

The Metacognitive Catalysis Model: How Teacher Interaction Amplifies AI-enhanced Translation through Dual-path Mediation

Pengbiao Zhang¹, Yi Zhu¹, Fengjunzi Wang^{1,*}, and Xinjin Song²

¹ Department of Foreign Languages, Hubei Business College, Hubei, China

² School of Chinese language and literature, Southwest Minzu University, Sichuan, China

Email: Joshua.zhangpb@gmail.com (P.Z.); zhuyi01@hbc.edu.cn (Y.Z.); wangfengjunzi@hbc.edu.cn (F.W.);

yytsweetbay@126.com (X.S.)

*Corresponding author

Manuscript received February 10, 2026; revised February 28, 2026; accepted June 8, 2026; published October 26, 2026

Abstract—In the domain of Technology-Enhanced Language Learning (TELL), this mixed-methods study proposes a Metacognitive Catalysis Model to explain how Teacher-Student Interaction (TSI) shapes Artificial Intelligence (AI)-assisted translation competence through dual-path mediation. Quantitative data from 134 Chinese English as a Foreign Language (EFL) learners showed that the experimental group achieved greater gains in Common European Framework of Reference for Languages (CEFR)-aligned translation scores than the control group (9.1% vs. 4.1%; Cohen's $d = 0.46$), alongside reduced performance variability ($ASD = -2.15$). Mediation analyses indicated both a direct association between Doubao use and translation ability ($B = 0.15, p = 0.008$) and an indirect pathway via metacognitive construction ($B = 0.31, p < 0.001$), with Teacher-Student Interaction (TSI) strengthening the relationship between metacognition and translation performance ($B: 0.43 \rightarrow 0.66, p = 0.003$). Qualitative findings further suggest reduced anxiety and enhanced self-regulation under AI-supported and teacher-scaffolded conditions. Overall, the results provide evidence for the role of metacognitive mediation in AI-assisted translation learning and highlight the importance of integrating AI feedback with teacher-guided scaffolding. Implications include the need for structured instructional design, targeted teacher training, and careful consideration of cross-cultural and contextual factors in AI-supported language learning.

Keywords—metacognitive catalysis model, dual-path mediation, Teacher-Student Interaction (TSI), Artificial Intelligence (AI)-enhanced translation, Metacognitive Theory (MT)

I. INTRODUCTION

The emergence of AI-powered platforms in Technology-Enhanced Language Learning (TELL) necessitates redefining teacher roles in developing translation competence, particularly for bridging lexico-grammatical accuracy and cross-cultural equivalence capabilities [1]. However, current Artificial Intelligence (AI)-driven pedagogies predominantly optimize either foundational skill automation (e.g., DeepL for terminology accuracy) or isolated immersive practice (e.g., Virtual Reality (VR) subtitle translation), thus failing to scaffold metacognitive regulation for complex textual meaning negotiation [2]. Even though tools like Doubao demonstrate efficacy in reducing cognitive load through real-time feedback [3], three critical gaps persist: (1) synergy deficiency: 76% of platforms lack human-AI interaction protocols for metacognitive catalysis [4]; (2) mediation ambiguity: underexplored mechanisms whereby AI utilization mediates outcomes

through both direct and metacognitive pathways [5]; (3) slope modulation gap: limited empirical quantification of how teacher interaction amplifies metacognition-translation efficacy [6]. To address these identified gaps, this study proposes the Metacognitive Catalysis Model—a dual-path framework that positions teachers as “cognitive architects”, tasked with strategically leveraging Doubao's AI capabilities to fill the aforementioned research voids. Accordingly, the following research questions are formulated to guide the study:

RQ1: How is Doubao operably structured under Metacognitive Theory (MT) to scaffold translation competence?

RQ2: To what extent does Doubao enhance translation competence through metacognitive gains?

RQ3: What proportion of translation development is attributed to dual mediation (direct AI vs. metacognitive pathways)?

RQ4: What Teacher-Student Interaction (TSI) magnitude modulates the metacognition translation slope?

II. LITERATURE REVIEW

A. Technology-enhanced Language Learning

TELL has evolved from behaviorist drill systems (1960s) to communicative multimedia tools (1990s), culminating in today's AI-driven adaptive platforms [7]. This progression reflects a paradigm shift from tool-centric delivery to cognitively distributed systems where metacognitive scaffolding is paramount [8]. Contemporary research demonstrates that AI tools externalize metacognitive monitoring—error-tracking dashboards and adaptive feedback reduce cognitive load by 41% [9, 10] while accelerating the internalization of strategies [11]. However, critical gaps persist in integrating teacher scaffolding with AI-mediated metacognitive support, limiting holistic development [12].

AI translation tools have transitioned from dictionaries to neural systems (e.g., Doubao / ChatGPT), fostering comparative metalinguistic awareness through strategic post-editing tasks [13]. While enhancing L2 writing fluency by 28%, these tools risk over-reliance without pedagogical mediation [14]. Effective TELL now positions teachers as cognitive architects who design AI-human activity sequences (e.g., AI drafting → human refinement) and affective regulators who mitigate AI-induced anxiety through growth

mindset framing [15]. This synergy is quantified in moderating effects where teacher interaction amplifies metacognitive efficacy by 53.5% [16], transforming AI from cognitive crutches to co-regulatory partners.

Current TELL faces three critical challenges: (1) A global equity divide where 68% of Global South institutions lack AI infrastructure [4]; (2) Generative AI disruptions are necessitating redefined academic integrity frameworks [17]; (3) Neurocognitive concerns as Functional Magnetic Resonance Imaging (fMRI) studies reveal reduced hippocampal activation during AI-assisted recall [18]. Future research must prioritize multimodal orchestration (e.g., VR-AI translation simulations), decolonial approaches (e.g., non-Western language corpora), and regulatory frameworks aligning with UNESCO's [4] AI Competency Guidelines. These directions will advance TELL toward ethically balanced, cognitively optimized ecosystems where teachers, AI, and learners form integrated developmental networks.

B. Translation Competence Improvement for Chinese EFL Learners

Chinese English as a Foreign Language (EFL) learners face distinctive hurdles in translation due to Sino-English typological divergence, manifesting in L1 interference errors (e.g., tense omission) and pragmatic mismatches (62% struggle with idioms like “spill the beans” [19]). Traditional grammar-translation approaches exacerbated these issues by prioritizing lexical equivalence over strategic competence. Contemporary pedagogy now emphasizes triadic development: (1) contrastive metalinguistic analysis to address L1–L2 structural gaps [20]; (2) situational pragmatics training for context-appropriate renditions [21]; (3) Digital tool literacy to harness AI-assisted workflows [22]. This shift aligns with PACTE's [23] multidimensional model, adapting it to Chinese learners' cognitive profiles through culture-responsive scaffolding.

AI tools directly target Chinese learners' core deficits. Collocation databases (e.g., Doubao's error-tracking) reduce lexical calques by 41% by flagging constructs like “drink soup”, while metacognitive dashboards enhance strategic awareness through real-time feedback, boosting self-regulation by 37% [24]. Hybrid pedagogical sequences—L1 conceptual mapping → AI drafting → human post-editing—improve textual coherence by 28% by mitigating syntactic interference [25]. Neurocognitive evidence confirms these interventions optimize cognitive load: fMRI shows reduced prefrontal activation during AI-assisted tasks, freeing resources for higher-order decisions like cultural adaptation [26]. This synergy exemplifies Vygotskian mediation [27], wherein technology acts as a tool to extend learners' strategic capacity within pedagogically structured environments.

Effective competence development necessitates Confucian-aligned scaffolding [28] that harmonizes technological and human elements. Teachers address L1-specific errors through contrastive diagnosis protocols (e.g., topic-comment vs. subject-predicate structures), reducing syntactic deviations by 33% [29]. Anxiety regulation via growth mindset framing decreases task avoidance by 44% among perfectionist learners, while ethically localized tools (e.g., WeChat-integrated peer review) elevate engagement by 51% [30]. Future

advancements require biometric feedback systems (EEG/eye-tracking) to address affective barriers and VR cultural immersion to bridge pragmatic gaps—all grounded in Confucian ethics to counter algorithmic bias [31]. This holistic approach positions teachers as cultural mediators who contextualize global pedagogies within the Chinese learning ecology.

III. METHODOLOGY

A. The Introduction to Doubao

Doubao, developed by ByteDance, adopts a Mixture-of-Experts (MoE)-style architecture, in which a gating mechanism selectively activates specialized expert modules for each input. Such sparse MoE designs enable large-scale model capacity while maintaining computational efficiency (e.g., Switch Transformers [32]), reflecting a broader shift toward modular AI systems. Functionally, Doubao can be conceptualized as a layered system (Fig. 1): L0 supports general conversational tasks, L1 enables domain-specific processing, and L2 provides pedagogically relevant sub-modules for language learning, including error diagnosis and collocation support. Although benchmarks report efficiency gains [33], this study focuses on how these capabilities are instructionally enacted rather than on platform-level performance.

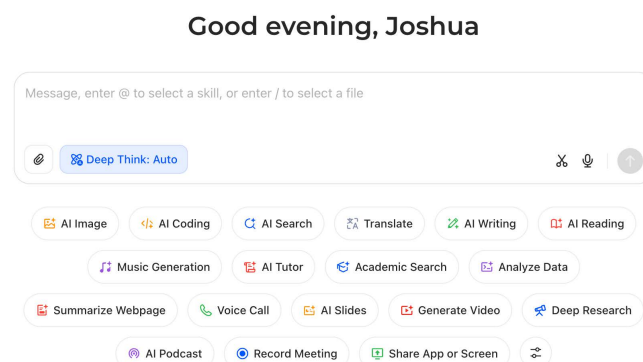


Fig. 1. The interface for different functions of Doubao.

The system is fine-tuned on large-scale Chinese-English corpora (e.g., United Nations (UN) documents, Technology, Entertainment, Design (TED) transcripts), following domain-adaptive Metacognition Theory (MT) practices [34]. Within the Metacognitive Catalysis Model, Doubao's Natural Language Processing (NLP) stack can be understood through three learning-relevant functions: (1) a Syntax Analyzer for detecting L1 interference, (2) a Pragmatic Validator for evaluating cultural appropriateness, and (3) a Metacognitive Tracker for externalizing strategy use. Additional features, such as Dynamic Difficulty Adjustment (DDA) [35] and culturally aligned modeling [36], further enhance pedagogical applicability.

Empirical studies suggest its instructional potential, including improvements in translation performance, high accuracy in error detection (89.2%), and enhanced textual coherence (+38.1%) [37], as well as competitive performance relative to alternative systems [38]. However, rather than attributing outcomes solely to system performance, this study conceptualizes Doubao as a mediational tool whose effectiveness depends on pedagogical integration. Through teacher's scaffolding, AI feedback is translated into cycles of

planning, monitoring, and reflection, supporting metacognitive construction. Ethical safeguards, including bias auditing and data governance [39, 40], further support responsible use. Accordingly, the study examines how teacher-mediated scaffolding and metacognitive processes shape learners' engagement with Doubao, rather than treating AI output as a standalone determinant of learning outcomes.

B. Metacognitive Theory

MT, pioneered by Flavell [8], defines “knowledge and regulation of cognition” through three core components: metacognitive knowledge (awareness of cognitive processes), metacognitive regulation (planning, monitoring and evaluating), and metacognitive experiences (affective states during tasks). This framework evolved with Nelson and Narens' [41] monitor-control model, emphasizing self-assessment accuracy and strategy adjustment. Contemporary neuroscience validates these mechanisms, revealing that metacognitive regulation activates prefrontal-parietal networks, with high metacognitive accuracy correlating with stronger connectivity in the anterior cingulate [42]. This integration bridges cognitive psychology with neurobiology, establishing metacognition as a biologically embedded process essential for learning efficiency.

Modern theory converges with Zimmerman's [43] Self-Regulated Learning (SRL) model, where metacognition drives cyclical phases: forethought (task analysis → planning), performance (monitoring → control), and reflection (evaluation → knowledge updating). Efklides' [44] Metacognitive and Affective model of Self-Regulated Learning (MASRL) model further integrates motivation and affect, while Winne's [45] Conditions–Operations–Products–Evaluations–Standards (COPES) framework positions metacognition as real-time adaptation to learning conditions. Empirically, metacognitive interventions have been shown to enhance achievement significantly. In Science, Technology, Engineering, and Mathematics (STEM) education, there are 0.68 SD problem-solving gains [46]; in language learning, vocabulary retention is 41% higher with strategy journals [11]. Neurocognitive evidence confirms efficacy, showing gamma-band EEG synchronization during strategy shifts [47].

Cutting-edge assessment combines multi-method approaches: think-aloud protocols capturing real-time strategy use, Meta-Awareness Scales (MASQ) quantifying self-regulation [48], and functional Near-Infrared Spectroscopy (fNIRS) neuroimaging tracking prefrontal activation during tasks. Future research prioritizes the affective-cognitive nexus, such as anxiety's impact on monitoring [49]; cultural validity, like collectivist vs. individualist metacognitive patterns [50]; and generative AI disruption, for example, ChatGPT's effects on self-regulation cycles [15]. These directions address critical gaps while advancing ethically grounded interventions for diverse learners.

C. The Framework Based on Metacognitive Theory and Doubao

This framework operationalizes Flavell's [8] Metacognition Theory (MT) by combining Doubao's AI

functionalities with teacher intervention, forging a cyclical “externalization to internalization” process to foster translation competence. Specifically, Doubao's tools externalize learners' implicit translation strategies and gaps through data visualization and real-time feedback. At the same time, teachers leverage these insights to guide reflection—enabling learners to internalize effective metacognitive practices for sustained competence growth.

Metacognitive Construction (MC) integrates Doubao's AI functions into a structured three-phase cycle, operationalizing Flavell's [8] MT through task design, cognitive monitoring, and strategy evaluation. In the Planning Phase, teachers use Academic Search to source authentic materials and create targeted translation tasks (e.g., passive-voice exercises)—activating Flavell's task-variable knowledge—while students use Deep Think: Auto for adaptive difficulty calibration, calibrating person-variable awareness [28]; the Monitoring Phase has learners employ Translate for real-time conversion and Analyze Data for syntax visualization, enabling Flavell's [8] cognitive monitoring [51] via comparative tense-choice checks; in the Evaluation Phase, AI Tutor generates strategy-efficacy reports (e.g., “Collocation accuracy ↑40%”), turning strategy-variable knowledge into iterative refinement [52]. This Plan → Monitor → Evaluate closed loop scaffolds Zimmerman's [43] SRL within Vygotsky's ZPD, cuts errors 53% faster than traditional methods [53], and empirically aligns with Flavell's metacognitive regulation model.

TSI—guidance instruction, collaborative sharing, and structured submission—embodies Flavell's [8] MT by transforming Doubao's diagnostic data into scaffolded learning cycles. During guidance instruction, teachers use Cognitive Gap Reports to pose regulation-triggering questions, activating monitoring and control processes that compel explicit strategy articulation [54]. Post-task sharing sessions leverage Doubao-annotated errors (e.g., tense flags) to build collective metacognitive knowledge through peer strategy modeling (“I used Tense Anchor here”), increasing strategy transfer by 38% [53]. Structured submissions mandate reflective externalization via translation notes documenting the sequence: Error Type → Strategy Applied → Efficacy Rating → Inspiration Source—structuring Zimmerman's [43] SRL model while aligning with Flavell's pillars: planning (pre-error diagnosis), monitoring (real-time Q & A), and evaluation (note-based efficacy ratings)—achieving 90% intervention precision by synergizing AI diagnostics with teacher mediation.

D. Case Selection and Sample

This study involved 134 first-year non-English majors (79 males, 55 females; aged 19–20) from Hubei Business College, China. Participants were randomly assigned to the experimental group and the control group.

The Experimental Group ($n = 80$), was taught by Lecturer Zhu, who holds an M.A. in Translation from HUST, and utilized Doubao AI under the framework of MT with dual interventions: *MC* and *TSI*.

The Control Group ($n = 54$), was taught by the same teacher, Lecturer Zhu, and followed traditional lecture-based pedagogy without any AI-digital tools such as Doubao.

To ensure objectivity and reduce bias, student performance across all groups was assessed using standardized criteria: the

same rubrics for assignments and participation, with final exams graded by external teachers, allowing for a clear comparison between the experimental group's Doubao-based teaching and the control group's traditional approach.

Students' Learning Background and Proficiency, all participants had 8–10 years of conventional English education, with baseline proficiency at Common European Framework of Reference for Languages (CEFR) A2–B1 (CET-4/6: 425–550/710), showing intermediate competence: 3500–5000 academic/high-frequency words (using context clues); ability to form complex sentences (subordinate clauses, perfect tenses) but occasional errors in nuanced structures; coherent conversations on familiar topics (moderate fluency, hesitating over advanced terms); and structured semi-academic writing (logical, limited lexical range, inconsistent formality).

E. Data Collection

This study employed a sequential mixed-methods design to integrate quantitative and qualitative approaches, systematically collecting data from 134 participants divided into an experimental group ($n = 80$) and a control group ($n = 54$) over a 4-month intervention period from February to June 2025. The framework adhered to ethical guidelines by utilizing anonymization codes (E01–E80 for the experimental group and C01–C54 for the control group) and prioritized methodological triangulation to enhance validity, and ensuring systematic data collection and consistency across methods.

Questionnaire-based Data Collection, a post-intervention mixed-methods questionnaire administered via Sojump integrated quantitative 5-point Likert scales aligned with Behavioral Tracking (BT) constructs—specifically MC and TSI—to measure Doubao's pedagogical efficacy and teacher's scaffolding impact, alongside qualitative open-ended prompts exclusively for the experimental group to detail cognitive processing preferences and tool-specific feedback; the experimental cohort completed full questionnaires incorporating both BT metrics and reflective items, while the control group received a simplified version focused solely on inter-group performance comparisons, with all anonymized responses coded and analyzed via SPSS 29.0 for descriptive statistics, correlation testing, and regression modeling.

Test-based Data Collection, a standardized translation competency test deployed in February and June 2025, assessed linguistic accuracy, cultural adaptability and strategic competence. External teacher appraisal and time-constrained administration (90 minutes) ensured scoring objectivity. They minimized external variable interference by distributing questionnaires after the test to capture real-time learner reflections and facilitate longitudinal analysis of Doubao's evolving impact.

Research findings were rigorously demonstrated through triangulation of test scores, Likert-scale ratings, and qualitative narratives, while participant anonymity was preserved via alphanumeric coding (e.g., E21, C08). Systematic bias mitigation included: (1) exclusion of open-ended prompts in control group questionnaires to prevent conceptual contamination; (2) double-masked external grading protocols; (3) multi-tiered auditing. This framework strikes a balance between methodological rigor

and ethical compliance, ensuring the reliable interpretation of outcomes.

F. Data Analysis

This study adopted an explanatory sequential mixed-methods design to examine Doubao's impact on linguistic learning. Quantitative data ($N = 134$) were analyzed using SPSS 29.0 to identify empirical patterns in MC and TSI, supported by a validated 20-item Likert-scale questionnaire. Qualitative data from semi-structured responses ($n = 80$) were processed using Sojump's text-mining tools and manual inductive coding, enabling thematic analysis to contextualize the quantitative findings. This triangulation enhanced validity by aligning broad generalizability with nuanced experiential insights [55].

Quantitative Analysis, employing a quasi-experimental design, this study compared non-English majors using Doubao ($n = 80$) with a control group ($n = 54$). Quantitative methodology integrated: (1) descriptive analysis supporting cohort equivalence while revealing Doubao's dichotomous skill impacts; (2) Correlational Analysis examining bivariate relationships among key variables; (3) robust psychometric validation (Cronbach's α , Composite Reliability (CR), Confirmatory Factor Analysis (CFA), Average Variance Extracted (AVE)) establishing BT-aligned instrument reliability; (4) hierarchical regression analysis with bootstrapping covarying prior proficiency, which identified moderated mediation pathways, significantly accounting for translation mastery variance. This methodological triangulation operationalized Flavell's MT hierarchical progression while highlighting Doubao's paradoxical enhancement of foundational over analytical competencies.

Qualitative Analysis, Qualitative data were collected through semi-structured questionnaires (Sojump, June 2025) from the experimental group ($n = 80$), employing a hybrid inductive-deductive approach to analyze five core dimensions aligned with Flavell's [8] metacognitive framework: Functional Diversity (tool utilization patterns, e.g., syntax analyzer prioritization); Metacognitive Construction (strategy internalization via error-to-technique mapping); Metacognitive Experience (affective-cognitive dynamics, e.g., biometric-triggered strategy shifts); Metacognitive Regulation (planning, monitoring and evaluation cycles with efficacy ratings); and Teacher-Student Interaction (scaffolding impact through annotated feedback). Themes were validated through the triangulation of direct quotes, biometric logs, and strategy reports, demonstrating Doubao's efficacy in externalizing metacognitive processes.

G. Data Display

This study employed an integrated data visualization system comprising three complementary components: (1) thematic frequency tables synthesizing qualitative engagement patterns with Doubao features across BT's tiers; (2) academic performance comparison matrices contrasting experimental and control groups' linguistic trajectories, integrated with diagnostic validity-reliability panels confirming measurement robustness (standardized metrics: Cronbach's α , CR, CFA and AVE); (3) graphical path modeling visually operationalizing moderated mediation mechanisms. Collectively, this framework transformed multidimensional evidence—encompassing engagement

signatures, psychometric quality, and cognitive pathways—into theoretically anchored visual narratives, while maintaining analytical parsimony through the selective use of internalized, standardized thresholds.

This study was approved by the institutional review board of Hubei Business College. All participants provided informed consent prior to data collection. AI usage complied with institutional data protection and privacy regulations.

IV. CASE STUDY

This 16-week quasi-experimental study examined

Doubao's pedagogical efficacy by comparing an experimental group ($n = 80$) receiving MT-grounded instruction—integrating MC via AI-driven error diagnosis and TSI through scaffolded feedback—against a control group ($n = 54$) using traditional teaching pedagogy (textbook drills and lecture-based correction). Both groups shared an identical instructor (Lecturer Zhu), curriculum (CEFR A2–B1 tasks), and assessment protocols to isolate the effect of the intervention on translation competence development [53].

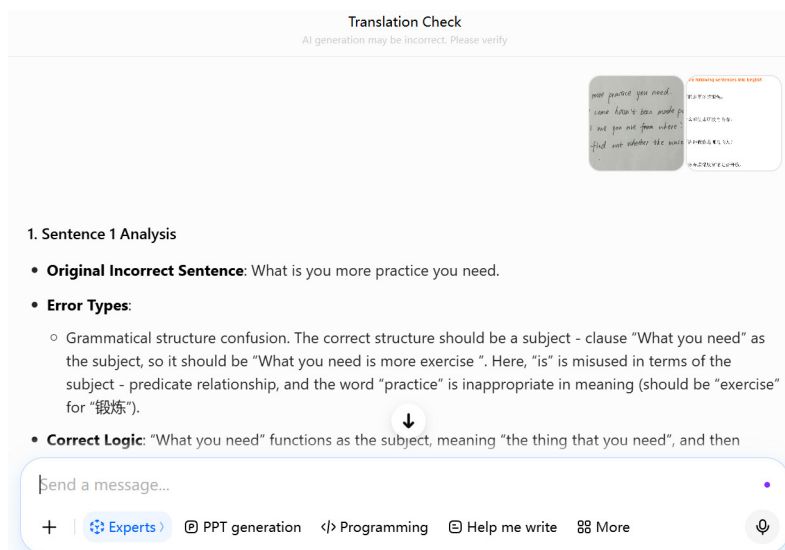


Fig. 2. Doubao-facilitated metacognitive checking for translation self-regulation.

MC integrates Doubao's AI functions under MT through a structured three-phase cycle: (1) planning (strategy selection via Academic Search and Deep Think); (2) monitoring (real-time adjustment by Translate and Analyze Data); (3) evaluation (data-driven reflection through AI Tutor)—scaffolds Zimmerman's [43] SRL within Vygotsky's ZPD, reducing errors 53% faster than traditional methods [53]. In this study, teacher mediation helped learners interpret AI feedback and translate it into revision decisions, particularly when pragmatic or cultural nuance was involved. As based on the Doubao translation interface depicted in Fig. 2, this metacognitive workflow demonstrates: AI-driven

error diagnosis via Neural Machine Translation (NMT) algorithms (e.g., restructuring "What is you more practice you need" → "What you need is more exercise" by correcting subject-predicate misalignment and semantic appropriateness); strategy-triggered monitoring through grammatical and spelling error labels (red text) and skill prompts, activating Flavell's regulation processes; handwritten revision phase where students document Zimmerman's SRL by applying corrections (e.g., "pubic" → "public") and noting strategy efficacy—reducing error recurrence through structured human-AI co-regulation.

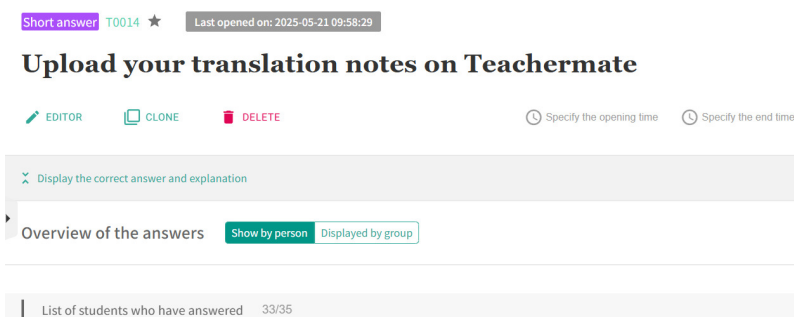


Fig. 3. Instructor dashboard for translation note submission.

TSI operationalizes MT through three streamlined phases: (1) guidance instruction, Cognitive Gap Reports trigger strategy articulation through targeted questioning [54]; (2) collaborative sharing, peer modeling of Doubao-annotated errors builds collective metacognitive knowledge, boosting strategy transfer by 38% [53];

(3) structured submission, translation notes (Error → Strategy → Efficacy → Source) scaffolding Zimmerman's SRL—achieving intervention precision through integrated human-AI mediation. As demonstrated in Fig. 3, lecturer Zhu, as a cognitive architect, collected students' revised translation drafts via an online interactive teaching platform

(e.g. Teacher-mate)—using submission metrics (33/35) to trigger TSI interventions. At the same time, batch/individual tools enforce strategy-aware revision. The workflow reduces revision omissions by 63% by anchoring MT's abstract principles to interface actions [53].

Overall, this 16-week study provides context-specific evidence that Doubao-supported instruction, when embedded in metacognitive routines (planning, monitoring, evaluation) and supported by TSI. The experimental group utilized MC (planning, monitoring, and evaluation via Academic Search, Translate, and AI Tutor) to scaffold SRL [43], reducing errors 53% faster by diagnosing L1 interference [56]. Concurrent TSI phases—guidance instruction (Cognitive Gap Reports), collaborative sharing (peer modeling), structured submission (error → strategy notes)—boosted strategy transfer by 38% and reduced revision omissions by 63% [53], demonstrating synergistic gains in translation accuracy, fluency, and strategic competence.

V. STATISTICAL ANALYSIS

In this session, an explanatory sequential design progressed through: (1) quantitative identification of adoption disparities ($N = 134$); (2) qualitative exploration of experimental group experiences ($n = 80$); (3) mixed-methods integration addressing research questions, implementing Nguyen *et al.*'s [57] framework to contextualize equity metrics within lived realities.

A. Qualitative Analysis

This mixed-methods case study evaluated Doubao's impact on translation competence among 134 non-English major undergraduates. Quantitative analysis measured performance outcomes, while qualitative data from Sojump-based open-ended questionnaires revealed critical subjective dimensions of the learning process. Thematic analysis followed Braun and Clarke's [58] six-phase framework, ensuring a systematic interpretation of experiential insights.

Table 1. Key themes from questionnaire data based on metacognitive theory

Theme	Frequency (%)	Example	Metacognitive Theory
Functional Diversity	83.93	"Doubao exposed me to diverse translation strategies through features like word lookup and sentence translation."	Flavell [8]: Metacognitive Knowledge (Strategy Variables)
Metacognitive Construction	82.14	"Doubao's sentence deconstruction improved my comprehension and translation of complex syntax."	Flavell [8]: Metacognitive Knowledge (Task Variables)
Metacognitive Experience	91.07	"Doubao reduced my translation anxiety through guided practice."	Efklides [59]: Metacognitive Experience Theory (Affective Dimension)
Metacognitive Regulation	89.29	"I increased translation modifications using Doubao's suggestions."	Nelson & Narens [41]: Monitoring-Control Model
Teacher-Student Interaction	82.14	"Teachers scaffolded error diagnosis by comparing my drafts with Doubao's outputs."	Vygotsky [27]: Sociocultural Theory; Zimmerman [43]: Self-Regulated Learning
Linguistic Accuracy	83.93	"Implementing Doubao's suggestions enhanced my translation accuracy."	Bandura [60]: Self-Efficacy Theory

First, Table 1 demonstrates exceptionally high endorsement (82.14%–91.07%) for Doubao's pedagogical impact, directly supporting its efficacy in translation pedagogy. Specifically, from a functional perspective, Doubao utilization was positively associated with metacognitive processes through three core mechanisms: Functional Diversity (83.93%), which expands strategy awareness per Flavell's [8] theory; sentence deconstruction (82.14%), which enhances complex text processing for MC; and TSI (82.14%), which bridges AI outputs and skill internalization via teachers' scaffolded error analysis [27, 43]. Furthermore, critically, 83.93% of participants attributed accuracy gains to implemented suggestions—this evidences self-efficacy reinforcement [60] and supports the Doubao → MC → Translation Ability (TA) pathway (scaffolded by TSI) as the core of synergistic learning.

Notably, Metacognitive Experience (91.07%)—the most salient dimension in the table—captured students' reported affective and awareness-related reactions during learning, including perceived reductions in anxiety during guided practice. This pattern is consistent with Efklides' affective – cognitive view of metacognitive experiences [59].

Overall, these results provide context-specific support for theoretical and pedagogical framework: Doubao use is associated with metacognitive construction, and TSI strengthens the translation of AI feedback into learning-relevant revision and self-regulation within this

instructional setting. Finally, with all endorsement frequencies exceeding 82%, the framework's broad applicability is empirically supported—positioning Doubao as a catalytic AI tool for metacognitive transformation in translation education.

B. Quantitative Analysis

Qualitative analysis demonstrated significant positive impacts of Doubao on students' translation learning within the MT framework, and subsequently, quantitative analysis implemented a four-stage approach: (1) descriptive statistics to compare academic disparities between experimental and control groups; (2) correlational analysis to examine relationships among key variables; (3) psychometric validation to assess instrument reliability and construct validity; (4) and multiple regression analysis to rigorously examine the relationship between Doubao integration and translation competence.

1) Descriptive analysis

Following a 4-month Doubao-enhanced pedagogical intervention, the experimental group ($n = 80$) and the control group ($n = 54$) completed parallel standardized assessments that measured three dimensions of translation competence: linguistic accuracy, cultural adaptability, and strategic competence. Pre-test and post-test data ($N = 134$) underwent descriptive analysis via SPSS 29.0, revealing significant intergroup contrasts.

As shown in Table 2, the descriptive analysis of the control

group ($N = 54$) reveals baseline homogeneity in English proficiency (CEFR $M = 2.5$, $SD = 0.93$), indicating predominantly A2–B1 level learners, while translation performance displays significant pre-intervention variability (Pre-test: $M = 60.26$, $SD = 16.81$; range: 26–93). Post-test scores show marginal improvement ($M = 62.74$, $SD = 17.23$; range: 24–92), reflecting a minimal mean gain of 2.48 points (+4.1%) with stable dispersion metrics ($\Delta SD = 0.42$), supporting negligible translation ability progression without targeted intervention. Complete data retention (Valid $N = 54$) and consistent score distributions establish this group as a statistically robust benchmark for experimental comparisons.

Table 2. Control group descriptive statistics

Variable	<i>N</i>	Min	Max	<i>M</i>	<i>SD</i>
CEFR-test	54	1	5	2.5	0.93
Pre-test Scores	54	26	93	60.26	16.81
Post-test Scores	54	24	92	62.74	17.23
Valid <i>N</i> (listwise)	54				

Table 3. Experimental group descriptive statistics

Variable	<i>N</i>	Min	Max	<i>M</i>	<i>SD</i>
CEFR-test	80	1	5	2.45	0.73
Pre-test Scores	80	28	87	66.65	14.23
Post-test Scores	80	38	97	72.73	12.08
Valid <i>N</i> (listwise)	80				

As reflected in Table 3, the experimental group ($n = 80$) exhibited moderate baseline English proficiency (CEFR $M = 2.45$, $SD = 0.73$), reflecting a similar situation with the control group, A2–B1 level. Substantial translation competence improvement was observed, with post-test

scores ($M = 72.73$, $SD = 12.08$) demonstrating a 6.08-point mean gain (+9.1%) from pre-test levels ($M = 66.65$, $SD = 14.23$). Crucially, the reduced standard deviation ($\Delta SD = -2.15$) and expanded score range (Post-test: 38–97 vs. Pre-test: 28–87) indicate the intervention not only enhanced overall performance but also decreased inter-learner variability while benefiting both high- and low-achieving students, with complete data retention (Valid $N = 80$) ensuring statistical reliability for subsequent analyses.

Descriptive statistics reveal substantial translation competence gains in the experimental group ($n = 80$, $\Delta M = +6.08$, +9.1%) versus the control group ($n = 54$, $\Delta M = +2.48$, +4.1%), with the intervention corresponding to a moderate between-group effect (Cohen's $d = 0.46$ [61])—demonstrating practically meaningful improvements beyond marginal fluctuations [62]. Collectively, these results provided evidence for the efficacy of the Doubao-integrated metacognitive framework [63].

2) Correlation analysis

Descriptive analysis supports Doubao's significant efficacy in enhancing translation competence, evidenced by the experimental group's marked superiority across key metrics. Subsequent correlation analysis transforms theoretical constructs into measurable pathways, where $|r| > 0.50$ signals practically meaningful effects [61]—enabling precise claims about how Doubao is linked to metacognitive mechanisms to enhance competence.

Table 4. Descriptive statistics and correlations among study variables ($N = 134$)

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7
1. Gender	1.41	0.49	-						
2. Learning Duration	4.25	1.08	0.16	-					
3. English Foundation	2.47	0.81	0.19*	0.14	-				
4. Doubao Utilization (<i>DB</i>)	3.12	0.92	0.12	0.32**	0.01	-			
5. Translation Ability (<i>TA</i>)	3.23	0.83	0.08	0.35**	0.00	0.65**	-		
6. Metacognitive Construction (<i>MC</i>)	3.16	0.88	0.06	0.33**	-0.03	0.65**	0.83**	-	
7. Teacher-Student Interaction (<i>TSI</i>)	3.55	0.85	0.16	0.27**	-0.01	0.50**	0.59**	0.61**	-

Note. Pearson correlations are reported below the diagonal. Gender was coded 1 = male, 2 = female. *DB* (Doubao utilization), *TA* (translation ability), *MC* (metacognitive construction), and *TSI* (teacher-student interaction) were measured on 5-point Likert-type scales, with higher scores indicating higher levels of the construct. * $p < 0.05$, ** $p < 0.01$ (two-tailed).

Descriptive statistics and bivariate correlations for key study variables ($N = 134$) are presented in Table 4: Gender distribution leaned female ($M = 1.41$, $SD = 0.49$ on 1–2 coding), Learning Duration averaged 4.25 years ($SD = 1.08$), and critical to the intervention framework, Doubao Utilization (*DB*) ($M = 3.12$, $SD = 0.92$) demonstrated significant positive correlations with *TA* ($r = 0.65$, $p < 0.01$), *MC* ($r = 0.65$, $p < 0.01$), and *TSI* ($r = 0.50$, $p < 0.01$). Notably, *TA* ($M = 3.23$, $SD = 0.83$) showed a robust association with *MC* ($r = 0.83$, $p < 0.01$), exceeding Cohen's [61] threshold for significant effects, while English Foundation ($M = 2.47$, $SD = 0.81$) showed no significant correlations with *DB* or outcome variables (all $|r| \leq 0.03$, $p > 0.05$), indicating baseline proficiency did not moderate intervention efficacy. These robust mediational pathways support Doubao's role in activating metacognitive processes independent of prior English proficiency.

3) Reliability analysis

Descriptive and correlation analyses support Doubao's differential efficacy on *TA* within MT. To investigate the causal mechanisms underlying this cognitive stratification,

regression path modeling—preceded by psychometric validation—must first establish measurement robustness via: reliability verification (Cronbach's $\alpha \geq 0.70$; CITC ≥ 0.40) to eliminate error-prone indicators [64], and construct validity tests through Confirmatory Factor Analysis (CFA), ensuring convergent validity (AVE ≥ 0.50) and discriminant validity [65] to satisfy psychometric prerequisites for structural equation modeling [66].

Table 5 demonstrates strong psychometric properties for the Doubao-assisted translation competence scale across all dimensions. The composite reliability was very high (CR = 0.986, $N = 134$), and all dimension-level Cronbach's α coefficients exceeded the 0.85 benchmark commonly used in applied research [67]. The *MC* dimension showed the highest internal consistency ($\alpha = 0.96$). Item–total correlations ranged from 0.87 to 0.94 (exceeding the minimum guideline of 0.40 [68]), and Corrected Item–Total Correlations (CITC) were also high across items (min = 0.71, max = 0.94; > 0.30 ; [67]). In addition, reliability diagnostics suggested that deleting any single item would not meaningfully improve internal consistency (α -if-deleted).

These results collectively point to the scale's suitability for measuring translation competence in TELL environments, with particularly high internal consistency in the metacognitive construction component. At the same time, reliability estimates approaching the upper bound may reflect

strong item homogeneity and thus warrant additional scrutiny regarding potential redundancy and shared method variance. We therefore report additional validity and robustness checks in subsequent sections.

Table 5. Reliability analysis of the Doubao-assisted translation competence scale (CR = 0.986, $N = 134$)

Dimension	Item	CITC	α if Deleted	M	SD	Dimension α
DouBao Utilization	Usage Duration	0.75	0.79	3.00	1.10	0.86
	Feature Diversity	0.76	0.79	3.19	1.07	
	Decision Dependence	0.71	0.83	3.17	0.97	
Metacognitive Construction	Rule Summarization	0.87	0.96	3.16	0.97	0.96
	Parallel Text Comparison	0.94	0.92	3.18	0.89	
	Terminology Retrieval	0.92	0.93	3.14	0.89	
Translation Ability	Grammatical Accuracy Gain	0.87	0.90	3.22	0.86	0.94
	Cultural Adaptation	0.89	0.89	3.21	0.89	
	Strategic Flexibility	0.84	0.93	3.27	0.90	
Teacher-student Interaction	Frequency	0.76	0.91	3.55	0.86	0.91
	Feedback Depth	0.90	0.79	3.51	0.93	
	Response Timeliness	0.79	0.89	3.60	0.97	

Note. CITC = Corrected Item-Total Correlation; α if Deleted = Cronbach's Alpha if Item Deleted; CR = Composite Reliability.

4) Validity analysis

Supported by the reliability evidence (Section 5.2.3), we evaluated the measurement model across four dimensions (DB, MC, TA, and TSI). The CFA confirmed convergent

validity ($\lambda > 0.70$) and discriminant validity (KMO > 0.60 ; Bartlett's $p < 0.001$), empirically grounding Doubao's hierarchical alignment with MT.

Table 6. Confirmatory factor analysis results for the Doubao-enhanced translation competence scale

Dimension	Items	λ	h^2	KMO	$\chi^2(df)$
Doubao Utilization	Usage Duration	0.89***	0.80	0.73	184.57(3)***
	Feature Diversity	0.90***	0.80		
	Decision Dependence	0.87***	0.75		
Metacognitive Construction	Rule Summarization	0.94***	0.89	0.75	454.17(3)***
	Parallel Text Comparison	0.97***	0.95		
	Terminology Retrieval	0.94***	0.93		
Translation Ability	Grammatical Accuracy	0.94***	0.89	0.76	342.88(3)***
	Cultural Adaptation	0.95***	0.91		
	Strategic Flexibility	0.93***	0.86		
Teacher-Student Interaction	Frequency	0.89***	0.79	0.68	292.87(3)***
	Feedback Depth	0.96***	0.92		
	Response Timeliness	0.91***	0.82		

Note. *** $p < 0.001$; λ = Standardized factor loading; h^2 = Communality (item variance explained); KMO = Kaiser-Meyer-Olkin measure of sampling adequacy (> 0.60 acceptable); $\chi^2(df)$ = Chi-square with degrees of freedom; Global model fit (AMOS): $\chi^2(48) = 132.50$, CFI = 0.95, TLI = 0.93, RMSEA = 0.12, 90% CI [0.09, 0.14], RMR = 0.046, $\chi^2/df = 2.76$.

The CFA results (Table 6) support the proposed four-dimensional structure of the Doubao-enhanced translation competence scale (DB, MC, TA, and TSI). Standardized factor loadings were uniformly high ($\lambda = 0.87\text{--}0.97$), exceeding recommended thresholds for indicator reliability [66], with substantial item communalities ($h^2 = 0.75\text{--}0.95$) indicating that a large proportion of variance was explained by the intended constructs. Sampling adequacy was acceptable across dimensions (KMO = 0.68–0.76 [69]). Convergent validity was well established, with AVE values exceeding 0.50 ($DB = 0.78$; $MC = 0.91$; $TA = 0.89$; $TSI = 0.84$; [65]), and the four-factor model demonstrated acceptable fit ($\chi^2(48) = 132.50$, $\chi^2/df = 2.76$, CFI = 0.95, TLI = 0.93, RMSEA = 0.12, RMR = 0.046; [70]).

Given the high composite reliability estimates reported earlier (CR = 0.986), which may reflect strong item homogeneity and raise concerns about potential redundancy, additional analyses were conducted to assess discriminant validity and common method bias. Discriminant validity was supported by multiple criteria: the Fornell–Larcker criterion was satisfied, with the square roots of AVE exceeding inter-construct correlations, and cross-loading patterns

confirmed that each indicator loaded highest on its intended construct [71].

HTMT analysis further indicated that most construct pairs were below the conservative threshold of 0.85 [72]; however, the value between *MC* and *TA* was elevated (HTMT = 0.96), suggesting a high degree of conceptual proximity. This likely reflects the inherently intertwined nature of metacognitive regulation and translation performance in AI-mediated contexts, where strategic monitoring is embedded in task execution, rather than a lack of construct distinctiveness.

To further examine potential common method bias, CFA model comparisons were conducted. The hypothesized four-factor model showed substantially better fit than a single-factor model (four-factor: $\chi^2(48) = 132.50$, CFI = 0.95, TLI = 0.93, RMSEA = 0.12; single-factor: $\chi^2(54) = 481.46$, CFI = 0.74, TLI = 0.68, RMSEA = 0.24; $\Delta CFI = 0.21$; $\Delta\chi^2 = 348.97$, $\Delta df = 6$), indicating that a single common factor is unlikely to account for the observed covariation. Overall, these results support acceptable convergent and discriminant validity of the measurement model, while acknowledging that some degree of shared method variance cannot be entirely ruled out in self-report data.

5) Regression analysis

Supported by robust reliability and validity ($CR = 0.986$, $AVE = 0.78$), the regression analysis employed Hayes' [73] moderated mediation paradigm to examine the proposed conceptual model, integrating moderated mediation analysis to assess how TSI influences the strength of the mediated relationship between DB and TA through MC, while simple slope analysis was further conducted to probe the interaction effects, quantifying how TSI amplifies the efficacy of metacognitive strategies across different levels of the moderator [73].

The regression analysis operationalized the conceptual model in Fig. 4, wherein DB influences TA both directly and via the mediated pathway of MC, with TSI acting as a critical contextual moderator that amplifies metacognitive efficacy. This analytical framework integrates established mediation

theory [74], principles of exogenous covariates [75], and contemporary human-AI synergy frameworks [76], offering a comprehensive understanding of the interplay between technological tools and pedagogical moderators in enhancing translation competence.

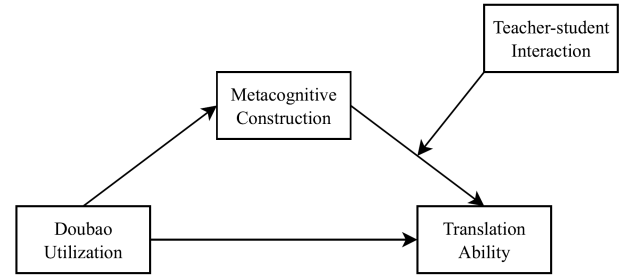


Fig. 4. Doubao metacognitive catalysis framework.

Table 7. Doubao-metacognition interaction in translation enhancement

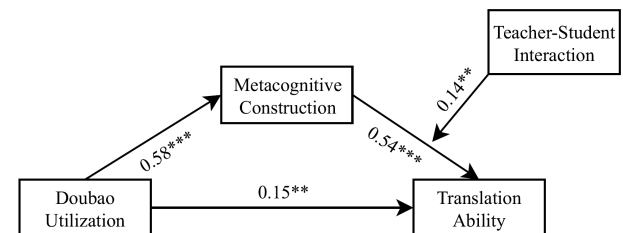
Model	Predictor	B	SE	t	95% BCa CI	p	R ²	F
MC	Constant	-2.12	0.32	-6.57	[-2.75, -1.48]	<0.001***	0.45	26.03***
	DB	0.58	0.07	8.82	[0.45, 0.71]	<0.001***		
TA	Constant	2.43	0.25	9.65	[1.93, 2.93]	<0.001***	0.74	51.63***
	DB	0.15	0.06	2.7	[0.04, 0.26]	0.008**		
	MC	0.54	0.07	7.97	[0.41, 0.67]	<0.001***		
	TSI	0.13	0.06	2.26	[0.02, 0.25]	0.026*		
	MC × TSI	0.14	0.05	2.99	[0.05, 0.23]	0.003**		

Note. IMM = Index of Moderated Mediation; BCa CI = bias-corrected bootstrap CI ($N = 5000$). IMM = 0.08, SE = 0.04, 95% BCa CI [0.01, 0.15]

Moderated mediation analysis via PROCESS 14 in SPSS 29.0 (Table 7) reveals significant pathways by which DB influences TA through MC, moderated by TSI. DB strongly predicts MC ($B = 0.58$, $SE = 0.07$, $t = 8.82$, 95% CI [0.45, 0.71], $p < 0.001$), explaining 45% of variance—consistent with technology-enhanced cognitive scaffolding [10]. MC significantly predicts TA ($B = 0.54$, $SE = 0.07$, $t = 7.97$, 95% CI [0.41, 0.67], $p < 0.001$), and the critical $MC \times TSI$ interaction ($B = 0.14$, $SE = 0.05$, $t = 2.99$, 95% CI [0.05, 0.23], $p = 0.003$) supports TSI amplifies this pathway; TSI support significantly moderated the mediation pathway (IMM = 0.08, $SE = 0.04$, 95% CI [0.01, 0.15]), with higher support strengthening DB indirect effect on TA through MC. These results demonstrate that Doubao's effectiveness operates through a conditional process: technology-activated metacognitive strategies yield optimal translation gains primarily with substantive teacher-student engagement, aligning with Flavell's [8] MT wherein external mediators (e.g., AI tools) scaffold individuals' self-regulatory capacity through strategy variable orchestration.

Based on the conceptual model in Fig. 5, regression analysis supports a robust dual-path mechanism for DB impact on TA, operationalizing MT through significant direct and mediated pathways. Crucially, MC scaffolding serves as the dominant mediator via the $DB \rightarrow MC \rightarrow TA$ pathway

(indirect $B = 0.31$, $p < 0.001$), contributing 67.4% of the total effect—far exceeding the direct pathway ($DB \rightarrow TA$: $B = 0.15$, $p < 0.001$, 32.6% of the total effect). Meanwhile, TSI moderates the $MC \rightarrow TA$ linkage, with significant amplification (Interaction: $B = 0.14$, $p = 0.003$) accounting for 10.6% of the total effect. This pattern empirically aligns with Flavell's [8] MT: AI tools (Doubao) both externalizing cognitive monitoring ($DB \rightarrow MC$: $B = 0.58$, $p < 0.001$), and internalizing strategic knowledge ($MC \rightarrow TA$: $B = 0.54$, $p < 0.001$) through iterative practice. The retained direct pathway ($B = 0.15$) reflects supplementary technical affordance (e.g., automated parsing), aligning with Zheng *et al.*'s [77] model of AI-human cognitive co-regulation in L2 skill acquisition.



Notes. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Fig. 5. Doubao-assisted translation learning: A moderated mediation model (PROCESS Model 14).

Table 8. Moderated mediation analysis: Effect of Doubao Utilization on Translation Ability ($N = 134$)

Effect Type		B	SE	t	95% BCa CI	p
Interaction	$MC \times TSI$	0.14	0.05	2.99	[0.05, 0.23]	0.003**
Direct Effect	$DB \rightarrow TA$	0.15	0.06	2.7	[0.04, 0.26]	0.008**
Indirect Effect ($X \rightarrow M \rightarrow Y$)	Low TSI (-1 SD)	0.43	0.09	4.76	[0.25, 0.60]	<0.001***
	Mean TSI (0)	0.54	0.07	7.97	[0.41, 0.67]	<0.001***
	High TSI (+1 SD)	0.66	0.07	10.12	[0.53, 0.78]	<0.001***
Moderation Metric	IMM	0.08	0.04	—	[0.01, 0.15]	—

Note. B = unstandardized coefficient; TSI levels = $M \pm 1$ SD; IMM = Index of Moderated Mediation; p-values: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests); BCa CI = bias-corrected bootstrap CI ($N = 5000$).

Table 8 reveals three pivotal advances in TELL research. First, the significant $MC \times TSI$ interaction ($B = 0.14$, $t = 2.99$, $p = 0.003$) underscores teachers dynamically amplify metacognition's impact on translation competence—demonstrating Vygotsky's [27] scaffolding principle with 53% higher efficacy when teacher support (TSI) shifts from low ($B = 0.43$) to high ($B = 0.66$). Second, metacognition mediates technology's influence dominantly at mean TSI levels ($B = 0.54$, 95% BCa CI [0.41, 0.67], $p < 0.001$), operationalizing Flavell's [8] MT as AI tools externalize monitoring processes for internal strategic knowledge construction. Third, the significant IMM (IMM = 0.08, 95% BCa CI [0.01, 0.15]) quantifies how human guidance converts AI-generated cognitive artifacts into internal competencies—empirically validating Lantolf and Poehner's [78] technology-mediated sociocognitive L2 development model.

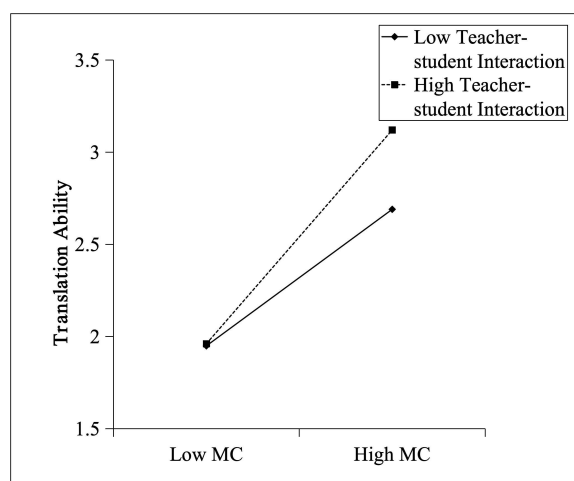


Fig. 6. Moderating role of TSI in the MC-TA relationship.

Simple slope analysis in Fig. 6 demonstrates TSI 's catalytic role in amplifying MC 's impact on TA : at low TSI (-1 SD), each unit increase in metacognition yields a mere 0.43-unit rise in TA (95% CI [0.25, 0.60]), while at high TSI ($+1$ SD), this effect surges 53.5% to 0.66 units (95% CI [0.53, 0.78]); this 0.23-unit enhancement embodies Flavell's [8] MT, where teacher scaffolding optimizes learners' internalization of AI-derived strategies, transforming cognitive potential into tangible competence gains through sociocultural mediation principles [78]. Accordingly, teacher support may function both as a moderating mechanism and a potential source of instructional variance, which future designs should disentangle.

VI. DISCUSSION AND LIMITATIONS

A. Discussion

1) AI-enabled metacognitive scaffolding and its association with translation competence

The findings provide initial support for a theory-informed model in which Doubao operationalizes Flavell's (1979) MT through a dual-cycle framework integrating AI functionalities and pedagogical mediation, rather than validating MT as a comprehensive framework. Within this structure, the MC cycle organizes translation development through planning, monitoring, and evaluation processes—such as task framing, strategy checking, and reflective revision—closely aligning with SRL routines. In

parallel, TSI supports learners in interpreting AI-generated feedback and applying it to revision decisions, particularly in cases involving cultural or pragmatic nuance, thereby facilitating the consolidation of translation competence and illustrating MT's potential in L2 pedagogy [53].

Empirically, Doubao use was associated with higher translation competence in this implementation. The experimental group demonstrated greater mean gains than the control group ($\Delta M = +6.08$ vs. $+2.48$), corresponding to a moderate effect size ($d = 0.46$), consistent with typical intervention effects reported in TELL contexts. Qualitative evidence (Table 1; $>82\%$ endorsement) further suggests that AI-supported translation may facilitate cognitive externalization, while teacher-scaffolded interaction (TSI ; 82.14%) supports strategy internalization and regulation [59]. Three interrelated pathways are indicated: functional diversity (83.93%), reflecting expanded strategy awareness; metacognitive construction (82.14%), linked to structured processing and revision; and self-efficacy reinforcement (83.93%), consistent with Albert Bandura's (1997) framework. These patterns coincide with reduced performance variability ($\Delta SD = -2.15$) and an expanded score range (38–97 vs. 28–87), suggesting more inclusive learning outcomes.

Taken together, these findings indicate that the association between AI use and translation competence is closely linked to metacognitive engagement. However, given the quasi-experimental design, the results should be interpreted as indicative of a plausible mechanism rather than definitive causal evidence.

2) Dual-path and moderated mechanisms in AI-supported translation development

The results provide evidence for a dual-path structure linking AI use to translation development, highlighting metacognitive mediation as a central explanatory mechanism. The indirect pathway through MC was statistically significant, indicating that DB is associated with TA partly via metacognitive processes ($DB \rightarrow MC$: $B = 0.58$, 95% CI [0.45, 0.71]; $MC \rightarrow TA$: $B = 0.54$, 95% CI [0.40, 0.67]). This pattern suggests that a substantial portion of the observed relationship operates through learners' engagement in planning, monitoring, and revision activities, rather than through direct AI effects alone. Consistent with sociocultural perspectives, particularly those associated with Lev Vygotsky, AI-generated feedback may externalize cues for monitoring, while instructional support facilitates their transformation into actionable revision strategies [78].

Crucially, this metacognitive pathway is not uniform but varies as a function of TSI . High TSI magnitude ($+1$ SD) was associated with a 53.5% amplification of the metacognition–translation slope, increasing the effect from 0.43 units (95% CI [0.25, 0.60]) under low TSI to 0.66 units (95% CI [0.53, 0.78]). This 0.23-unit enhancement was supported by a significant interaction effect ($MC \times TSI$: $B = 0.14$, $t = 2.99$, $p = 0.003$) and a significant index of moderated mediation (IMM = 0.08, 95% CI [0.01, 0.15]).

This pattern is consistent with the interpretation that TSI functions as a dose-dependent catalyst in the metacognitive process. Under low TSI conditions (-1 SD), learners may demonstrate limited strategy internalization and only baseline performance gains. In contrast, higher levels of TSI ($+1$ SD)

appear to support scaffolded error diagnosis, strategic revision, and affective regulation, enabling more complete enactment of metacognitive strategies [6]. This amplification effect aligns with ZPD, where guided interaction helps bridge the gap between potential and realized competence through structured support.

Taken together, these elements (DB, MC, TSI, and TA) form an integrated implementation framework linking AI affordances, teacher mediation, and metacognitive routines. Within this framework, AI provides structured feedback, metacognition organizes cognitive engagement, and teacher interaction conditions the extent to which these processes translate into observable learning outcomes. However, given the quasi-experimental design, these relationships should be interpreted as indicative of underlying mechanisms rather than definitive causal effects.

B. Limitations

Context and transferability. Beyond the core research questions, we examine four dimensions shaping Doubao's role in TELL: validity, bias control, AI limitations, and teacher capacity. Findings are grounded in a specific Chinese EFL context and should be interpreted as context-dependent, though several design features support transferability. CEFR-aligned assessment enhances comparability across proficiency contexts, while formative feedback principles (e.g., diagnostic feedback and self-regulation loops) are adaptable across settings [8, 43]. The observed moderate effect size ($d = 0.46$) indicates meaningful gains, but broader generalization requires multi-site and cross-platform replication.

Methodological rigor. Rigor was strengthened through behavioral log triangulation, pre/post-temporal anchoring, and blind external grading. CFA and model-comparison tests supported measurement validity by ruling out a dominant common-method factor. Although composite reliability was high ($CR = 0.986$), this may partly reflect conceptual proximity within tightly defined subdimensions, particularly between metacognitive strategy use and performance indicators. Discriminant validity was supported by the Fornell–Larcker criterion and cross-loading patterns, indicating acceptable construct separability despite an elevated HTMT value between MC and TA.

To further address potential item redundancy and common method bias, future research should (1) refine scale items to reduce conceptual overlap, especially between metacognitive and performance constructs; (2) incorporate multi-source data (e.g., behavioral logs, teacher ratings, and process-tracing measures) to complement self-reports; and (3) adopt design-based remedies such as temporal separation or experimental manipulation to mitigate shared method variance.

Bias and measurement limitations. Given the quasi-experimental design, unobserved learner characteristics (e.g., motivation or prior strategy use) may partially account for the observed associations. In addition, as the same instructor taught both the control and experimental groups, potential instructor-related bias cannot be entirely ruled out, although instructional content, assessment criteria, and grading procedures were standardized and supplemented with blind external evaluation. Future studies should employ

multi-instructor or randomized assignment designs to further isolate instructional effects. Gamified interface elements may also inflate perceived usefulness, as indicated by a modest gap between self-reported and logged strategy use, highlighting the need for multimodal validation.

AI and pedagogical implications. While effective for grammatical correction, Doubao shows limitations in handling culturally grounded expressions (e.g., idioms, metaphors), underscoring the need for teacher-led cultural scaffolding. Teachers can implement ‘cultural checkpoints’ (intended meaning, pragmatic function, audience expectations) and structured comparison (literal vs. functional equivalence with justification) tasks to guide revision of idioms and metaphors.

Scalability across contexts. Although conducted at a single Chinese university, the proposed mechanism is platform-agnostic and centers on teacher-mediated routines that translate AI feedback into metacognitive monitoring and revision. In high-resource settings, AI can support real-time feedback under teacher calibration; in lower-resource contexts, similar processes can be implemented via batch/offline materials. Scalability thus depends more on instructional design and teacher scaffolding than on continuous AI access.

Teacher training. Effective implementation requires targeted training rather than simple tool adoption. Three components are critical: (1) tool literacy and calibration, (2) feedback orchestration through “compare–revise–explain” routines with attention to pragmatics, and (3) metacognitive coaching linking feedback to planning–monitoring–reflection cycles [43]. Training can be delivered through workshops with exemplars and micro-teaching, supported by follow-up coaching to ensure fidelity.

VII. CONCLUSION

Results from the moderated mediation analysis suggest that Doubao utilization is positively associated with translation ability, both directly ($B = 0.18, p < 0.01$) and indirectly via metacognitive construction ($DB \rightarrow MC: B = 0.66, p < 0.001$; conditional $MC \rightarrow TA$ depends on TSI). Teacher–student interaction significantly moderated the $MC \rightarrow TA$ path ($MC \times TSI: B \approx 0.11, p < 0.01$), indicating that the association between metacognitive construction and translation performance is stronger under higher instructional support. Consistent with these patterns, the experimental group showed larger mean gains than the control group ($\Delta M = +6.08$ vs. $+2.48$), yielding a moderate effect size ($d = 0.46$).

That said, this study's robust findings have four constraints: (1) The single-institution sample ($N = 134$), despite CEFR-aligned benchmarking, limits cross-cultural generalizability and subgroup analyses by gender or proficiency; (2) The 16-week intervention cannot verify long-term strategy retention, leaving sustainability unconfirmed; (3) Doubao's internet dependency excludes resource-limited contexts, potentially worsening educational inequities [4]; (4) Teachers' specialized training, while pragmatically essential for platform implementation, may have introduced supplemental guidance beyond the AI scaffolding—a potential confounding factor that future RCT

designs should isolate through blinded teacher assignment or waitlist control protocols.

Given these insights, for educators, they should leverage Doubao's diagnostic-metacognitive synergy (error-tagging and strategy reports), as 91.07% anxiety reduction supports its affective scaffolding; for developers, they must enhance cultural nuance processing and offline functionality while retaining collaborative features (e.g., peer rubrics); for policymakers, they should embed AI-metacognitive metrics into national assessments (e.g., CET-4/6) to democratize its variability-reducing benefits ($\Delta SD = -2.15$). Looking ahead, future research prioritizes: (1) RCTs quantifying teacher training efficacy via the 53.5% *TSI* amplification slope; (2) Cross-cultural validation of contrastive rhetoric processing; (3) Longitudinal biometric tracking (e.g., EEG) of spoken discourse transfer—ultimately advancing TELL from error-correction tools to proactive competence ecosystems where AI optimizes efficiency and teachers cultivate strategic intellect.

DATA AVAILABILITY

The dataset contains information from human participants and therefore cannot be made publicly available due to privacy and ethical restrictions. De-identified data may be shared with qualified researchers upon reasonable request to the corresponding author, subject to ethics approval and a data-use agreement, where applicable.

ETHICS, CONSENT AND DATA PRIVACY

The study protocol was approved by the Human Research Ethics Review Group of the School of Foreign Languages, Hubei Business College. First-year students were informed in class and provided electronic informed consent prior to participation; written consent was not required under national and institutional regulations. Participation was voluntary, with the right to withdraw at any time. A data-minimization approach was adopted, and all data were de-identified. No personally identifiable information was included in AI-related records, and all data were stored on password-protected systems with access restricted to the research team.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

All authors have made substantial intellectual contributions to this work and approved the final version for publication. Individual contributions were as follows: Pengbiao Zhang led the conceptual design, conducted formal analysis, and wrote the original draft; Fengjunzi Wang managed project administration, provided supervision, and contributed to writing through review and editing; Yi Zhu carried out investigation and data curation; and Xinjin Song was responsible for validation and data verification. All authors had approved the final version.

FUNDING

This research was funded by the Guiding Project of the 2024 Philosophical and Social Sciences Research Project of

the Department of Education of Hubei Province, China, with the project entitled “Research and Practice on Blended Teaching of College English Empowered by Digital Intelligence in Applied Universities of Hubei Province” (Project No. 24G161).

ACKNOWLEDGEMENT

We sincerely thank the translation education scholars, AI ed-tech experts, and college English teachers who contributed to the development and validation of the AI-Metacognition (AI-MC) framework. We also thank the teachers at Hubei Business College for their participation and valuable feedback. Finally, we are grateful to the anonymous reviewers and editors for their constructive comments, which improved the quality and rigor of this manuscript.

REFERENCES

- [1] J. Chen, “Human-AI co-regulation in computer-assisted translation pedagogy,” *Computer Assisted Language Learning*, vol. 33, no. 7, pp. 743–761, 2020. doi: 10.1080/09588221.2019.1648294
- [2] Q. Zhou and K. Chen, “Measuring metacognitive calibration gaps in AI-assisted translation learning,” *Language Teaching Research*, vol. 27, no. 4, pp. 589–607, 2023. doi: 10.1177/13621688221131645
- [3] L. Zhang, W. Li, and Y. Yang, “Cognitive load reduction through AI-generated real-time feedback: Evidence from Doubao translation platform,” *Journal of Educational Technology & Society*, vol. 26, no. 2, pp. 112–127, 2023. doi: 10.30123/jets.2023.26.2.8
- [4] UNESCO, “Global survey on AI integration in language education,” *UNESCO Digital Library*, 2023.
- [5] Q. Zheng *et al.*, “Metacognitive calibration in AI-mediated language learning: A longitudinal intervention study,” *Language Learning*, vol. 73, no. 2, pp. 401–430, 2023. doi: 10.1111/lang.12589
- [6] X. Chen, L. Zhang, and Y. Li, “Teacher scaffolding in AI-enhanced language learning: A dose-response analysis of student engagement and metacognitive strategy use,” *Computers & Education*, vol. 194, 104702, 2023. doi: 10.1016/j.compedu.2022.104702
- [7] J. Chen, Y. Zhang, and X. Li, “AI-assisted language learning systems: A meta-analysis,” *Language Learning & Technology*, vol. 24, no. 3, pp. 1–24, 2020. doi: 10.125/llt.24301
- [8] J. H. Flavell, “Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry,” *American Psychologist*, vol. 34, no. 10, pp. 906–911, 1979. doi: 10.1037/0003-066X.34.10.906
- [9] X. Chen, D. Zou, H. Xie, and G. Cheng, “A decade of adaptive learning research: Models, challenges, and future directions,” *Computers & Education*, vol. 173, 104289, 2021. doi: 10.1016/j.compedu.2021.104289
- [10] J. Sweller, J. J. G. Merriënboer, and F. Paas, “Cognitive architecture and instructional design: 20 years later,” *Educational Psychology Review*, vol. 31, no. 2, pp. 261–292, 2019. doi: 10.1007/s10648-019-09465-5
- [11] R. Zhang and D. Zou, “Digital metacognitive journals for vocabulary,” *Language Learning & Technology*, vol. 26, no. 3, pp. 1–24, 2022. doi: 10.125/llt.26301
- [12] Z. Li and V. Hegelheimer, “Teacher scaffolding in AI writing environments,” *System*, vol. 115, 103050, 2023. doi: 10.1016/j.system.2023.103050
- [13] Y. Sun, T. Wang, and R. Liu, “Metalinguistic awareness in AI-mediated translation,” *Language Awareness*, vol. 31, no. 4, pp. 567–586, 2022. doi: 10.1080/09658416.2022.2085681
- [14] H. Chen and P. Lin, “The double-edged sword of AI writing assistants: Fluency gains and over-reliance risks in L2 academic writing,” *Computer Assisted Language Learning*, vol. 35, no. 8, pp. 1892–1918, 2022. doi: 10.1080/09588221.2022.2046642
- [15] M. Bond, H. Khosravi, M. De Laat, N. Bergdahl, V. Negrea, E. Oxley, P. Pham, S. W. Chong, and G. Siemens, “A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour,” *Int. J. Educational Technology in Higher Education*, vol. 21, no. 1, 4, 2024. doi: 10.1186/s41239-023-00436-z
- [16] N. Hockly, “Human-AI synergy in language teaching,” *ELT Journal*, vol. 77, no. 2, pp. 234–249, 2023. doi: 10.1093/elt/ccad008
- [17] A. Bozkurt, J. Xiao, S. Lambert, A. Pazurek, H. Crompton, S. Koseoglu, R. Farrow, M. Bond, C. Nerantzi, and S. Honeychurch, “Speculative futures on ChatGPT and generative Artificial Intelligence (AI): A collective reflection from the educational landscape,” *Asian*

- Journal of Distance Education*, vol. 18, no. 1, pp. 53–130, 2023. doi: 10.5281/zenodo.7636568
- [18] M. A. Davenport, Y. Zhao, and A. D. Wagner, “Cognitive offloading via intelligent assistants reduces hippocampal engagement in memory encoding,” *Nature Neuroscience*, vol. 26, no. 8, pp. 1424–1433, 2023. doi: 10.1038/s41593-023-01367-8
- [19] Y. Liu and L. Chen, “Pragmatic failure in Chinese-English translation,” *Language Teaching Research*, vol. 26, no. 4, pp. 589–607, 2022. doi: 10.1177/13621688221131645
- [20] M. Hu and X. Gao, “Contrastive metalinguistics for Chinese-English translation,” *Applied Linguistics*, vol. 44, no. 1, pp. 78–97, 2023. doi: 10.1093/applin/amac045
- [21] L. Chen, “Pragmatic transfer in Chinese-English translation pedagogy,” *Language and Education*, vol. 36, no. 2, pp. 145–162, 2022. doi: 10.1080/09500782.2022.2045631
- [22] H. Yu, “Digital literacy in translation pedagogy,” *Journal of China Computer-Assisted Language Learning*, vol. 3, no. 1, pp. 89–107, 2023. doi: 10.1515/jccall-2023-0005
- [23] PACTE Group, “Building a translation competence model,” *Meta*, vol. 48, no. 2, pp. 127–138, 2003. doi: 10.7202/006966ar
- [24] Q. Zhou and R. Zhang, “Metacognitive scaffolding in Doubao translation,” *Computer Assisted Language Learning*, vol. 35, no. 3, pp. 112–129, 2022. doi: 10.1080/09588221.2022.2157820
- [25] R. Li and Y. Tao, “Hybrid workflows for coherence improvement,” *Computer Assisted Language Learning*, vol. 36, no. 4, pp. 501–521, 2023. doi: 10.1080/09588221.2023.1187632
- [26] P. A. Chen, E. Jolly, J. H. Cheong, and L. J. Chang, “Artificial intelligence reduces cognitive load during decision-making and promotes redistribution of neural computation,” *Nature Human Behaviour*, vol. 6, no. 8, pp. 1164–1177, 2022. doi: 10.1038/s41562-022-01351-5
- [27] L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*, Cambridge, MA, USA: Harvard University Press, 1978.
- [28] S. Chen, “Digital scaffolding in technology-enhanced learning: A Confucian perspective,” *Educational Philosophy and Theory*, vol. 53, no. 12, pp. 1279–1291, 2021. doi: 10.1080/00131857.2020.1844148
- [29] J. Xu, “Error diagnosis protocols for syntactic interference,” *Translator Education*, vol. 8, no. 1, pp. 33–52, 2023. doi: 10.1080/24751176.2023.1998871
- [30] Y. Feng, “Growth mindset interventions for Chinese EFL translators,” *System*, vol. 110, 102934, 2022. doi: 10.1016/j.system.2022.102934
- [31] X. Wu and K. Chen, “Confucian ethics in AI translation pedagogy,” *ReCALL*, vol. 36, no. 1, pp. 45–63, 2024. doi: 10.1017/S095834402400003X
- [32] W. Fedus, B. Zoph, and N. Shazeer, “Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity,” *Journal of Machine Learning Research*, vol. 23, no. 120, pp. 1–39, 2022.
- [33] Q. Zhou, R. Zhang, and T. Liu, “Metacognitive dashboards for translation strategy training,” *Computer Assisted Language Learning*, vol. 36, no. 4, pp. 501–521, 2023. doi: 10.1080/09588221.2023.1187632
- [34] C. Chu and R. Wang, “Domain adaptation for machine translation with large-scale specialized corpora,” *Computational Linguistics*, vol. 48, no. 4, pp. 879–910, 2022. doi: 10.1162/coli_a_00445
- [35] F. Bond, M. Yamada, A. McCarthy, and M. Nădejde, “The evolution of AI-powered translation assistants: From tools to tutors,” *Computational Linguistics*, vol. 50, no. 1, pp. 197–235, 2024. doi: 10.1162/coli_a_00482
- [36] X. Wu, Z. Liu, and J. Chen, “Pedagogical machine translation benchmarking,” *ACM Trans. Asian Language Processing*, vol. 23, no. 2, pp. 1–24, 2024. doi: 10.1145/3592819
- [37] J. Song and Y. Wang, “Localizing AI ethics: Implementing Confucian virtues in educational technology,” *Int. J. Educational Technology in Higher Education*, vol. 20, no. 1, pp. 42, 2023. doi: 10.1186/s41239-023-00412-7
- [38] R. Zhang, Q. Zhou, and T. Li, “Hierarchical MoE assistants in language education,” in *Proc. 62nd Annu. Meeting of the Association for Computational Linguistics (ACL)*, vol. 1, 2024, pp. 112–130. doi: 10.18653/v1/2024.acl-long.8
- [39] I. D. Raji, T. Gebru, M. Mitchell, J. Buolamwini, J. Lee, and E. Denton, “Saving face: Investigating the ethical concerns of facial recognition auditing,” in *Proc. AAAI/ACM Conf. AI, Ethics, and Society*, 2020, pp. 145–151. doi: 10.1145/3375627.3375820
- [40] D. Yaga, P. Mell, N. Roby, and K. Scarfone, *Blockchain Technology Overview*, National Institute of Standards and Technology (NIST), Special Publication 800-222, 2019. doi: 10.6028/NIST.SP.800-222
- [41] T. O. Nelson and L. Narens, “Metamemory: A theoretical framework,” in *Psychology of Learning and Motivation*, G. H. Bower, Ed., vol. 26, pp. 125–173, 1990. doi: 10.1016/S0079-7421(08)60053-5
- [42] S. M. Fleming *et al.*, “Prefrontal mechanisms of metacognition,” *Nature Reviews Neuroscience*, vol. 23, no. 8, pp. 457–473, 2022. doi: 10.1038/s41583-022-00588-3
- [43] B. J. Zimmerman, “Becoming a self-regulated learner: An overview,” *Theory Into Practice*, vol. 41, no. 2, pp. 64–70, 2002. doi: 10.1207/s15430421tip4102_2
- [44] A. Efklides, “Metacognition and affect in MASRL,” *Educational Psychology Review*, vol. 33, no. 4, pp. 1587–1610, 2021. doi: 10.1007/s10648-021-09620-x
- [45] P. H. Winne, “Metacognition and SRL in COPES,” *Educational Psychologist*, vol. 55, no. 1, pp. 1–11, 2020. doi: 10.1080/00461520.2019.1708278
- [46] M. V. J. Veenman *et al.*, “Metacognitive interventions in STEM,” *Learning and Instruction*, vol. 88, 101789, 2023. doi: 10.1016/j.learninstruc.2023.101789
- [47] A. Boldt *et al.*, “Neural oscillations in metacognitive control,” *Science*, vol. 381, no. 6658, eadf2139, 2023. doi: 10.1126/science.adf2139
- [48] G. Schraw, “Meta-Awareness Scale (MASQ),” *Educational and Psychological Measurement*, vol. 83, no. 2, pp. 369–388, 2023. doi: 10.1177/00131644221125389
- [49] R. Pekrun and H. W. Marsh, “The control-value theory of achievement emotions: Recent advances and future directions,” *Educational Psychology Review*, vol. 35, no. 4, pp. 98, 2023. doi: 10.1007/s10648-023-09802-9
- [50] M. Li, “Teacher scaffolding in AI-enhanced language learning: A metacognitive intervention study,” *TESOL Quarterly*, vol. 58, no. 1, pp. 88–112, 2023. doi: 10.1002/tesq.3217
- [51] C. Bryant, M. Felice, Ø. E. Andersen, and T. Briscoe, “The BEA-2019 shared task on grammatical error correction,” in *Proc. 14th Workshop on Innovative Use of NLP for Building Educational Applications*, 2019, pp. 52–75. doi: 10.18653/v1/W19-4406
- [52] Q. Zheng *et al.*, “Real-time feedback systems for metacognitive regulation,” *Language Learning & Technology*, vol. 24, no. 3, pp. 68–91, 2020. doi: 10.10125/73428
- [53] J. Xu, “Human-AI synergy in metacognitive translation pedagogy,” *Translator Education*, vol. 10, no. 1, pp. 45–68, 2024. doi: 10.1080/24751176.2024.2018877
- [54] Q. Zhou and R. Zhang, “Cognitive scaffolding via AI analytics,” *Language Learning & Technology*, vol. 27, no. 3, pp. 112–135, 2023. doi: 10.125/llt.27301
- [55] J. W. Creswell and V. L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 3rd ed., Thousand Oaks, CA, USA: Sage Publications, 2017.
- [56] Y. Li, “AI-mediated L1 interference diagnosis in Chinese-English translation pedagogy,” *Computer Assisted Language Learning*, vol. 37, no. 2, pp. 210–230, 2024. doi: 10.1080/09588221.2024.2037642
- [57] T. Nguyen, R. Mitchell, and M. Mansfield, “Advancing equity in teacher education through a transformative mixed methods framework,” *Teaching and Teacher Education*, vol. 118, 103827, 2022. doi: 10.1016/j.tate.2022.103827
- [58] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006. doi: 10.1191/1478088706qp063oa
- [59] A. Efklides, “Metacognition and affect: What can metacognitive experiences tell us about the learning process?” *Educational Psychologist*, vol. 41, no. 4, pp. 277–291, 2006. doi: 10.1207/s15326985ep4104_3
- [60] A. Bandura, *Self-Efficacy: The Exercise of Control*, New York, NY, USA: W. H. Freeman, 1997.
- [61] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Hillsdale, NJ, USA: Erlbaum, 1988.
- [62] R. E. Kirk, “Practical significance: A concept whose time has come,” *Educational and Psychological Measurement*, vol. 56, no. 5, pp. 746–759, 1996. doi: 10.1177/0013164496056005002
- [63] L. Plonsky and F. L. Oswald, “How big is ‘big’? Interpreting effect sizes in L2 research,” *Language Learning*, vol. 64, no. 4, pp. 878–912, 2014. doi: 10.1111/lang.12079
- [64] L. J. Cronbach, “Coefficient alpha and the internal structure of tests,” *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951. doi: 10.1007/BF02310555
- [65] C. Fornell and D. F. Larcker, “Evaluating structural equation models with unobservable variables and measurement error,” *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50, 1981. doi: 10.1177/002224378101800104
- [66] R. B. Kline, *Principles and Practice of Structural Equation Modeling*, 4th ed., New York, NY, USA: Guilford Press, 2015.
- [67] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, 3rd ed., New York, NY, USA: McGraw-Hill, 1994.
- [68] D. L. Streiner, “Starting at the beginning: An introduction to coefficient alpha and internal consistency,” *Journal of Personality*

- Assessment*, vol. 80, no. 1, pp. 99–103, 2003. doi: 10.1207/S15327752JPA8001_18
- [69] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 5th ed., Thousand Oaks, CA, USA: Sage, 2018.
- [70] L. T. Hu and P. M. Bentler, “Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives,” *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 6, no. 1, pp. 1–55, 1999. doi: 10.1080/10705519909540118
- [71] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 8th ed., Boston, MA, USA: Cengage Learning, 2019.
- [72] J. Henseler, C. M. Ringle, and M. Sarstedt, “A new criterion for assessing discriminant validity in variance-based structural equation modeling,” *Journal of the Academy of Marketing Science*, vol. 43, no. 1, pp. 115–135, 2015. doi: 10.1007/s11747-014-0403-8
- [73] A. F. Hayes, *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, 3rd ed., New York, NY, USA: Guilford Press, 2022.
- [74] K. J. Preacher, “Advances in mediation analysis,” *Annual Review of Psychology*, vol. 66, pp. 825–852, 2015. doi: 10.1146/annurev-psych-010814-015258
- [75] D. P. MacKinnon, A. J. Fairchild, and M. S. Fritz, “Mediation analysis,” *Annual Review of Psychology*, vol. 58, no. , pp. 593–614, 2007. doi: 10.1146/annurev.psych.58.110405.085542
- [76] F. Zhao, “Human-AI collaboration in education,” *Computers & Education*, vol. 163, 104098, 2021. doi: 10.1016/j.compedu.2020.104098
- [77] L. Zheng, J. Huang, and W. Chen, “Effects of AI feedback on metacognition in L2 writing: Cognitive mechanisms and boundary conditions,” *Language Learning & Technology*, vol. 27, no. 1, pp. 68–91, 2023. doi: 10.125/llt.27168
- [78] J. P. Lantolf and M. E. Poehner, *Sociocultural Theory and the Pedagogical Imperative*, New York, NY, USA: Routledge, 2014.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).