expressions.

Given these limitations, examining MI in metaphor learning provides an opportunity to consider awareness of cognitive regulation, motivational beliefs, and learning outcomes within a less explored EFL domain. Although MI has been associated with reflective awareness, strategic regulation, and language learning more broadly (Kim & Whitehead, 2025; Teng, 2022), its contribution to metaphor learning performance has received considerably less empirical attention. This gap is particularly relevant in Chinese university EFL classrooms, where PPP remains a prevalent teacher-led approach involving explicit instruction, controlled practice, and language production. As a conventional instructional approach, PPP provides a useful comparison for examining whether MI can offer additional benefits by engaging learners more actively in planning, monitoring, and evaluating their metaphor learning. The present study therefore adopts a quasi-experimental design to compare MI and PPP in terms of Chinese EFL undergraduates' MAML, SEML, and MLP. By examining these outcomes together, the study aims to clarify the potential of a metacognitive approach to support learners' regulation, perceived capability, and learning outcomes in a cognitively demanding area of language learning. Specifically, the study addresses the following research questions (RQs):

RQ1. Is there a significant difference in metacognitive awareness of metaphor learning between Chinese EFL undergraduates receiving metacognitive instruction and those receiving conventional PPP instruction?

RQ2. Is there a significant difference in self-efficacy in metaphor learning between Chinese EFL undergraduates receiving metacognitive instruction and those receiving conventional PPP instruction?

RQ3. Is there a significant difference in metaphor learning performance between Chinese EFL undergraduates receiving metacognitive instruction and those receiving conventional PPP instruction?

Drawing on previous research, this study assumes that MI may generate greater benefits than conventional PPP instruction in boosting learners' MAML, SEML, and MLP. The following hypotheses are therefore proposed:

H1: Chinese EFL undergraduates receiving metacognitive instruction will demonstrate significantly higher metacognitive awareness of metaphor learning than those receiving conventional PPP instruction.

H2: Chinese EFL undergraduates receiving metacognitive instruction will demonstrate significantly higher self-efficacy in metaphor learning than those receiving conventional PPP instruction.

H3: Chinese EFL undergraduates receiving metacognitive instruction will demonstrate significantly higher metaphor learning performance than those receiving conventional PPP instruction.

Literature Review

Metacognitive Instruction and Metacognitive Awareness of EFL Learning

Metacognition was first introduced by Flavell (1979) to describe learners' awareness and control of their own thinking processes. Building on this perspective, Schraw and Dennison (1994) conceptualized metacognitive awareness as comprising two interrelated dimensions: metacognitive knowledge and metacognitive regulation. Metacognitive knowledge refers to learners' understanding of cognition, including declarative, procedural, and conditional knowledge; whereas metacognitive regulation concerns the processes used to manage learning through planning, information management, monitoring, debugging, and evaluation. Drawing on this framework, MI emphasizes explicit guidance that helps learners plan, monitor, solve problems, and evaluate their learning while engaging in authentic communicative activities (Goh & Vandergrift, 2021).

A substantial body of research has examined the role of MI across different areas of EFL and ESL learning. Existing studies have mainly focused on listening (e.g., Kim and Whitehead, 2025; Pei et al., 2023; Robillos & Bustos, 2022), and writing (e.g., Wang et al., 2025; Teng, 2020, 2021, 2022; Lee & Mak, 2018; Ruan, 2014). Although these studies differ in their instructional settings and research designs, they generally suggest that MI facilitates language learning. For example, Wang et al. (2025) found that metacognitive intervention combined with collaborative writing improved lexical diversity, accuracy, and fluency, while Teng (2020, 2022) reported gains in metacognitive knowledge, regulation, and writing performance. Lee and Mak (2018) and Ruan (2014) similarly showed that explicit metacognitive training enhanced awareness of when and how to use learning strategies. More recently, Kim and Whitehead (2025) found that MI increased EFL learners' awareness of oral corrective feedback. Together, these findings suggest that MI may foster learners' ability to comprehend and produce language.

Evidence also indicates that the effectiveness of MI is not uniform across learning contexts. Teng (2021), for example, showed that learners who received metacognitive prompts outperformed those who did not, while metacognitive regulation emerged as the strongest predictor of learning outcomes. In contrast, Pei et al. (2023) reported that online metacognitive listening instruction improved listening performance but did not produce significant gains in either metacognitive awareness or self-efficacy. This inconsistency suggests that improved language performance does not necessarily reflect corresponding development in learners' metacognitive competence. Despite the growing evidence supporting MI in language education, relatively little attention has been paid to its application in metaphor learning. Understanding and producing metaphors require learners to monitor their interpretation, evaluate alternative meanings, and adjust their reasoning during language processing (Seidenberg & Bernstein, 1988). These cognitive demands suggest that MI may provide an effective framework for supporting metaphor learning by encouraging learners to regulate both comprehension and production processes.

Metacognitive Instruction and Self-efficacy in EFL Learning

Bandura (1997) defined self-efficacy as individuals' beliefs about their capability to organize and perform the actions required to achieve specific goals. Building on this framework, Self-efficacy has subsequently been conceptualized as a multidimensional construct that captures academic, emotional, and social dimensions (Muris, 2001; Won et al., 2024). Academic self-efficacy reflects learners' confidence in completing academic tasks (Bandura, 1997; Schunk, 1991); emotional self-efficacy concerns the regulation of negative emotions and the expression of positive emotions (Muris, 2001); and social self-efficacy refers to learners' confidence in their ability to build and maintain effective relationships with teachers and peers (Patrick et al., 2007). Although these dimensions have received attention across different educational contexts, limited research has investigated the effects of metacognitive instruction on SEML. Exploring this relationship may help explain how learners' confidence interacts with strategic behaviour and cognitive engagement during metaphor learning.

Recent research has provided further evidence of a positive relationship between metacognitive instruction and self-efficacy in EFL learning (e.g., Alfahad, 2025; Taghani & Razavi, 2022; Bai et al., 2022; Rahimirad & Zare-ee, 2015). In particular, both Alfahad (2025) and Rahimirad and Zare-ee (2015) found that metacognitive listening instruction enhanced EFL learners' listening self-efficacy, with Alfahad (2025) additionally reporting gains in metacognitive awareness and listening achievement. Similar effects have been observed in other instructional contexts. Taghani and Razavi (2022), for example, found that EFL learners receiving metacognitive training demonstrated significantly higher self-efficacy than those receiving conventional instruction. Using data from 1,954 secondary school students in Singapore, Bai et al. (2022) likewise identified a positive association between academic self-efficacy and both classroom engagement and metacognitive self-regulation, together with lower avoidance coping. Taken together, the literature indicates that metacognitive instruction is closely linked to learners' confidence in managing language-learning tasks. However, increased self-efficacy does not necessarily

correspond to improved learning outcomes. Milliner and Dimoski (2024), for instance, found that low-proficiency Japanese EFL learners reported greater listening confidence following metacognitive instruction but showed no significant improvement in listening performance. Thus, although existing evidence generally supports the role of MI in fostering self-efficacy, increased self-efficacy does not necessarily translate into improved language-learning performance, with its effects potentially varying across learner and instructional contexts. Against this background, research on metacognitive instruction has largely examined self-efficacy in conventional areas of EFL learning, such as listening and writing, while metaphor learning has received considerably less attention. Little is known about whether MI can also build learners' confidence in interpreting and using metaphorical language. Examining this issue would extend research on the motivational dimensions of MI to the specific context of EFL metaphor learning.

Metaphor Learning performance in EFL Context

Conceptual metaphor theory conceptualizes metaphor as a cognitive mechanism that shapes thought and language use rather than merely a rhetorical device (Lakoff & Johnson, 1980). From this perspective, successful metaphor learning involves recognizing the conceptual relations underlying metaphorical meanings. Metaphor awareness supports this process by helping learners identify conceptual schemas and preserve figurative expressions (Boers, 2000; Gutiérrez Pérez, 2018). Building on this perspective, Littlemore (2001) defined metaphorical competence as the ability to interpret, infer, produce, and process metaphorical language, which extends beyond awareness to the accurate, appropriate, and creative use of metaphors in communication (Zhou et al., 2022). Recent research further views L2 metaphorical competence as a multidimensional construct involving linguistic, conceptual, and metalinguistic processes, with metaphor knowledge largely emphasized (O'Reilly & Yan, 2025; Chen et al., 2025). In line with this perspective, the present study defines MLP as learners' ability to recognize, interpret, and produce metaphors appropriately across communicative contexts (Niu & Ganapathy, 2025). Previous studies have mainly examined metaphor use through English written production, including academic writing, standardized writing tasks, and classroom compositions (e.g., Saneie Moghadam & Ghafar Samar, 2020; Chen, 2020; Kathpalia & Carmel, 2011; Littlemore et al., 2014). It is noted that the existing approaches to metaphor assessment have largely focused on identifying and quantifying metaphor use, with comparatively less attention to learners' creativity, contextual integration, and linguistic accuracy. To capture these dimensions, the present study adopts a multidimensional framework adapted from Littlemore et al. (2014), which assesses metaphor categorization, contextual integration, creative originality, complex metaphor construction, and linguistic accuracy.

Despite the recognized importance of metaphor in L2 learning, systematic metaphor instruction remains relatively uncommon in EFL classrooms, partly because of constrained curricular resources, teacher preparation, and the cognitive demands of metaphor learning (Boussaid, 2024; MacArthur & Alejo-González, 2024; Littlemore & Low, 2006; Hu, 2019). Metaphor comprehension and production require learners to identify conceptual relations, monitor interpretations, and evaluate language choices in

context (Seidenberg & Bernstein, 1988; O'Reilly & Yan, 2025). These processes closely correspond to the metacognitive functions of planning, monitoring, and evaluation. Although explicit and strategy-based approaches have shown potential for improving metaphor learning among Chinese EFL learners (Wang & Yuan, 2024; Chen & Lai, 2015), the role of learners' metacognitive regulation in this process remains less clear. MI may be well suited to metaphor learning because it encourages learners to plan, monitor, and reflect on their learning, yet few studies have examined its contribution to MLP among Chinese EFL undergraduates.

Methodology

To address the research questions and objectives, the present study initially adopted an explanatory sequential mixed-methods design consisting of a quantitative phase followed by a qualitative phase. Due to the scope and word limitations of this article, only the quantitative findings are reported here, while the qualitative findings are presented in a separate publication.

Participants

This study collected data from a college in an Eastern Chinese university in the second semester of the 2023–2024 academic year. Ethical approval was granted by the institute, and informed permissions were obtained from all participants. The participants were sophomores and juniors who had studied English for an average of 10 years. They voluntarily participated in the study and were studying majors of art and engineering. Initially, 76 college students were recruited for the study using convenience sampling method. The final sample comprised 70 valid participants (31 males and 39 females; $M = 20.67$ years, range = 19–23), with 35 participants in each group. Six students were excluded due to absence, leave during the post-test, or incomplete or invalid responses. Before the intervention, the two groups did not differ significantly in MAML, SEML, or MLP, and participants reported limited opportunities for metaphor training both inside and outside the classroom.

Instruments

Metacognitive awareness inventory of metaphor learning

The Metacognitive Awareness Inventory of Metaphor Learning (MAIML) was administered to both the EG and CG before and after the intervention to assess participants' metacognitive awareness in metaphor learning. Adapted from the Metacognitive Awareness Inventory developed by Schraw and Dennison (1994), which has demonstrated strong reliability and validity (Teng, 2020), the instrument was contextualized for metaphor learning (i.e., On a regular basis, I ask myself if I'm achieving success in my efforts to learn metaphors (item 1)). The MAIML consists of 52 items distributed across eight components: (1) declarative knowledge(DK), (2) procedural knowledge(PK), (3) conditional knowledge(CK), (4) planning(PLA), (5) information management strategies(IMS), (6) monitoring(MON), (7) debugging strategies(DEB), and (8) evaluation(EVA). These components form two subscales, i.e.,

metacognitive knowledge (MK, sections 1–3) and metacognitive strategy (MS, sections 4–8). Responses were rated on a six-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree). Lower scores indicate weaker metacognitive awareness and perceived use of MK and MS, while higher scores signify enhanced awareness and affirmative participation. The original MAI had strong internal consistency ($\alpha = 0.88$ for each subscale; $\alpha = 0.93$ overall).

Questionnaire for Measuring Self-Efficacy of Metaphor Learning (QMSEML)

The SEML questionnaire was derived from the Questionnaire for Measuring Self-Efficacy in Youths (QMSY) created by Muris (2001). The QMSY consists of 24 items covering three subscales: emotional, social, and academic self-efficacy, with 8 questions in each subscale. The QMSY was contextualized to form the QMSEML (i.e., How well can you get others to help you when you get stuck on schoolwork of metaphor learning? (Question 1)). Responses were rated on a five-point Likert scale (1 = not at all, 5 = very well), where higher scores reflect stronger self-efficacy in metaphor acquisition. The instrument has adequate internal consistency (overall $\alpha = 0.88$; emotional $\alpha = 0.86$; social $\alpha = 0.85$; academic $\alpha = 0.88$).

Metaphor Learning Performance Test (MLPT)

The MLPT was conducted using a writing section with a theme: the choice after graduation, which aims to test students' proficiency to depict a picture, elucidate its meaning, and provide comments using metaphors. To examine participants' metaphor learning achievement, a summative Rubric for Writing Metaphor Assessment (RMWA) was created and validated after referring to Littlemore et al. (2014). The RMWA owns five subcategories, i.e., Contextual Organization and Support for Metaphors (COSM), Types and Clusters of Metaphors (TCM), Creativity and Originality of Metaphors (COM), Complexity and Combination of Metaphors (CCM), and Grammar, Spelling, and Sentences (GSS). Each subcategory is evaluated on a 0–4 score and the score of each participant ranges from 0 to 20 points. The RMWA significantly fills a critical gap in MLPT assessment.

Pilot study

The pilot study followed self-validation and expert review. Four language experts assessed the clarity, suitability, and readability of the MAIML, QMSEML, and MLPT and suggested revisions for the target participants. The pilot intervention was conducted from April to August 2023 with 31 EG and 25 CG EFL undergraduates similar to those in the main study. Both groups received eight weeks of instruction, with the EG following metacognitive instruction and the CG receiving PPP instruction. Post-test analyses revealed high internal consistency for MAIML (Cronbach's $\alpha = 0.947$) and QMSEML (Cronbach's $\alpha = 0.927$), with further adjustments applied to the MLPT rubric. Participants' feedback validated average completion duration of around 10 minutes for consent, 30 minutes for surveys, and 60 minutes for the MLPT. Overall, the pilot validated the instruments for extensive deployment.

Intervention procedures

The EG and CG had the same class schedule, teaching materials, and writing assignments, differing mainly in instructional approach. The EG was based on Schraw and Dennison's (1994) framework of metacognitive intervention and integrated into 15 instructional weeks to promote engagement, self-regulation, and self-efficacy in metaphor learning. The CG followed PPP instruction, with a focus on guided practice and the communicative use of metaphors without explicit metacognitive instruction. Both groups used *Metaphor: A Practical Introduction* (Kövecses, 2010) and researcher-developed materials suited to learners' proficiency levels. The 17-week intervention comprised 30 class hours, with two academic hours scheduled each week on weekends. Each weekly session had a 20-minute pre-learning exercise, a 50-minute instructional phase, and a 20-minute post-learning task. The first and final weeks were reserved for pre- and post-assessments, which included the MAIML, QMSEML, and a 300–400-word essay. The data collected were coded and analyzed using IBM SPSS Statistics Version 30.0. For the MLPT, two EFL raters with expertise in metaphor studies jointly established the scoring criteria and independently scored the essays. Scores from the 70 students' pre- and post-tests ($N = 700$) were then examined. Inter-rater reliability showed substantial agreement ($\kappa = 0.705$) and a significant positive correlation ($r = .833, p < .05$). Although one rater was more lenient, the scores showed good overall consistency.

Findings

Differences in Metacognitive Awareness of Metaphor Learning

For RQ1, the EG and CG showed no significant pre-test differences in overall MAML or any of its subcomponents (see Table 1). The EG ($M = 151.66, SD = 27.49$) and the CG ($M = 157.80, SD = 26.67$) did not differ significantly in overall MAML ($t(68) = -0.95, p = .346$). No significant between-group differences were found for MK, MS, or their respective sub-dimensions, including DK, PK, CK, PLA, IMS, MON, DEB, and EVA (all $p > .05$). These findings indicate similar baseline MAML levels across the two groups, which supports the subsequent comparison of post-test outcomes and treatment effects.

Table 1
Descriptive statistics and Independent Samples T-test for MAML Pretest Mean Scores Between Groups

	EG		CG		Mean Diff.	Std. Error Diff.	t	df	Sig. (2-tailed)
	M	SD	M	SD					
MAIML	151.66	27.49	157.80	26.67	-6.14	6.47	-0.95	68	.346
MK	48.63	8.69	50.29	8.93	-1.66	2.11	-0.79	68	.434
DK	22.03	4.28	22.83	4.12	-0.80	1.00	-0.80	68	.428
PK	11.26	2.32	11.49	2.12	-0.23	0.53	-0.43	68	.668
CK	15.34	2.49	15.97	3.02	-0.63	0.66	-0.95	68	.345
MS	103.03	18.90	107.51	17.87	-4.49	4.40	-1.02	68	.311
PLA	22.17	3.93	23.37	3.49	-1.20	0.89	-1.35	68	.181
IMS	27.83	5.07	28.49	4.77	-0.66	1.18	-0.56	68	.578
MON	20.03	4.16	20.77	3.82	-0.74	0.95	-0.78	68	.439
DEB	15.43	2.99	16.31	2.94	-0.89	0.71	-1.25	68	.216
EVA	17.57	3.50	18.57	3.56	-1.00	0.84	-1.19	68	.240

Table 2 reports the results of the independent samples t-test on the post-test MAML scores for the EG and the CG. Specifically, the EG had a higher mean score ($M = 220.80$, $SD = 28.88$) than the CG ($M = 201.49$, $SD = 27.84$), and the difference was significant ($t(68) = 2.85$, $p = .006$). Similar results were found across the two MAML dimensions, with the EG scoring significantly higher than the CG on both MK ($t(68) = 2.97$, $p = .004$) and MS ($t(68) = 2.76$, $p = .007$). Within MK, significant between-group differences appeared in DK, PK, and CK (all $p < .05$). Within MS, significant differences were identified in all five strategy components, including PLA, IMS, MON, DEB, and EVA (all $p < .05$). These results provided preliminary support for H1, as the EG scored significantly higher than the CG on overall MAML and all sub-dimensions. To further assess the treatment effect while controlling for pre-test MAML, an ANCOVA was conducted with post-test MAML as the dependent variable and instructional approach as the independent variable.

Table 2

Descriptive statistics and Independent Samples T-test for Mean MAML Post-test Scores Between Groups

	EG		CG		Mean Diff.	Std. Error Diff.	t	df	Sig. (2- tailed)
	M	SD	M	SD					
MAIML	220.80	28.88	201.49	27.84	19.31	6.78	2.85	68	.006
MK	71.49	9.47	64.97	8.86	6.51	2.19	2.97	68	.004
DK	33.49	4.59	30.63	4.27	2.86	1.06	2.70	68	.009
PK	16.97	2.38	15.51	2.27	1.46	0.56	2.62	68	.011
CK	21.03	2.71	18.83	2.77	2.20	0.66	3.36	68	.001
MS	149.31	19.56	136.51	19.28	12.80	4.64	2.76	68	.007
PLA	30.46	4.05	27.86	3.95	2.60	0.96	2.72	68	.008
IMS	41.60	5.63	38.51	5.46	3.09	1.33	2.33	68	.023
MON	29.51	3.65	26.86	3.93	2.66	0.91	2.93	68	.005
DEB	21.83	2.99	19.63	3.05	2.20	0.72	3.05	68	.003
EVA	25.91	3.64	23.66	3.77	2.26	0.89	2.55	68	.013

Before conducting the ANCOVA, the assumptions of normality and linearity were examined and met. However, Levene's test was significant ($F(1, 68) = 5.712, p = .020$), which indicated a slight departure from the homogeneity of variance assumption. Given this violation, a more conservative significance criterion was adopted, and the 99% confidence interval (CI) was reported for the group difference.

Table 3
Comparison of Adjusted Mean MAML Post-test Scores Between Groups

	Group	Cov.	Adjusted post-test			
			Mean	Mean Diff.	Std. Error	99% CI
MAML	Experimental	154.73	223.35	24.42	4.19	[13.32, 35.52]
	Control		198.93			
MK	Experimental	49.46	72.16	7.87	1.38	[4.22, 11.51]
	Control		64.29			
DK	Experimental	22.43	33.80	3.48	0.72	[1.57, 5.40]
	Control		30.32			
PK	Experimental	11.37	17.06	1.63	0.39	[0.61, 2.65]
	Control		15.43			
CK	Experimental	15.66	21.25	2.64	0.47	[1.41, 3.88]
	Control		18.61			
MS	Experimental	105.27	151.17	16.52	2.92	[8.76, 24.27]
	Control		134.66			
PLA	Experimental	22.77	30.92	3.52	0.69	[1.70, 5.34]
	Control		27.40			
IMS	Experimental	28.16	41.87	3.63	0.91	[1.22, 6.04]
	Control		38.24			
MON	Experimental	20.40	29.77	3.18	0.62	[1.53, 4.82]
	Control		26.60			
DEB	Experimental	15.87	22.16	2.87	0.49	[1.56, 4.18]
	Control		19.29			
EVA	Experimental	18.07	26.30	3.02	0.62	[1.38, 4.66]
	Control		23.28			

Table 4
Tests of Between-Subjects Effects for Adjusted Mean MAML Post-test Scores

	Source	Type III Sum of Squares	df	Mean Square	F	Sig. (2-tailed)	Partial Eta Squared
MAML	Group	10299.03	1	10299.03	34.05	< .001	.34
	Error	20267.07	67	302.49			
MK	Group	1073.36	1	1073.36	32.72	< .001	.33
	Error	2198.01	67	32.81			
DK	Group	210.14	1	210.14	23.22	< .001	.26
	Error	606.35	67	9.05			
PK	Group	46.44	1	46.44	17.98	< .001	.21
	Error	173.08	67	2.58			
CK	Group	120.69	1	120.69	32.26	< .001	.33
	Error	250.67	67	3.74			
MS	Group	4701.16	1	4701.16	31.90	< .001	.32
	Error	9872.77	67	147.36			
PLA	Group	210.88	1	210.88	26.27	< .001	.28
	Error	537.87	67	8.03			
IMS	Group	229.45	1	229.45	15.98	< .001	.19
	Error	961.90	67	14.36			
MON	Group	175.17	1	175.17	26.28	< .001	.28
	Error	446.69	67	6.67			
DEB	Group	140.65	1	140.65	33.71	< .001	.34
	Error	279.59	67	4.17			
EVA	Group	156.42	1	156.42	23.76	< .001	.26
	Error	441.14	67	6.58			

As shown in Table 3 and Table 4, the ANCOVA results showed a significant difference in post-test MAML scores between the two groups after pre-test scores were entered as a covariate. The adjusted mean score was higher in the EG ($M = 223.35$) than in the CG ($M = 198.93$). The group difference reached statistical significance ($F(1, 67) = 34.05, p < .001$), with a large effect size ($\eta^2p = .34$). For Metacognitive Knowledge, the EG ($M = 72.16$) also showed a significantly higher score than the CG ($M = 64.29$), $F(1, 67)$

= 32.72, $p < .001$, $\eta^2_p = .33$. Significant group effects were found across all MK subcomponents, i.e., DK, $F(1, 67) = 23.22$, $p < .001$, $\eta^2_p = .26$; PK, $F(1, 67) = 17.98$, $p < .001$, $\eta^2_p = .21$; and CK, $F(1, 67) = 32.26$, $p < .001$, $\eta^2_p = .33$. Similarly, for Metacognitive Strategies, the EG obtained significantly higher adjusted mean ($M = 151.17$) than the CG ($M = 134.66$), $F(1, 67) = 31.90$, $p < .001$, $\eta^2_p = .32$. The EG scored significantly higher across all five strategy dimensions, for instance, PLA, $F(1, 67) = 26.27$, $p < .001$, $\eta^2_p = .28$; IMS, $F(1, 67) = 15.98$, $p < .001$, $\eta^2_p = .19$; MON, $F(1, 67) = 26.28$, $p < .001$, $\eta^2_p = .28$; DEB, $F(1, 67) = 33.71$, $p < .001$, $\eta^2_p = .34$; and EVA, $F(1, 67) = 23.76$, $p < .001$, $\eta^2_p = .26$. These findings confirmed that the EG scored significantly higher than the CG, with substantial effect sizes across all MAML components.

Differences in Self-Efficacy in Metaphor Learning

Table 5 displays pre-intervention descriptive statistics for the SEML between groups. The EG attained a total SEML mean of 70.71 ($SD = 13.39$) whereas the CG generated a marginally higher total SEML mean of 74.31 ($SD = 11.40$), with no statistically significant difference observed ($t(68) = -1.21$, $p = .230$). The CG had slightly higher scores on ESEML ($p = .165$), SSEML ($p = .234$), and ASEML ($p = .334$), but none of these differences was statistically significant. These findings indicate similar baseline levels of SEML across the two groups before the intervention.

Table 5

Descriptive statistics and Independent Samples T-test for Mean SEML Pretest Scores Between Groups

	EG		CG		Mean Diff.	Std. Error Diff.	t	df	Sig. (2-tailed)
	M	SD	M	SD					
SEML	70.71	13.39	74.31	11.40	-3.60	2.97	-1.21	68	.230
ESEML	24.03	4.46	25.37	3.49	-1.34	0.96	-1.40	68	.165
SSEML	23.51	4.20	24.66	3.75	-1.14	0.95	-1.20	68	.234
ESEML	23.17	5.05	24.29	4.51	-1.11	1.14	-0.97	68	.334

By contrast, Table 6 reports independent-samples t-test outputs for post-intervention SEML mean scores after the instructional treatments. The EG obtained higher overall SEML scores ($M = 93.34$, $SD = 10.05$) against the CG ($M = 88.60$, $SD = 9.66$), and the difference was statistically significant ($t(68) = 2.01$, $p = .048$). The EG scored significantly higher on ASEML than the CG ($t(68) = 2.09$, $p = .040$). The differences in ESEML and SSEML approached, but did not reach the .05 significance level ($p = .066$ and $.075$, respectively). These results show a difference in SEML between the two groups following the intervention.

Table 6

Descriptive statistics and Independent Samples T-test for Mean SEML Post-test Scores Between Groups

	EG		CG		Mean Diff	Std. Error Diff.	t	df	Sig. (2-tailed)
	M	SD	M	SD					
SEML	93.34	10.05	88.60	9.66	4.74	2.36	2.01	68	.048
ESEML	31.00	3.50	29.51	3.15	1.49	0.80	1.87	68	.066
SSEML	30.66	3.34	29.26	3.13	1.40	0.77	1.81	68	.075
ESEML	31.69	3.76	29.83	3.67	1.86	0.89	2.09	68	.040

Subsequently, an ANCOVA compared post-test SEML scores between the EG and CG, with pre-test SEML scores as the covariate. The assumptions of normality, linearity, and homogeneity of variance were satisfied before the analysis. For SEML, the adjusted post-test means were 94.13 for the EG and 87.81 for the CG, as reported in Tables 7 and 8. The ANCOVA showed a significant group effect ($F(1, 67) = 9.94$, $p = .002$, $\eta^2p = .13$). Further ANCOVA tests revealed significant group differences across all three SEML subdimensions. The adjusted mean for ESEML was 31.25 in the EG and 29.26 in the CG ($F(1, 67) = 7.58$, $p = .008$, $\eta^2p = .10$). For SSEML, the EG had an adjusted mean of 30.89, whereas the CG had 29.02 ($F(1, 67) = 7.65$, $p = .007$, $\eta^2p = .10$). ASEML showed the largest group difference, with adjusted means of 31.92 for the EG and 29.59 for the CG ($F(1, 67) = 9.68$, $p = .003$, $\eta^2p = .13$). Thus, after adjustment for pre-intervention scores, the overall SEML measure and all three subdimensions showed significant group differences.

Table 7

Comparison of Adjusted Mean SEML Post-test Scores Between Groups

	Group	Cov.	Adjusted post-test			
			Mean	Mean Diff.	Std. Error	95% CI
SEML	Experimental	72.51	94.13	6.31	2.00	[2.32 ,10.31]
	Control		87.81			
ESEML	Experimental	24.70	31.25	1.99	0.72	[0.55, 3.44]
	Control		29.26			
SSEML	Experimental	24.09	30.89	1.87	0.68	[0.52,3.23]
	Control		29.02			
ASEML	Experimental	23.73	31.92	2.34	0.75	[0.84 ,3.83]
	Control		29.59			

Table 8
Tests of Between-Subjects Effects for Adjusted Mean SEML Post-test Scores

	Source	Type III Sum of Squares	df	Mean Square	F	Sig. (2-tailed)	Partial Eta Squared
SEML	Group	683.07	1	683.07	9.94	.002	.13
	Error	4602.37	67	68.69			
ESEML	Group	67.54	1	67.54	7.58	.008	.10
	Error	597.41	67	8.92			
SSEML	Group	60.18	1	60.18	7.65	.007	.10
	Error	527.12	67	7.87			
ASEML	Group	94.13	1	94.13	9.68	.003	.13
	Error	651.61	67	9.73			

Differences in Metaphor Learning Performance

Prior to the intervention, the EG ($M = 8.97$, $SD = 2.05$) and the CG ($M = 9.19$, $SD = 2.24$) showed no statistically significant divergence in total MLP ($t(68) = 0.43$, $p = .667$), as illustrated in Table 9. Likewise, comparisons of the MLP sub-scores showed no significant group differences, namely COSM ($t(68) = 0.93$, $p = .357$), TCM ($t(68) = 0.98$, $p = .329$), COM ($t(68) = -0.80$, $p = .424$), CCM ($t(68) = 0.19$, $p = .850$), and GSS ($t(68) = 0.39$, $p = .696$). All these baseline comparisons pointed to comparable metaphor-learning proficiency levels for the two groups before instructional work began.

Table 9
Descriptive statistics and Independent Samples T-test for Mean MLP Pretest Scores Between Groups

	EG		CG		Mean Diff.	Std. Error Diff.	t	df	Sig. (2-tailed)
	M	SD	M	SD					
MLP	8.97	2.05	9.19	2.24	0.21	0.50	0.43	68	.667
COSM	1.79	0.48	1.90	0.62	0.11	0.12	0.93	68	.357
TCM	2.34	0.51	2.46	0.53	0.11	0.12	0.98	68	.329
COM	1.61	0.55	1.53	0.53	-0.09	0.11	-0.80	68	.424
CCM	1.26	0.70	1.29	0.62	0.03	0.15	0.19	68	.850
GSS	1.97	0.48	2.01	0.53	0.04	0.11	0.39	68	.696

Table 10 presents the post-test MLP scores after the intervention. The mean score was 14.16 ($SD = 2.48$) for the EG and 13.34 ($SD = 2.58$) for the CG. The between-group difference was not significant ($t(68) =$

1.41, $p = .165$). Across the five MLP components, the EG had higher mean scores than the CG, but none of the between-group differences was statistically significant. The corresponding results were COSM, $t(68) = 0.86, p = .396$; TCM, $t(68) = 0.66, p = .512$; COM, $t(68) = 1.47, p = .147$; CCM, $t(68) = 1.39, p = .168$; and GSS, $t(68) = 1.64, p = .106$. These results did not show significant post-test differences between the two groups. An ANCOVA was therefore conducted to examine group differences after adjustment for pre-test MLP.

Table 10
Descriptive statistics and Independent Samples T-Test for Mean MLP Post-test Scores Between Groups

	EG		CG		Mean Diff.	Std. Error Diff.	t	df	Sig. (2-tailed)
	M	SD	M	SD					
MLP	14.16	2.48	13.34	2.58	0.81	0.58	1.41	68	.165
COSM	2.97	0.59	2.86	0.67	0.11	0.13	0.86	68	.396
TCM	3.26	0.58	3.17	0.59	0.09	0.13	0.66	68	.512
COM	2.61	0.60	2.43	0.67	0.19	0.13	1.47	68	.147
CCM	2.39	0.60	2.20	0.65	0.19	0.13	1.39	68	.168
GSS	2.93	0.71	2.69	0.60	0.24	0.15	1.64	68	.106

The assumptions for ANCOVA were met. As shown in Table 11 and Table 12, after adjustment for pre-test MLP, the EG had an adjusted mean of 14.27, compared with 13.23 for the CG (see). The ANCOVA yielded a significant group effect ($F(1, 67) = 16.01, p < .001, \eta^2p = .19$), which indicates a large in magnitude. Across all MLP subdimensions, the adjusted means were higher for the EG than for the CG. CCM showed a significant group effect ($F(1, 67) = 6.20, p = .02, \eta^2p = .09$). The remaining components also favored the EG, but none reached the .05 significance level. The corresponding results were COSM, $F(1, 67) = 3.31, p = .07, \eta^2p = .05$; TCM, $F(1, 67) = 3.73, p = .06, \eta^2p = .05$; COM, $F(1, 67) = 1.50, p = .23, \eta^2p = .02$; and GSS, $F(1, 67) = 3.88, p = .05, \eta^2p = .06$. Taken as a whole, significant group differences were found for total MLP and the CCM subdimension, whereas the other MLP subdimensions did not reach statistical significance.

Table 11
Comparison of Adjusted Mean MLP Post-test Scores Between Groups

	Group	Cov.	Adjusted post-test			
			Mean	Mean Diff.	Std. Error	95% CI
MLP	Experimental	9.08	14.27	1.04	0.26	[0.52, 1.56]
	Control		13.23			
COSM	Experimental	1.84	3.01	0.19	0.11	[-0.02, 0.40]
	Control		2.82			
TCM	Experimental	2.40	3.30	0.18	0.09	[-0.01, 0.36]
	Control		3.13			
COM	Experimental	1.57	2.58	0.12	0.10	[-0.08, 0.32]
	Control		2.46			
CCM	Experimental	1.27	2.40	0.21	0.08	[0.04, 0.37]
	Control		2.19			
GSS	Experimental	1.99	2.94	0.27	0.14	[-0.003, 0.54]
	Control		2.67			

Table 12
Tests of Between-Subjects Effects for Adjusted Mean MLP Post-test Scores

	Source	Type III Sum of Squares	Df	Mean Square	F	Sig. (2-tailed)	Partial Eta Squared
MLP	Group	18.82	1	18.82	16.01	< .001	.19
	Error	78.79	67	1.18			
COSM	Group	0.64	1	0.64	3.31	.07	.05
	Error	12.89	67	0.19			
TCM	Group	0.54	1	0.54	3.73	.06	.05
	Error	9.77	67	0.15			
COM	Group	0.25	1	0.25	1.50	.23	.02
	Error	11.17	67	0.17			
CCM	Group	0.74	1	0.74	6.20	.02	.09
	Error	8.00	67	0.12			
GSS	Group	1.25	1	1.25	3.88	.05	.06
	Error	21.52	67	0.32			

Discussion

Effect of MI on Metacognitive Awareness of Metaphor Learning

The results support RQ1 and H1, as learners in the MI group showed significantly higher levels of MAML than those in the PPP group. This difference may be attributed to the greater emphasis on learners' control. Through planning, monitoring, evaluation, and strategy adjustment, learners became more actively involved in managing their metaphor learning. The finding is consistent with Flavell's (1979) and Schraw and Dennison's (1994) view that metacognitive awareness develops through conscious control of learning. It also supports previous EFL research on the effects of metacognitive instruction (Kim and Whitehead, 2025; Teng, 2022, 2020; Lee & Mak, 2018; Ruan, 2014). The present study extends this evidence to metaphor learning, which involves more than knowledge of linguistic forms. Learners need to identify conceptual relationships, consider alternative interpretations, and select appropriate strategies for unfamiliar metaphors (Littlemore, 2009; Seidenberg & Bernstein, 1988). Metacognitive control may help learners manage these interpretive demands more effectively.

The metacognitive knowledge findings provide further insight into this effect. The EG outperformed the CG in declarative, procedural, and conditional knowledge, with the largest gain in conditional knowledge.

This result indicates stronger awareness of how and when strategies should be applied, rather than simply knowing which strategies are available. Such knowledge supports effective self-regulation (Schraw, 1998) and accords with Lee and Mak (2018) and Ruan (2014). It may be particularly valuable in metaphor learning, where learners need to adapt their approach to different conceptual relationships. Beyond MK, metacognitive regulation also improved across planning, monitoring, debugging, and evaluation. The gain in debugging is especially relevant because metaphor comprehension may require learners to revise an initial interpretation rather than retrieve a fixed rule or definition. This finding extends Hacker et al. (2009) by indicating that metacognitive regulation may support learning that requires both figurative and conceptual processing. Lastly, these findings extend MI research beyond the language skills that have received most attention, such as listening and writing (Wang et al., 2025; Teng, 2022, 2020). They further show that MI can support metaphor learning by helping learners manage its interpretive and strategic demands.

Effect of MI on Self-Efficacy in Metaphor Learning

The significant difference between the two groups in SEML provides support for H2 and answers RQ2. This result adds to previous evidence from EFL contexts (Rahimirad & Zare-ee, 2015; Taghani & Razavi, 2022) by extending it to metaphor learning area. Bandura's (1997) theory suggests that self-efficacy develops through successful experiences, support, and a sense of control over learning. In the present study, goal setting, reflection, and performance evaluation gave learners greater regulation over their learning and opportunities to recognise their progress. These experiences may have facilitated greater confidence in handling metaphor-related tasks. The finding is in line with Rahimirad and Zare-ee (2015) and Taghani and Razavi (2022), both of whom reported positive effects of metacognitive instruction on self-efficacy.

Across the three dimensions, academic self-efficacy showed the largest improvement. This result is plausible because metaphor learning requires learners to interpret abstract meanings and apply conceptual knowledge in new contexts. It also accords with Bai et al. (2022), who found that ASEML was related to learner engagement, metacognitive self-regulation, and adaptive learning behaviors. At the same time, ESEML and SSEML also improved significantly after pre-test differences were controlled, although the gains were smaller. This broader finding suggests that MI may influence learner confidence beyond academic performance. Additionally, in line with Muris (2001) and Won et al. (2024), this study confirms that reflective activities and peer interaction may have helped learners manage negative emotions and feel more confident in metaphor-related **tasks**. Beyond these dimension-specific changes, the findings also add to the discussion of self-efficacy and learning outcomes. For instance, Milliner and Dimoski (2024) found that MI increased learners' confidence in listening, but did not produce significant gains in listening outcome. In contrast, the present study found that MI significantly improved self-efficacy and performance in a structured metaphor-learning context. The contrasting results may reflect differences in instructional design, learner characteristics, and task demands. Overall, the results indicate that MI supports the development of SEML. By strengthening learners' perceived competence

and sense of control, metacognitive training provides motivational support for engaging with complex figurative language.

Effect of MI on Metaphor Learning Performance

The results show that MI significantly improved EFL learners' metaphor learning. Although the two groups performed similarly before the intervention, the EG achieved significantly higher post-test scores after controlling for pretest scores. This finding is consistent with previous studies indicating that MI can facilitate EFL learning outcomes (Wang et al., 2025; Alfahad, 2025; Pei et al., 2023; Teng, 2022). More importantly, it extends metaphor instruction research, which has largely focused on metaphor knowledge and explicit explanation (Wang & Yuan, 2024; Chen & Lai, 2015), by revealing that learners' active regulation of cognitive processes can also advance metaphor learning.

The improvement in MLP following the MI intervention can be interpreted from the perspective of CMT (Lakoff & Johnson, 1980). Metaphor learning requires learners to recognize conceptual relationships and construct meanings beyond literal expressions. The metacognitive activities may have deepened metaphor understanding and supported more appropriate use. This finding is consistent with previous arguments that highlights the roles of metaphor awareness and strategic processing in developing metaphorical competence (Gutiérrez Pérez, 2018; Littlemore, 2009; Boers, 2000). The subcomponent results further clarify these effects. The significant gain in CCM suggests that MI may facilitate higher-order metaphorical thinking and creative language use. This is in line with Littlemore (2009), who viewed metaphorical competence as involving both the comprehension and manipulation of metaphorical meanings. However, COSM, TCM, COM and GSS did not differ significantly between the groups. One possible explanation is that PPP instruction provides structured input and controlled practice, which may also support metaphor production to some extent (Anderson, 2017). In addition, it may also reflect the nature of original metaphor production, which requires sustained exposure to varied metaphorical input and sufficient opportunities for meaningful use (Boers, 2000). Creative and sufficient production of metaphors may therefore develop more gradually than metaphor comprehension. Taken together, the results confirm RQ3 and H3, with the metacognitive group showing better metaphor learning outcomes than the conventional instruction group.

Conclusion and Implication

This study provides evidence that MI can enhance metaphor learning among Chinese EFL undergraduates. Compared with conventional PPP instruction, the MI group achieved higher levels of MAML, SEML, and MLP. The MI was particularly effective in developing conditional knowledge, debugging skills, and academic self-efficacy, and it also improved performance in the complexity and combination of metaphor use. These findings suggest that effective metaphor learning requires not only knowledge of figurative language but also motivational support, such as self-efficacy, and the ability to regulate the cognitive processes involved in interpreting and using metaphorical language.

Beyond these findings, the study offers theoretical and pedagogical implications. Theoretically, it brings together metacognition theory, self-efficacy theory, and conceptual metaphor theory to explain how

cognitive regulation and self-efficacy may support metaphor learning. It also extends research on metacognitive instruction beyond conventional language skills (Pei et al., 2023; Teng, 2022; Wang et al., 2025) and highlights cognitive regulation and learner self-efficacy as relevant dimensions of metaphor pedagogy (Boers, 2000; Littlemore, 2009; Wang & Yuan, 2024). Pedagogically, teachers can incorporate planning, monitoring, debugging, and evaluation into metaphor activities while providing regular opportunities for reflection, guided practice, and successful task completion. Such an approach may help learners develop greater self-efficacy and autonomy when they encounter unfamiliar metaphorical meanings.

This study has several limitations. First, the sample was limited to Chinese EFL undergraduates, which may restrict the generalizability of the findings to other contexts. Future research could include learners at different proficiency levels and in other EFL settings. Second, MLP relied mainly on written tasks. Future studies could incorporate oral production, comprehension measures, and authentic communicative activities to provide a broader assessment of metaphorical competence. Third, although the study employed an explanatory sequential mixed-methods design, this article reports only the quantitative findings due to its scope. Future research could integrate qualitative evidence to provide a fuller account of how learners develop metacognitive awareness and self-efficacy during metaphor learning. Finally, the lack of a delayed post-test limits evidence of the long-term effects of the intervention. Future longitudinal studies could include delayed assessments to examine whether gains in MAML, SEML, and MLP are maintained over time.

Declarations

Ethical approval

The study was conducted in accordance with the principles of the Declaration of Helsinki. Participation was voluntary, and no identifiable personal data were collected. The study was reviewed and approved in accordance with the research ethics requirements of Universiti Sains Malaysia prior to data collection.

Informed consent

Informed consent was obtained from all participants involved in the study.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Author Contribution

Baiwen Niu conceived and designed the study, developed the methodology, conducted the investigation, curated and analyzed the data, performed the software-related work, and drafted the manuscript. Malini Ganapathy contributed to the methodology, supervised the study, and reviewed and edited the manuscript. Dongxue Han contributed to the investigation and visualization and performed the software-related work. Jianping Chen provided resources and contributed to the validation of the study. All authors read and approved the final manuscript.

Acknowledgement

We are grateful to the institute where the study was conducted and to all participants for their support, and we appreciate the editor and anonymous reviewers for their valuable feedback.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Alfahad, M. F. (2025). Implementing the metacognitive pedagogical cycle: Effects on EFL learners' listening achievement, metacognitive awareness, listening self-efficacy, and listening challenges. *Journal of Language Teaching and Research*, 16(3), 932–942. <https://doi.org/10.17507/jltr.1603.23>
2. Anderson, J. (2017). A potted history of PPP with the help of ELT journal. *ELT Journal*, 71(2), 218–227. <https://doi.org/10.1093/elt/ccw055>
3. Bai, B., Nie, Y., & Lee, A. N. (2022). Academic self-efficacy, task importance and interest: relations with English language learning in an Asian context. *Journal of Multilingual and Multicultural Development*, 43(5), 438–451. <https://doi.org/10.1080/01434632.2020.1746317>
4. Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
5. Boers, F. (2000). Enhancing metaphoric awareness in specialized reading. *English for Specific Purposes*, 19(2), 137–147. [https://doi.org/10.1016/s0889-4906\(98\)00017-9](https://doi.org/10.1016/s0889-4906(98)00017-9)
6. Bolognesi, M., & Horvat, A. W. (2022). *The metaphor compass: directions for metaphor research in language, cognition, communication, and creativity*. Routledge.
7. Boussaid, Y. (2024). The difficulties hindering EFL teachers from using metaphor-based teaching. *Australian Journal of Applied Linguistics*, 7(2), 1495. <https://doi.org/10.29140/ajal.v7n2.1495>
8. Chen, Y., & Lai, H. (2015). Developing EFL learners' metaphoric competence through cognitive-oriented methods. *International Review of Applied Linguistics in Language Teaching*, 53(4), 415–438.

<https://doi.org/10.1515/iral-2015-0019>

9. Chen, Y. C. (2020). Assessing the lexical richness of figurative expressions in Taiwanese EFL learners' writing. *Assessing Writing*, 43, 100414. <https://doi.org/10.1016/j.asw.2019.06.007>
10. Chen, Z., Guan, L., & Zhou, X. (2025). Is L2 learners' metaphorical competence essentially cognitive, linguistic, or personal?—A meta-analysis. *Journal of Intelligence*, 13(9), 117. <https://doi.org/10.3390/jintelligence13090117>
11. 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>
12. Goh, C. C., & Vandergrift, L. (2021). *Teaching and learning second language listening: metacognition in action*. Routledge.
13. Gutiérrez Pérez, R. (2018). The development of a metaphoric competence. A didactic proposal of educational innovation. *Innovation in Language Learning and Teaching*, 13(4), 331–357. <https://doi.org/10.1080/17501229.2018.1466891>
14. Hacker, D. J., Dunlosky, J., & Graesser, A. C. (2009). *Handbook of Metacognition in Education*. Routledge.
15. Hu, Z. (2019). The study of metaphoric competence in China. *Language and Semiotic Studies*, 5(3), 134–144. <https://doi.org/10.1515/lass-2019-050306>
16. Kathpalia, S. S., & Carmel, H. L. H. (2011). Metaphorical competence in ESL student writing. *RELC Journal*, 42(3), 273–290. <https://doi.org/10.1177/0033688211419379>
17. Kim, T., & Whitehead, G. E. (2025). Metacognitive Instruction and Oral Corrective Feedback: Impacts on Metacognitive Awareness and Speaking Accuracy in Learners of English as a Foreign Language. *RELC Journal*, 00336882251370760. <https://doi.org/10.1177/00336882251370760>
18. Kövecses, Z. (2010). *Metaphor: A practical introduction*. Oxford University Press.
19. Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
20. Lee, I., & Mak, P. (2018). Metacognition and metacognitive instruction in second language writing classrooms. *TESOL Quarterly*, 52(4), 721–724. <https://doi.org/10.1002/tesq.436>
21. Littlemore, J. (2001b). Metaphoric competence: A possible language learning strength of students with a holistic cognitive style? *TESOL Quarterly*, 35(3), 459–491. <https://doi.org/10.2307/3588031>
22. Littlemore, J. (2009). Applying cognitive linguistics to second language learning and teaching. *Palgrave Macmillan*. <https://doi.org/10.1057/9780230245259>
23. Littlemore, J., & Low, G. (2006). Metaphoric competence, second language learning, and communicative language ability. *Applied Linguistics*, 27(2), 268–294. <https://doi.org/10.1093/applin/aml004>
24. Littlemore, J., Krennmayr, T., Turner, J., & Turner, S. (2014). An investigation into metaphor use at different levels of second language writing. *Applied Linguistics*, 35(2), 117–144. <https://doi.org/10.1093/applin/amt004>

25. Low, G. D. (1988). On teaching metaphor. *Applied Linguistics*, 9(2), 125–147.
<https://doi.org/10.1093/applin/9.2.125>
26. MacArthur, F., & Alejo-González, R. (2024). Beyond idioms, the use of metaphor in ELF academic settings: A comprehensive review. *Journal of Pragmatics*, 219, 48–57.
<https://doi.org/10.1016/j.pragma.2023.11.002>
27. Milliner, B., & Dimoski, B. (2024). The effects of a metacognitive intervention on lower-proficiency EFL learners' listening comprehension and listening self-efficacy. *Language Teaching Research*, 28(2), 679–713. <https://doi.org/10.1177/13621688211004646>
28. Muris, P. (2001). A brief questionnaire for measuring self-efficacy in youths. *Journal of Psychopathology and Behavioral Assessment*, 23(3), 145–149.
<https://doi.org/10.1023/a:1010961119608>
29. Niu, B., & Ganapathy, M. (2025). Revisiting the assessment of metaphor learning performance through English language learners' English writings: A systematic review. *Open Journal of Modern Linguistics*, 15, 1–21. <http://dx.doi.org/10.4236/ojml.2025.151001>
30. O'Reilly, D., & Yan, L. (2025). Playing with second language metaphor: An exploration with advanced Chinese learners of English. *Applied Linguistics*, 46(1), 53–74.
<https://doi.org/10.1093/applin/amad067>
31. Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, 99(1), 83–98. <https://doi.org/10.1037/0022-0663.99.1.83>
32. 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, 1285059. <https://doi.org/10.3389/fpsyg.2023.1285059>
33. Rahimirad, M., & Zare-ee, A. (2015). Metacognitive strategy instruction as a means to improve listening self-efficacy among Iranian undergraduate learners of English. *International Journal of Instruction*, 8(1), 117–132.
34. Robillos, R. J., & Bustos, I. G. (2022). Learners' listening skill and metacognitive awareness through metacognitive strategy instruction with pedagogical cycle. *International Journal of Instruction*, 15(3), 393–412. <https://doi.org/10.29333/iji.2022.15322a>
35. Ruan, Z. (2014). Metacognitive awareness of EFL student writers in a Chinese ELT context. *Language Awareness*, 23(1–2), 76–91. <https://doi.org/10.1080/09658416.2013.863901>
36. Saneie Moghadam, M., & Ghafar Samar, R. (2020). Metaphor in second language academic writing. *Language Awareness*, 29(2), 1–16. <https://doi.org/10.1080/09658416.2020.1786577>
37. Sasson, I., & Tifferet, S. (2025). Enhancing undergraduate metacognitive awareness and self-efficacy: Effective instructional practices for research question formulation. *European Journal of Education*, 60(1), e12888. <https://doi.org/10.1111/ejed.12888>
38. Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26, 113–125.
<https://doi.org/10.1023/A:1003044231033>

39. Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460–475. <https://doi.org/10.1006/ceps.1994.1033>
40. Schunk, D. H. (1991). Self-Efficacy and academic motivation. *Educational Psychologist*, 26, 207–231. <https://doi.org/10.1080/00461520.1991.9653133>
41. Seidenberg, P. L., & Bernstein, D. K. (1988). Metaphor comprehension and performance on metaphor-related language tasks. *Remedial and Special Education*, 9(2), 39–45. <https://doi.org/10.1177/074193258800900208>
42. Su, L., Noordin, N., & Yang, D. (2023). Unraveling the path to success: Exploring self-regulated language learning among Chinese college EFL learners. *SAGE open*, 13(4), 21582440231218537. <https://doi.org/10.1177/21582440231218537>
43. Taghani, A., & Razavi, M. R. (2022). The effect of metacognitive skills training of study strategies on academic self-efficacy and academic engagement and performance of female students in Taybad. *Current Psychology*, 41(12), 8784–8792. <https://doi.org/10.1007/s12144-020-01278-y>
44. Teng, M. F. (2020). Tertiary-level students' English writing performance and metacognitive awareness: A group metacognitive support perspective. *Scandinavian Journal of Educational Research*, 64(4), 551–568. <https://doi.org/10.1080/00313831.2019.1595712>
45. Teng, M. F. (2021). The effectiveness of incorporating metacognitive prompts in collaborative writing on academic English writing skills. *Applied Cognitive Psychology*, 35(3), 659–673. <https://doi.org/10.1002/acp.3789>
46. Teng, M. F. (2022). Effects of cooperative–metacognitive instruction on EFL learners' writing and metacognitive awareness. *Asia Pacific Journal of Education*, 42(2), 179–195. <https://doi.org/10.1080/02188791.2020.1835606>
47. Wang, C., & Yuan, Y. (2024). A study on the influence of conceptual metaphor on college English teaching. *Journal of Language Testing & Assessment*, 7, 16–23. <https://doi.org/10.23977/langta.2024.070103>
48. Wang, K., Zhang, L. J., & Cooper, M. (2025). Metacognitive instruction for improving the effectiveness of collaborative writing for EFL learners' writing development. *The Asia-Pacific Education Researcher*, 34, 661–673. <https://doi.org/10.1007/s40299-024-00886-7>
49. Won, S., Kapil, M. E., Drake, B. J., & Paular, R. A. (2024). Investigating the role of academic, social, and emotional self-efficacy in online learning. *The Journal of Experimental Education*, 92(3), 485–501. <https://doi.org/10.1080/00220973.2023.2183375>
50. Zhou, X., Younas, M., Omar, A., & Guan, L. (2022). Can second language metaphorical competence be taught through instructional intervention? A meta-analysis. *Frontiers in Psychology*, 13, 1065803. <https://doi.org/10.3389/fpsyg.2022.1065803>