Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Generative AI and Tertiary EFL Teachers' Pedagogical Shifts: The Dual Mediating Roles of Cognitive Regulation Scaffolds and Personalized Learning Design

Jianggang Zhang

20180103900@jsahvc.edu.cn

Jiangsu Agri-animal Husbandry Vocational College

Shuai Zhang

Yantai University

Zhuang Sun

Jilin University

Peng Tang

Jiangsu Agri-animal Husbandry Vocational College

Chunhua Zhao

Nanjing Medical University

Research Article

Keywords: Generative AI, Tertiary English Language Teaching (ELT), Cognitive Regulation Scaffolds, Personalized Learning Design, AI-TPACK

Posted Date: March 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9058296/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

Traditional tertiary English Language Teaching (ELT) has long faced an intractable dilemma: standardized, one-size-fits-all instruction fails to address the heterogeneous linguistic and cognitive needs of second language learners, while teachers' limited instructional capacity severely restricts the delivery of fine-grained cognitive scaffolding and individualized differentiated feedback. Generative Artificial Intelligence (Generative AI) offers unprecedented technical affordances to resolve these long-standing pain points, yet extant research remains predominantly student-centric, with no empirical studies unpacking the causal mediating mechanisms between Generative AI integration and sustained pedagogical change in tertiary ELT. To address this critical gap, this study employs a fully integrated sequential exploratory mixed-methods design with 205 in-service tertiary EFL teachers across 28 Chinese higher education institutions. Findings reveal that Generative AI drives systematic, multi-dimensional shifts in teachers' pedagogical practices, primarily channeled through two synergistic mediating pathways: Generative AI-enabled Cognitive Regulation Scaffolds Design (CRSD), and Generative AI-facilitated Personalized Learning Design (PLD). Quantitative analyses confirm significant independent, serial mediating, and synergistic interaction effects of the two pathways, with effects moderated by teachers' AI-specific Technological Pedagogical Content Knowledge (AI-TPACK), AI literacy, and resolution of core implementation dilemmas. Full-chain endogeneity treatment, reverse model validation, and exhaustive competitive model testing further confirm the causal order and optimality of the dual-pathway framework. This study constructs and empirically validates a revised, theory-grounded framework for higher education ELT, advances scholarship in Generative AI-enabled language education by refining foundational second language acquisition (SLA) theories, and informs evidence-based teacher professional development in the AI era.

Highlights

1. Constructs and empirically validates a hierarchically theory-grounded dual-pathway framework, one of the first studies to systematically unpack the causal mechanisms linking Generative AI integration to sustained pedagogical change among tertiary EFL teachers, with a priori theoretical justification for the serial causal ordering of dual mediators.

2. Systematically validates the independent, serial mediating, and synergistic interaction effects of CRSD and PLD, with exhaustive competitive model testing to confirm the uniqueness and statistical optimality of the dual-pathway mechanism.

3. Implements a fully integrated sequential exploratory mixed-methods design, triangulating qualitative insights on implementation dilemmas and non-transformative AI use with quantitative moderated mediation analysis, addressing the qualitative-quantitative disconnect pervasive in extant research.

4. Provides a full-chain systematic treatment of endogeneity, addressing reverse causality, omitted variable bias, and instrumental variable exclusivity constraints, establishing robust causal inference that exceeds conventional cross-sectional research standards in educational technology.

5. Refines and extends foundational SLA and educational technology theories for the Generative AI era, re-centers teacher agency to challenge dominant techno-deterministic narratives, and delivers evidence-based practical guidance for stratified teacher professional development.

1. Introduction

The global proliferation of Generative AI has triggered unprecedented transformations in educational practice, with higher education language education emerging as one of the most significantly impacted disciplines (Kasneci et al., 2023). For tertiary ELT practitioners, Large Language Models (LLMs) offer unique capabilities to automate routine instructional tasks, generate contextually relevant academic English learning materials, and provide adaptive support for diverse learners, addressing long-standing structural constraints in traditional university English classrooms (Crompton et al., 2024). These affordances are particularly salient in the Chinese tertiary ELT context, defined by large class sizes (60 + students per class), a uniform national curriculum framework, and intense high-stakes assessment pressure (CET-4/6 and TEM-4/8). For teachers seeking to implement learner-centered pedagogy in this constrained context, Generative AI represents a potentially transformative solution to long-standing capacity limits (Liu & Xiao, 2024). Recent evidence confirms that Chinese tertiary EFL teachers have shown growing willingness to integrate Generative AI into daily instruction, yet the actual pedagogical impacts of this integration remain vastly under-examined (Doo & Bonk, 2024; Li & Yu, 2025).

The rapid integration of Generative AI into tertiary ELT has outpaced academic research into its actual impact on university teachers' instructional practice. While popular discourse often frames Generative AI as either a revolutionary panacea for educational inequity or an existential threat to the teaching profession (Bozkurt et al., 2024), empirical research has yet to systematically examine how tertiary English teachers leverage these tools in daily instructional work, and how such use reshapes their core pedagogical decision-making (Chen & Tawilapakul, 2026). This gap is consequential: teachers are the ultimate arbiters of technology integration in the classroom; their instructional designs, pedagogical beliefs, and professional practices determine whether Generative AI's potential is realized, misused, or ignored entirely (Mishra et al., 2024). Without centering teachers in our analysis of AI in ELT, we cannot explain why the technology drives transformative change for some educators, while yielding only superficial efficiency gains for others.

At the core of tertiary ELT pedagogy lie two foundational instructional goals long constrained by teachers' limited time and capacity: designing cognitive regulation scaffolds that foster learners' self-regulated language learning (Zimmerman, 2000), and creating personalized learning experiences that deliver comprehensible input matched to each learner's proficiency level, academic major, and career needs (Krashen, 1985). Generative AI offers the potential to scale these two core pedagogical actions in ways previously impossible for individual tertiary English teachers (Yeh, 2024). Yet, to date, no research has empirically validated these two practices as the core causal mechanisms linking Generative AI use to transformative pedagogical change in tertiary ELT contexts, nor has research established a theoretically justified causal order between these two mechanisms. This leaves the field with no

coherent framework for understanding how, why, and under what conditions Generative AI reshapes ELT teaching practice.

The primary purpose of this study is to empirically investigate the mechanisms through which Generative AI influences the pedagogical practices of in-service tertiary EFL teachers in Chinese higher education. We move beyond simple correlational analyses to unpack the “black box” between Generative AI adoption and pedagogical change, proposing and testing a dual-pathway model that positions CRSD and PLD as core mediating variables, with a priori theoretical justification for their serial causal ordering. We further explore how these mediating pathways interact, how their effects are conditioned by teachers’ AI-TPACK, AI literacy, and resolution of implementation dilemmas, and exhaustively test alternative competitive explanations to confirm the uniqueness of our core model.

Four core research questions guide the study:

1. How do in-service tertiary EFL teachers in China perceive and utilize Generative AI Application in ELT (GAA) to design CRSD and PLD activities, and what implementation dilemmas shape this process?
2. What is the nature and multi-dimensional structure of the shifts in teachers’ core pedagogical practices resulting from sustained GAA integration?
3. To what extent does GAA predict shifts in tertiary EFL teachers' pedagogical practices, and do GAA-enabled CRSD and PLD mediate this relationship independently, serially, and synergistically, after accounting for endogeneity and competitive explanations?
4. How do teachers’ AI-TPACK, AI literacy, and resolution of implementation dilemmas moderate the strength of the indirect relationships between GAA and shifts in pedagogical practices via the dual mediating pathways?

This manuscript is organized into six sequential sections. Following this introduction, Section 2 reviews extant literature and identifies three critical research gaps. Section 3 outlines the hierarchical theoretical foundation, defines core constructs, presents the conceptual framework, and formulates testable hypotheses. Section 4 details the fully integrated sequential exploratory mixed-methods research design. Section 5 presents integrated qualitative and quantitative findings aligned directly with the four research questions. Section 6 discusses the core findings, outlines theoretical contributions and practical implications, acknowledges study limitations, and proposes a logically tied future research agenda. A brief concluding summary closes the paper. All supplementary materials (interview protocol, full survey instruments, complete statistical outputs, coding manual) are included in the Online Supplementary Materials (OSM), with in-text citations to the OSM where relevant.

2. Literature Review and Research Gaps

2.1 Current State of Research on Generative AI in Tertiary ELT

Existing research on Generative AI in tertiary ELT falls into three dominant streams, each leaving critical gaps in our understanding of AI's impact on teachers' instructional practice.

First, the overwhelming majority of empirical studies adopt a student-centric perspective, focusing on how direct learner interaction with Generative AI shapes language learning processes and outcomes. This work has demonstrated Generative AI's potential as an adaptive conversational partner for oral proficiency development (Qiao & Zhao, 2023), a feedback provider for L2 academic writing (Zhang, 2024), and a scaffold for self-regulated learning (SRL) (Chen & Alibakhshi, 2025). While invaluable for establishing Generative AI's learning potential, these studies treat the technology as a stand-alone intervention, overlooking the teacher's pivotal role in designing, orchestrating, and moderating these AI-mediated learning experiences (Lai, 2024). Even in studies of learner-AI interaction, the teacher's instructional design decisions fundamentally shape the quality of that interaction, yet this critical link between teacher design and AI learning outcomes remains unexamined in extant tertiary ELT literature.

Second, a smaller but growing body of research has shifted focus to tertiary educators, yet remains predominantly centered on teachers' attitudes, perceptions, and behavioral intentions to adopt Generative AI. Studies have explored pre-service teachers' perceptions of AI integration (Wu et al., 2025), in-service teachers' acceptance of tools like ChatGPT (Alzahrani, 2024), and barriers to adoption including academic integrity concerns and lack of technical skills (Gustilo et al., 2024). While essential for understanding the precursors to technology integration, this research overwhelmingly stops short of examining teachers' actual, sustained pedagogical behaviors after AI adoption, failing to bridge the well-documented "intention-behavior gap" in educational technology research (Celik et al., 2022). Even small-scale qualitative studies describing teachers' AI use behaviors do not quantify the magnitude of pedagogical changes or unpack the causal mechanisms driving them, leaving the "black box" of AI-driven instructional transformation unexamined.

Third, even studies that link Generative AI use to pedagogical or learning outcomes rarely provide robust, theory-driven explanations of the intermediate processes connecting cause and effect (Noroozi et al., 2024). Existing research may demonstrate a positive correlation between AI use frequency and student engagement, or between AI adoption and teacher work efficiency, but it seldom explains how and why AI drives instructional transformation (Zhao, 2025). Critically, extant studies rely on prior empirical findings to justify causal ordering of mediating pathways, rather than a priori theoretical derivation, risking circular argumentation and undermining the validity of causal claims. This gap prevents the field from moving beyond descriptive accounts of technology use to rigorous, explanatory accounts of educational impact.

2.2 Critical Research Gaps

Our systematic review identifies three interconnected, critical gaps that this study addresses:

Gap 1: Lack of empirical inquiry into AI-driven shifts in tertiary EFL teachers' core pedagogical behaviors

There is a dearth of multi-method empirical research systematically documenting how Generative AI reshapes tertiary EFL teachers' core instructional practices, including lesson planning, material design, assessment and feedback, and classroom interaction. Existing work notes anecdotal changes in teachers' workload and role perceptions, but has not unpacked the multi-dimensional nature of these pedagogical shifts, nor triangulated self-reported changes with classroom observations and instructional artifact analysis (Chen & Tawilapakul, 2026; Liu & Xiao, 2024). Furthermore, no research has systematically integrated qualitative insights on implementation barriers into quantitative model testing to explain why AI adoption fails to drive change for some teachers, leaving the field with no explanation for the heterogeneous outcomes of AI integration in ELT.

Gap 2: Unpacked causal black box of mediating mechanisms between Generative AI and pedagogical change

No existing empirical studies have proposed or tested the mediating roles of CRSD and PLD in the relationship between GAA and tertiary EFL teachers' pedagogical shifts, with theoretically justified causal ordering. Extant research is dominated by correlational analyses, with insufficient treatment of endogeneity, and no exhaustive testing of competitive mediating explanations to confirm the uniqueness of the proposed mechanisms. This gap leaves the field without a causal explanation of how and why AI drives instructional transformation in ELT, and no clear guidance for teachers on how to leverage AI for meaningful pedagogical change.

Gap 3: Absence of a theory-grounded, empirically validated framework for Generative AI-ELT integration

The field lacks a comprehensive, hierarchically cohesive framework that explains the relationship between Generative AI and pedagogical change in higher education, validated with empirical data from tertiary English teachers. Existing conceptual models of AI in language education are either descriptive, student-centric, untested in tertiary ELT settings, or rely on a theoretical patchwork of disconnected frameworks with no endogenous logical links (Li & Zhou, 2025). Such a framework is essential for guiding future research, teacher professional development, and educational policy design around AI integration, as it provides a coherent, testable model for understanding when and how Generative AI delivers on its transformative promise in ELT.

3. Theoretical Foundation, Conceptual Framework, and Research Hypotheses

3.1 Core Construct Definition and Operationalization

To ensure conceptual clarity and measurement consistency, we define and operationalize the study's core constructs, aligned with both foundational theory and our qualitative exploratory findings:

1. **Generative AI Application in Tertiary ELT (GAA)**: Sustained, pedagogically intentional use of LLM-based AI systems by in-service tertiary EFL teachers for core instructional purposes, operationalized

by the frequency, diversity, and pedagogical sophistication of tool use to co-create instructional content (Kohnke et al., 2023). This construct emphasizes teachers' active role in prompting, curating, and adapting AI outputs, distinct from passive use of pre-programmed educational software or superficial administrative use.

2. **Cognitive Regulation Scaffolds Design (CRSD)**: Teachers' use of Generative AI to create instructional supports that guide students through monitoring, directing, and regulating their own cognitive and metacognitive processes during academic English learning tasks (Zimmerman, 2000). Operationalized examples include AI-generated guided hints for reading comprehension, metacognitive reflection prompts, and structured feedback that fosters student self-correction.
3. **Personalized Learning Design (PLD)**: Teachers' use of Generative AI to tailor instructional content, activities, and feedback to individual students' proficiency levels, academic majors, career needs, and learning paces (Krashen, 1985). Operationalized examples include tiered reading texts aligned with different majors, targeted vocabulary lists based on individual error patterns, and adaptive conversational practice activities.
4. **Shifts in Pedagogical Practices**: Fundamental, sustained changes in teachers' instructional approach, operationalized across two hierarchical dimensions: (1) Surface-level behavioral change: material creation efficiency and administrative task automation; (2) Deep-level ideological change: instructional design shift from teacher-centric to student-centered planning, assessment shift from summative to formative, and teacher role reorientation from "content transmitter" to "learning architect" (Chen & Tawilapakul, 2026).
5. **Moderating Variables**: (1) **AI-TPACK**: Specialized integration of Generative AI tools with EFL pedagogy and content knowledge, the synergistic understanding of technology, pedagogy, and subject content required for effective technology integration (Mishra & Koehler, 2006); (2) **AI Literacy**: Competencies enabling teachers to critically evaluate, effectively use, and ethically engage with AI technologies in ELT (Long & Magerko, 2020); (3) **Implementation Dilemma Resolution**: Teachers' resolution of three core barriers (student over-reliance, AI hallucinations, academic integrity risks), identified in the qualitative phase as a critical boundary condition for dual-pathway activation.

3.2 Hierarchical Theoretical Foundation

This study integrates four foundational educational theories into a hierarchically cohesive, endogenous theoretical system, extending each theory to the Generative AI era. The four theories are structured into three logically nested layers, with a unified core logic: Generative AI resolves long-standing structural constraints in tertiary ELT by amplifying teachers' capacity to enact learner-centered pedagogy, contingent on their integrated professional knowledge.

Layer 1: Foundational Logic Layer – Defining Core Tertiary ELT Pain Points

Vygotsky's (1978) social constructivist theory (Zone of Proximal Development, ZPD) and Krashen's (1985) Input Hypothesis (i + 1) form the foundational logic, collectively defining the two intractable pain points of traditional tertiary ELT that Generative AI is uniquely positioned to resolve. Vygotsky's ZPD

theory posits that learning occurs through guided support within the gap between a learner's independent capacity and potential capacity with scaffolding; traditional large-class ELT cannot deliver contingent, individualized scaffolding to every learner's ZPD, a constraint Generative AI can alleviate at scale. Krashen's Input Hypothesis argues that successful second language acquisition requires consistent exposure to comprehensible input slightly beyond the learner's current level (i + 1), matched to their interests and needs; traditional standardized instruction cannot deliver individualized i + 1 input to diverse learners, a core pain point Generative AI addresses. Together, these two theories define the non-negotiable core of effective ELT: individualized cognitive scaffolding and personalized comprehensible input, establishing why CRSD and PLD are the only viable core mediating mechanisms.

Layer 2: Mechanism Explanation Layer – Defining How AI Drives Pedagogical Change

Zimmerman's (2000) SRL Theory forms the mechanism explanation layer, linking the foundational pain points to actionable teacher practice, and providing a priori theoretical justification for the serial causal order of the dual mediating pathways. SRL theory emphasizes that long-term language acquisition requires learners to develop metacognitive, motivational, and behavioral self-regulation skills. CRSD directly targets the development of these skills, as teachers use AI to design scaffolds that guide students to monitor, evaluate, and regulate their own learning processes.

Critically, this layer provides the theoretical derivation for the serial pathway *GAA → CRSD → PLD → Pedagogical Shifts*, eliminating circular argumentation. Designing CRSD requires teachers to engage in a deep, learner-centric cognitive process: they must first deconstruct the linguistic and cognitive demands of a task, map individual learners' ZPDs and metacognitive gaps, and design targeted scaffolds to bridge those gaps. This process builds the pedagogical reasoning capacity and nuanced understanding of learner variability that is a necessary prerequisite for high-quality PLD. In other words, teachers cannot design truly effective personalized learning experiences without first understanding the individual cognitive and metacognitive needs of their learners – a capacity built through CRSD design. This hierarchical logic provides an unassailable a priori theoretical justification for the serial causal order of the dual mediators.

Layer 3: Boundary Condition Layer – Defining When AI Drives Pedagogical Change

Mishra & Koehler's (2006) TPACK Framework forms the boundary condition layer, defining the individual difference factors that determine whether teachers can translate GAA into effective CRSD and PLD. The framework argues that effective technology integration requires a synergistic synthesis of technology, pedagogy, and content knowledge – not isolated technical skills. This layer explains why superficial AI use fails to drive pedagogical change: without integrated AI-TPACK and AI literacy, teachers cannot leverage AI to design high-quality CRSD and PLD, regardless of how frequently they use the tool. This layer also extends the TPACK framework to the Generative AI era, by validating AI-TPACK and AI literacy as critical moderators of the dual mediating pathways.

3.3 Rationale for Synergistic Interaction Effects

The synergistic interaction effect between CRSD and PLD is theoretically justified by the complementary nature of the two pathways. PLD ensures that academic English input is matched to the learner's i + 1 level, while CRSD equips the learner with the metacognitive skills to process that input effectively and build long-term learning autonomy. Neither pathway alone can maximize pedagogical transformation: personalized content without metacognitive scaffolds fails to build sustainable learner autonomy, while generic cognitive scaffolds without personalized adaptation do not address individual learner needs. Thus, the two pathways have a mutually reinforcing, synergistic effect on shifts in teachers' pedagogical practices (Yeh, 2024).

3.4 Conceptual Framework

Based on the hierarchical theoretical foundation, we propose a dual-pathway moderated mediation conceptual framework. The framework posits that GAA has a direct effect on Shifts in Pedagogical Practices, but that this effect is primarily mediated through the dual pathways of CRSD and PLD. These pathways function independently, in a theoretically justified serial sequence, and interactively. The entire model is moderated by teachers' AI-TPACK, AI literacy, and implementation dilemma resolution, with control variables included on all endogenous variables to address omitted variable bias. The full conceptual framework is presented in Fig. 1.

3.5 Research Hypotheses

Based on the conceptual framework and hierarchical theoretical rationale, we formulate six testable hypotheses:

- **H1**: GAA has a significant positive direct effect on shifts in tertiary EFL teachers' pedagogical practices.
- **H2**: CRSD significantly mediates the relationship between GAA and shifts in tertiary EFL teachers' pedagogical practices.
- **H3**: PLD significantly mediates the relationship between GAA and shifts in tertiary EFL teachers' pedagogical practices.
- **H4**: CRSD and PLD play a significant serial mediating role in the relationship between GAA and shifts in tertiary EFL teachers' pedagogical practices, following the theoretically justified order *GAA → CRSD → PLD → Pedagogical Shifts*.
- **H5**: There is a significant positive synergistic interaction effect between CRSD and PLD on shifts in tertiary EFL teachers' pedagogical practices.
- **H6**: Teachers' AI-TPACK level, AI literacy, and implementation dilemma resolution significantly moderate the strength of the indirect relationships between GAA and shifts in pedagogical practices via the dual mediating pathways.

4. Research Design and Methodology

4.1 Overall Research Design

A fully integrated sequential exploratory mixed-methods design (QUAL → QUAN) was selected to address the research questions (Creswell & Clark, 2017). This design ensures full integration of qualitative and quantitative phases at every step of the research process:

1. **Phase 1 (Qualitative Exploration)**: Answered RQ1 and RQ2 via semi-structured interviews, classroom observations, and instructional artifact analysis. This phase developed survey items for the quantitative phase, and identified core implementation dilemmas, counterexamples of non-transformative AI use, and boundary conditions directly incorporated into the quantitative model specification.
2. **Phase 2 (Quantitative Validation)**: Answered RQ3 and RQ4 via a large-scale survey, with model specifications directly informed by Phase 1 qualitative findings.

This fully integrated design eliminates the qualitative-quantitative disconnect in extant research, ensuring the two phases form a logically cohesive, mutually reinforcing analytical process. Full methodological details (pilot study, instrument development, analytical protocols) are presented in OSM Appendices A-D.

4.2 Participants and Ethical Approval

4.2.1 Qualitative Sample

A purposive maximum variation sampling strategy recruited 15 in-service tertiary EFL teachers from 6 provinces across eastern, central, and western China. Participants met three criteria: (1) minimum 3 months of Generative AI use in ELT; (2) full-time in-service status in Chinese higher education institutions; (3) variation in teaching experience (1–5 years, 6–15 years, 16 + years). Thematic saturation was reached at the 12th interview, with 3 additional interviews conducted for cross-validation. Demographic characteristics are presented in OSM Table 2.

4.2.2 Quantitative Sample

A stratified random sampling approach was used, with stratification variables including geographic region, university type, and teaching experience. The sampling frame consisted of in-service tertiary EFL teachers registered in national and provincial university teacher professional development networks across China. A total of 320 surveys were distributed, with 205 valid responses retained after data cleaning (effective response rate = 64.06%). The final sample included teachers from 28 higher education institutions, with teaching experience ranging from 1 to 32 years (M = 9.14, SD = 6.57). A priori sample size calculation via G*Power 3.1 confirmed a minimum required sample size of 138 (f^2=0.15, α = 0.05, power = 0.95, 12 predictors), with the final sample providing sufficient statistical power for all primary analyses. Full demographic characteristics are presented in OSM Table 3.

4.2.3 Ethical Approval

This research received full ethical approval from the Institutional Review Board (IRB) of Jilin University, P.R. China (IRB Reference No.: 2025-EDU-AI-0119; Approval Date: 02/04/2025). All procedures were conducted in strict accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to data collection, with all data anonymized and stored on a secure, password-encrypted university server.

4.3 Research Instruments

4.3.1 Qualitative Instruments

Three complementary qualitative instruments were used: a 12-item semi-structured interview protocol, a systematic instructional artifact analysis guide, and a 5-domain classroom observation protocol. Full instruments are presented in OSM Appendices A-C.

4.3.2 Quantitative Instruments

The survey consisted of eight sections measuring core variables on a 7-point Likert scale (1 = Strongly Disagree, 7 = Strongly Agree). All scales were adapted from previously validated instruments, refined based on Phase 1 qualitative findings, and validated via a 3-round expert review process to ensure content validity. Full survey items and psychometric properties are presented in OSM Appendix D. Scales included: GAA (5 items, $\alpha = .91$), CRSD (6 items, $\alpha = .90$), PLD (6 items, $\alpha = .92$), Shifts in Pedagogical Practices (8 items, $\alpha = .93$), AI-TPACK (12 items, $\alpha = .90$), AI Literacy (10 items, $\alpha = .88$), Implementation Dilemma Resolution (3 items, $\alpha = .86$), and demographic/control variables.

4.4 Data Collection and Analytical Strategies

In Phase 1, interviews were conducted via secure video conference between September and October 2025, audio-recorded with participant consent, and transcribed verbatim within 48 hours. Classroom observations and instructional artifact collection were conducted concurrently. Qualitative data were analyzed using Braun & Clarke's (2006) six-phase reflexive thematic analysis, managed via NVivo 20, with double-blind coding achieving a Cohen's Kappa coefficient of 0.88 (excellent reliability).

In Phase 2, the survey was distributed via national and provincial university teacher professional development networks, with data collection spanning October to November 2025. Quantitative data were analyzed using SPSS 28.0 and AMOS 26.0, with mediation and moderated mediation analyses conducted using SPSS PROCESS Macro 4.2 (Hayes, 2018). Analysis proceeded in eight sequential steps: (1) preliminary analyses and common method bias testing; (2) measurement model validation via confirmatory factor analysis (CFA); (3) structural model testing via structural equation modeling (SEM); (4) mediation effect testing via bootstrap analysis; (5) interaction and moderation analysis; (6) multi-group heterogeneity analysis; (7) competitive model validation; (8) endogeneity and robustness testing.

5. Results

5.1 Qualitative Results (Answering RQ1 and RQ2)

Thematic analysis of interview, observation, and artifact data yielded rich, triangulated insights into teachers' AI implementation and resultant pedagogical shifts, with thematic saturation reached at the 12th interview.

5.1.1 Dual Mediating Pathways and Implementation Dilemmas (RQ1)

Teachers conceptualized and implemented the two mediating pathways in distinct but complementary ways, with clear conceptual boundaries and mutually reinforcing practices.

Pathway 1: Generative AI as a "Metacognitive Dialogue Partner" for CRSD

Teachers used Generative AI not just to provide academic content answers, but to generate questions and prompts that scaffold students' metacognitive and cognitive regulation processes during academic English learning. Three interrelated subthemes emerged: (1) deconstruction of academic linguistic and task complexity, breaking down complex EAP reading and writing tasks into sequential steps aligned with students' ZPD; (2) prompting student self-correction and metacognitive reflection via Socratic-style feedback; (3) process-oriented learning support for academic English, including checklists and progress trackers for complex language tasks. Artifact analysis confirmed that 85% of collected lesson plans included Generative AI-generated process scaffolds for students' self-regulated academic English learning.

Pathway 2: Generative AI as a "Personalization Engine" for Differentiated Instruction

Teachers leveraged Generative AI to move beyond one-size-fits-all whole-class instruction and cater to individual learner variability at scale, operationalizing Krashen's i + 1 hypothesis for diverse learners. Three subthemes emerged: (1) tiered material generation for diverse proficiency and major groups; (2) major-aligned and career-focused content creation, adapting learning content to increase the relevance of comprehensible input; (3) adaptive practice and targeted remediation for high-stakes assessments, generating mini-lessons focused on individual students' specific error patterns.

Core Implementation Dilemmas and Critical Counterexamples

Teachers identified three persistent, interconnected implementation dilemmas that directly shaped their ability to activate the dual mediating pathways: (1) Student Over-Reliance vs. Learning Autonomy; (2) AI Hallucinations and Output Inaccuracy; (3) Ethical and Academic Integrity Risks. These findings directly informed the inclusion of implementation dilemma resolution as a moderator in the quantitative model. Two interviewed teachers provided clear counterexamples: they used Generative AI regularly (3–4 times weekly) but reported no meaningful pedagogical shifts, with AI use limited to superficial administrative tasks and no CRSD/PLD implementation. Further analysis revealed two core reasons for implementation

failure: (1) extremely low AI-TPACK and AI literacy, with no formal AI professional development; (2) unresolved implementation dilemmas, leading to risk-averse, non-instructional AI use. These cases informed the multi-group heterogeneity analysis in the quantitative phase.

5.1.2 Multi-Dimensional Pedagogical Shifts (RQ2)

The integration of Generative AI precipitated three fundamental, interconnected shifts in teachers' core pedagogical practices, with a clear distinction between surface-level behavioral change and deep-level ideological change:

1. Deep-level ideological shift: From content deliverer to learning architect, with teachers focusing on designing AI-mediated learning experiences rather than delivering static content;
2. Deep-level ideological shift: From static semester plans to dynamic, real-time instruction, adapting materials and activities in response to formative student understanding;
3. Surface-level behavioral shift: From delayed summative assessment to ongoing, formative dialogue, shifting feedback from a one-time summative event to a continuous conversation between student, AI, and teacher.

Notably, only teachers who activated the dual mediating pathways reported deep-level ideological shifts; teachers with only administrative AI use reported only minor surface-level efficiency gains. This finding informed the split analysis of surface vs. deep pedagogical change in the quantitative phase.

5.2 Quantitative Results (Hypothesis Testing, Answering RQ3 and RQ4)

5.2.1 Preliminary Analyses and Measurement Model Validation

Descriptive statistics revealed that teachers in the sample used GAA with moderate frequency ($M = 4.18$, $SD = 1.59$ on the 7-point scale). All core variables were significantly and positively correlated ($p < .001$), with correlation coefficients ranging from .31 to .69, providing preliminary support for the hypothesized relationships. Discriminant validity was confirmed via the Fornell-Larcker criterion. Harman's single-factor test revealed the first unrotated factor explained 32.45% of the total variance (below the 40% threshold), confirming no severe common method bias. Full descriptive statistics and common method bias testing results are presented in OSM Tables 4–5.

The CFA for the seven-factor measurement model demonstrated excellent fit to the data, with all fit indices meeting recommended thresholds (Hair et al., 2019): $\chi^2/df = 2.21$, CFI = .97, TLI = .96, RMSEA = .047 [.040, .054], SRMR = .039. All factor loadings were statistically significant ($p < .001$) and greater than .70. Convergent validity was established, with AVE values for all constructs ranging from .62 to .78 (all > .50) and CR values ranging from .90 to .95 (all > .70). Full reliability and convergent validity results are presented in OSM Table 6.

5.2.2 Structural Model and Direct Effect Testing (H1 Validation)

The hypothesized structural model showed excellent fit to the data: $\chi^2/df = 2.32$, CFI = .96, TLI = .95, RMSEA = .049 [.043, .056], SRMR = .043. The direct effect of GAA on shifts in pedagogical practices, after accounting for the two mediators, was significant and positive ($\beta = .20$, SE = .06, $p < .001$). Split analysis revealed that GAA had a significant direct effect on deep-level ideological change ($\beta = .24$, SE = .05, $p < .001$), but a non-significant direct effect on surface-level behavioral change ($\beta = .08$, SE = .07, $p = .25$). This provides strong statistical support for H1, and confirms that GAA primarily drives transformative ideological change, rather than superficial efficiency gains. Standardized path coefficients for the hypothesized structural model are presented in Fig. 2.

5.2.3 Independent and Serial Mediating Effect Testing (H2-H4 Validation)

SEM-based bootstrap analysis with 5,000 resamples revealed the following significant indirect effects, with results cross-validated via the SPSS PROCESS macro. Full bootstrap mediation test results are reported in Table 1.

Table 1
Bootstrap Mediation Effect Test Results

Pathway	Effect	SE	95% Bias-Corrected CI	Hypothesis	Result
Direct Effect: GAA → Pedagogical Shifts	.20	.06	[.09, .31]	H1	Supported
Indirect Effect 1: GAA → CRSD → Pedagogical Shifts	.17	.04	[.10, .25]	H2	Supported
Indirect Effect 2: GAA → PLD → Pedagogical Shifts	.23	.04	[.15, .31]	H3	Supported
Indirect Effect 3: GAA → CRSD → PLD → Pedagogical Shifts	.08	.02	[.05, .12]	H4	Supported
Total Indirect Effect	.48	.05	[.38, .58]	-	-
Total Effect	.68	.05	[.58, .78]	-	-

Note. N = 205. Bootstrap sample size = 5,000. An effect is statistically significant if its 95% CI does not include zero. The total indirect effect accounts for 70.59% of the total effect of GAA on pedagogical shifts.

H2, H3, and H4 were fully supported: the independent indirect paths through CRSD and PLD, and the serial mediation path, were all statistically significant (95% CIs did not include zero). Redundant mediator analysis confirmed that CRSD and PLD had no significant redundant explanatory power (VIF < 3.0 for both mediators), with each contributing unique variance to the prediction of pedagogical shifts. The total indirect effect accounted for 70.59% of the total effect, indicating that over 70% of GAA's impact on pedagogical shifts is realized through the dual mediating pathways.

5.2.4 Synergistic Interaction Effect Testing (H5 Validation)

Hierarchical regression analysis confirmed that, after controlling for demographic variables and the main effects of the two mediators, the interaction term (CRSD × PLD) was a significant positive predictor of pedagogical shifts ($\beta = .14$, SE = .04, $p < .01$, $\Delta R^2 = .02$). Simple slopes analysis showed the positive effect of PLD on pedagogical shifts was significantly stronger for teachers with high levels of CRSD (+ 1 SD: $\beta = .41$, $p < .001$) than for those with low levels (-1 SD: $\beta = .20$, $p < .001$). This provides strong support for H5. Simple slopes analysis results are presented in Fig. 3.

5.2.5 Moderation Effect Testing (H6 Validation)

Moderated mediation analysis confirmed that AI-TPACK, AI literacy, and implementation dilemma resolution significantly moderated the total indirect effect of GAA on pedagogical shifts via the dual pathways. For all three moderators, the total indirect effect was significantly stronger for teachers with high levels of the moderator, with significant indices of moderated mediation (95% CIs no zero). Full moderated mediation test results are presented in Table 2.

Table 2
Moderated Mediation Effect Test Results (Total Indirect Effect)

Moderator	Conditional Indirect Effect	SE	95% Bias-Corrected CI	Index of Moderated Mediation	95% CI
AI-TPACK	High (+ 1 SD): .56	.07	[.42, .70]	.11	[.05, .17]
	Low (-1 SD): .31	.07	[.17, .45]		
AI Literacy	High (+ 1 SD): .53	.07	[.39, .67]	.09	[.03, .15]
	Low (-1 SD): .34	.07	[.20, .48]		
Implementation Dilemma Resolution	High (+ 1 SD): .51	.07	[.37, .65]	.08	[.02, .14]
	Low (-1 SD): .36	.07	[.22, .50]		
Note. N = 205. Bootstrap sample size = 5,000.					

Disaggregated analysis revealed nuanced mechanistic differences across the moderators: AI-TPACK exerted the strongest moderating effect on the independent PLD pathway; AI Literacy most strongly

moderated the serial mediating pathway; Implementation Dilemma Resolution had the most pronounced effect on the student-facing PLD pathway. There was no significant moderating effect on the direct path from GAA to pedagogical shifts ($p > .05$) for any moderator. This provides full support for H6. Combined simple slopes for the three moderators are presented in Fig. 4.

5.2.6 Heterogeneity, Competitive Model, and Endogeneity Testing

Multi-group SEM analysis compared the model fit between the "teaching-focused AI use" group (n = 162) and "administrative-focused AI use" group (n = 43). Results revealed that the hypothesized model had excellent fit for the teaching-focused group ($\chi^2/df = 2.28$, CFI = .96, RMSEA = .048), but poor fit for the administrative-focused group ($\chi^2/df = 3.57$, CFI = .88, RMSEA = .082). The total indirect effect was significant for the teaching-focused group (Indirect Effect = .52, 95% CI [.41, .63]), but non-significant for the administrative-focused group (Indirect Effect = .09, 95% CI [-.03, .21]). This confirms that only pedagogically intentional AI use that activates the dual mediating pathways drives pedagogical change. Full multi-group SEM outputs are presented in OSM Table 11.

To validate the causal order and optimality of the hypothesized model, we systematically compared it with five alternative competitive models. The hypothesized model demonstrated significantly better fit to the data than all alternative models, with all competitive models showing non-significant indirect effects via alternative pathways. Full model fit comparison is presented in Table 3.

Table 3
Model Fit Comparison: Hypothesized vs. Competitive Models

Model	χ^2/df	CFI	TLI	RMSEA (90% CI)	SRMR	Core Indirect Effect Significance
Hypothesized Dual-Pathway Model	2.32	.96	.95	.049 [.043, .056]	.043	Significant
Reverse Serial Model (GAA → PLD → CRSD → Pedagogical Shifts)	3.18	.91	.90	.078 [.071, .085]	.072	Non-significant
Single Mediator Model (CRSD Only)	2.87	.93	.92	.062 [.055, .069]	.058	Significant, poorer fit
Single Mediator Model (PLD Only)	2.79	.93	.92	.060 [.053, .067]	.056	Significant, poorer fit
Competitive Model 1: Additional Mediator (Workload Reduction)	2.41	.95	.94	.052 [.046, .059]	.048	Non-significant alternative effect
Competitive Model 2: Additional Mediator (Teaching Self-Efficacy)	2.45	.95	.94	.053 [.047, .060]	.047	Non-significant alternative effect

Full competitive model outputs are presented in OSM Table 9.

Full-chain endogeneity and robustness testing was conducted to address core causal inference threats. Timeline analysis of retrospective survey items revealed that 89.27% of teachers who reported systematic pedagogical change began using GAA before their pedagogical shift occurred, directly addressing reverse causality concerns. Instrumental Variable (IV) 2SLS regression using “average GAA level of other teachers in the same university” (gold-standard IV in educational research) further confirmed that, after controlling for endogeneity, the positive predictive effect of GAA on pedagogical shifts remained significant ($\beta = 0.66$, $SE = 0.07$, $p < 0.001$), with the direction and significance of the coefficient fully consistent with the baseline model. Omitted variable bias testing and split-sample validation further confirmed the stability and generalizability of the model. Full 2SLS regression results are presented in OSM Table 10.

5.3 Integrated Qualitative and Quantitative Findings

Triangulation of qualitative themes and quantitative results provides a holistic narrative of the study’s core conclusions, presented in Table 4.

Table 4
Joint Display of Integrated Qualitative and Quantitative Findings

Integrated Finding	Qualitative Evidence (Themes & Representative Quotes)	Quantitative Evidence (Model Testing & Statistical Results)
1. Generative AI drives the design of cognitive scaffolds fostering students' self-regulated language learning	"I use ChatGPT to turn my academic writing instructions into a step-by-step checklist. It helps my students monitor their own progress on an essay instead of constantly asking me 'What's next?'" (Teacher C, Comprehensive University)	H2 Supported: Significant independent mediation via CRSD (Indirect Effect = .17, $p < .001$)
2. Generative AI enables scalable, deep personalization of ELT materials and pathways for diverse learners	"With Generative AI, differentiation is no longer just a theory we learn in professional development. I can create five different versions of a CET-4 homework assignment in minutes, each tailored to a different major and proficiency group." (Teacher M, Undergraduate College)	H3 Supported: Significant independent mediation via PLD (Indirect Effect = .23, $p < .001$)
3. The two pathways are synergistic; CRSD builds pedagogical capacity that enhances effective PLD implementation	"When I use Generative AI to build a good cognitive scaffold—like a step-by-step guide for academic reading comprehension—it makes it easier for me to then personalize that guide for different majors and proficiency levels. The two go hand-in-hand; you can't do one well without the other." (Teacher L, Vocational College)	H4 & H5 Supported: Significant serial mediation path (Indirect Effect = .08, $p < .001$) and synergistic interaction effect ($\beta = .14$, $p < .01$)
4. GAA primarily drives deep ideological pedagogical change, not just superficial behavioral adjustment	"My value to my students is no longer in knowing all the answers about English grammar or academic writing. My value is in asking the right questions—both to my students and to the Generative AI. I'm designing an entire ecosystem for their academic English learning, not just delivering content." (Teacher F, Comprehensive University)	H1 Supported: Significant direct effect on deep-level ideological change ($\beta = .24$, $p < .001$)
5. The effectiveness of GAA is contingent on teachers' integrated knowledge and implementation dilemma resolution	"Some of my university colleagues just use Generative AI to make their PowerPoints prettier or write worksheets faster. They don't have the pedagogical vision to use it well for college English teaching. You need to know why you're using the tool, not just how to click the buttons." (Teacher P, Comprehensive University)	H6 Supported: AI-TPACK, AI Literacy, and implementation dilemma resolution significantly moderated the total indirect effect ($p < .001$)

Integrated Finding	Qualitative Evidence (Themes & Representative Quotes)	Quantitative Evidence (Model Testing & Statistical Results)
6. Superficial administrative AI use (without dual pathway activation) does not drive pedagogical change	Counterexample teachers used AI regularly for administrative tasks, but reported no changes to their instructional approach, with no CRSD or PLD implementation	Multi-group SEM confirmed non-significant indirect effect for administrative-focused AI use group; competitive model testing confirmed non-significant effects of auxiliary mediators

6. Discussion

6.1 Core Findings Synthesis

This study provides robust, mixed-methods evidence that Generative AI is a potent catalyst for pedagogical change among in-service tertiary EFL teachers in Chinese higher education, but its impact is not automatic or deterministic. The core theoretical insight emerging from the findings is that Generative AI functions as a pedagogical amplifier: it magnifies the impact of high-quality, learner-centered instructional design, but cannot compensate for weak pedagogical reasoning or a lack of integrated technological, pedagogical, and content knowledge. This finding directly challenges the two dominant, reductionist narratives that pervade both popular and academic discourse around Generative AI in education: the techno-utopian narrative that frames AI as an automatic “silver bullet” for educational inequity, and the techno-dystopian narrative that frames AI as an existential threat to teacher agency (Bozkurt et al., 2024).

Instead, our findings demonstrate that the transformative potential of Generative AI is realized primarily through two synergistic mediating pathways: GAA-enabled CRSD and GAA-facilitated PLD, which together account for over 70% of GAA’s total impact on pedagogical change. The two pathways operate in a theoretically justified serial causal sequence: designing CRSD builds the pedagogical reasoning capacity and understanding of learner variability that is a necessary prerequisite for high-quality PLD. The effectiveness of these pathways is significantly moderated by teachers’ AI-TPACK, AI literacy, and resolution of core implementation dilemmas: teachers with higher levels of integrated professional knowledge and resolved implementation barriers are exponentially more likely to translate GAA into transformative pedagogical change. Critically, only pedagogically intentional AI use that activates the dual mediating pathways drives deep-level ideological change; superficial administrative AI use has no significant impact on teachers’ core pedagogical practices.

6.2 Critical Dialogue with Existing Literature

This study makes three key contributions to the existing literature, addressing the three core research gaps identified in the review.

First, this study addresses the long-standing student-centric bias in educational technology research (Zawacki-Richter et al., 2019) by re-centering the tertiary EFL teacher as the primary unit of analysis. While prior research has documented anecdotal changes in teachers' role perceptions (Bower et al., 2024; Mishra et al., 2024), our study is the first to quantify the magnitude of multi-dimensional shifts in EFL teachers' instructional design, material creation, feedback practices, and professional identity, and distinguish between superficial behavioral adjustment and deep ideological transformation. Our findings also directly explain the contradictory results from Hassan & Islam (2024), who found that many teachers fail to translate AI adoption into meaningful pedagogical change: we demonstrate that only when teachers use AI to design CRSD and PLD does the technology drive transformation; superficial, administrative use of AI does not activate the dual mediating pathways, resulting in no meaningful pedagogical change.

Second, this study unpacks the causal "black box" between Generative AI use and pedagogical change, addressing a key limitation of existing research that has been dominated by correlational analyses (Noroozi et al., 2024; Hew et al., 2024). While prior work has noted AI's potential to support personalized learning (Pesovski et al., 2024) and self-regulated learning (Dahri et al., 2024), our study is the first to integrate these two previously siloed constructs into a cohesive, theoretically justified framework. We provide a priori theoretical derivation of the serial causal order of the dual mediators, eliminating the circular argumentation of relying on prior empirical findings to justify causal ordering. This is a critical theoretical advance, providing the field with a unified, causal lens for understanding AI's impact on ELT.

Third, this study extends the TPACK framework to the Generative AI era, validating AI-TPACK, AI literacy, and implementation dilemma resolution as critical boundary conditions for effective AI integration in ELT. Our findings challenge the pervasive narrative that basic "button training" on AI tools is sufficient to drive educational change, demonstrating that integrated pedagogical and technological knowledge, rather than isolated technical skills, is the key determinant of effective AI integration. This aligns with prior research on TPACK in ELT (Zhang & Chen, 2022; Bostancıoğlu & Handley, 2018), but extends this work in two key ways: (1) we demonstrate that AI literacy interacts synergistically with AI-TPACK to moderate the effectiveness of AI integration; (2) we validate implementation dilemma resolution as a critical contextual moderator, as even teachers with high AI-TPACK and AI literacy cannot activate the dual pathways if they have not resolved core ethical and practical barriers to student-facing AI use.

6.3 Theoretical Contributions

This study makes three core, substantive theoretical contributions to the field of ELT and AI education:

1. The Hierarchically Integrated Dual-Pathway Framework of AI-Driven Pedagogical Transformation for Tertiary ELT

This study constructs and empirically validates the first hierarchically cohesive, theory-grounded, context-specific framework for understanding Generative AI integration in tertiary ELT. The framework

moves beyond descriptive accounts of technology use to provide a mechanistic, causal explanation of how AI drives pedagogical change, integrating four foundational educational theories into a logically nested, endogenous system. We provide a priori theoretical justification for the serial causal order of the dual mediators, and exhaustively test competitive models to confirm the uniqueness of the framework. This contribution provides the field with a robust, generalizable theoretical lens for future research on Generative AI in tertiary language education.

2. Refined Theoretical Propositions for Foundational SLA Theories in the Generative AI Era

This study refines and extends two foundational SLA theories, addressing long-standing theoretical limitations via empirical validation in the Generative AI context. First, we propose a revised Input Hypothesis Proposition: comprehensible input (i + 1) for successful second language acquisition requires two complementary, AI-scalable conditions: (1) personalized input matched to the learner's proficiency, academic background, and interests; and (2) concurrent cognitive regulation scaffolds that foster the metacognitive skills needed to process that input effectively. This directly addresses the long-standing limitation of Krashen's Input Hypothesis, which has been criticized for focusing exclusively on input while neglecting the learner's cognitive processing capacity. Second, we propose an Extended SRL Theory Proposition: Generative AI enables a shift from teacher-delivered SRL scaffolding to co-designed, scalable SRL environments, where the teacher acts as the architect of AI-mediated metacognitive supports, rather than the sole provider of scaffolding. This extends Zimmerman's SRL theory to the Generative AI era, providing a theoretical framework for understanding how AI can foster sustainable learner autonomy.

3. A Non-Technocentric Theory of Educational Technology Integration in Higher Education

This study challenges the technocentric assumptions that pervade much of the existing research on AI in higher education, demonstrating that technology itself is not a deterministic driver of change. Instead, we show that the transformative impact of Generative AI is entirely contingent on teachers' intentional instructional design decisions and integrated professional knowledge. This contribution provides a much-needed corrective to techno-utopian and techno-dystopian narratives, offering a nuanced, evidence-based theory of educational technology integration that centers teacher agency and pedagogical reasoning. This theory directly addresses the long-standing "AI vs. teacher" narrative in the field, empirically demonstrating that Generative AI cannot replace teachers, but can only amplify the impact of high-quality, learner-centered teaching.

6.4 Evidence-Driven Practical Implications

All practical strategies proposed in this study are 1:1 aligned with the study's empirical findings. Based directly on the study's core findings, we provide three targeted, evidence-based practical implications for stakeholders in tertiary ELT.

First, for in-service tertiary EFL teachers: adopt a sequential, dual-pathway approach to GAA, aligned with the study's serial mediation finding that CRSD is a necessary prerequisite for high-quality PLD. Novice teachers should prioritize foundational CRSD first, using pre-validated prompt templates to design standardized cognitive regulation scaffolds, building the pedagogical reasoning capacity needed for effective PLD. Proficient teachers should adopt the full dual-pathway integration approach, leveraging existing pedagogical expertise to design synergistic CRSD and PLD activities, following the validated serial causal sequence. Expert teachers should focus on advanced, transformative integration of the two pathways, and act as peer mentors for less experienced teachers. All teachers should implement a "scaffolded release" approach to student AI use, a standardized 3-step validation process for AI outputs, and clear classroom guidelines for ethical AI use to address the three core implementation dilemmas.

Second, for tertiary teacher educators and professional development providers: university teacher education and professional development programs must be fundamentally redesigned around the dual-pathway framework, with a focus on integrated AI-TPACK development. Programs should shift from decontextualized "button training" on AI tools to situated, content-specific training that focuses on using Generative AI to solve core problems of practice in tertiary ELT—specifically, designing CRSD and PLD. Training should be structured around the validated serial causal sequence, with sequential modules on Generative AI for CRSD, Generative AI for PLD, and synergistic integration of the two pathways. Programs should prioritize critical AI literacy and implementation dilemma resolution training, and differentiate professional development by teacher career stage and AI literacy level.

Third, for university administrators and educational policymakers: higher education leaders should foster a supportive ecosystem for responsible, pedagogically sound GAA in tertiary ELT. Policymakers and administrators should prioritize human capacity building over infrastructure investment: while providing access to premium AI tools is necessary, the study's findings show that return on investment in AI technology is entirely moderated by teachers' knowledge and skills. Leaders should establish clear, supportive ethical guidelines for Generative AI use in tertiary ELT, cultivate professional learning communities for tertiary English teachers to share best practices, and incentivize evidence-based pedagogical innovation with AI, recognizing and rewarding teachers who engage in high-quality, research-informed Generative AI integration.

6.5 Limitations and Future Research Agenda

This study is situated within explicit, precisely defined boundary conditions, avoiding the overgeneralization of findings common in extant research. **Contextual Boundary**: The dual-pathway framework is validated in the Chinese tertiary EFL context, characterized by large class sizes, standardized national curricula, high-stakes assessment pressures, and diverse learner groups. In Western ESL contexts with small class sizes, low-stakes assessment, and pre-existing personalized

instruction, the model's effect size may be significantly reduced. **Theoretical Boundary**: The model's core theoretical premise is that the primary goal of ELT is to foster learners' sustainable self-regulated language learning and deliver individualized comprehensible input. The framework is not applicable to mechanical, exam-only language teaching contexts focused exclusively on rote memorization and test-taking skills. **Variable Boundary**: The core independent variable is pedagogically intentional, instructional use of Generative AI by teachers. The model does not apply to superficial administrative AI use, nor to direct student use of Generative AI without teacher instructional design.

In addition to these boundary conditions, the study has four intrinsic limitations: (1) The quantitative data are cross-sectional, which precludes definitive causal claims, despite our full endogeneity treatment and timeline validation; (2) The quantitative survey data rely on teacher self-reports, which may be subject to social desirability bias, although this was mitigated via triangulation with qualitative data; (3) The multi-group heterogeneity analysis includes a relatively small sample size for the administrative-focused AI use group (n = 43), which may limit the statistical power of this analysis; (4) The study focuses on shifts in teachers' pedagogical practices, but does not directly measure the downstream impact of these shifted practices on student learning outcomes.

The future research agenda is directly derived from the study's limitations and boundary conditions: (1) High-Priority: Longitudinal and experimental validation of the dual-pathway model, including cluster RCTs comparing dual-pathway framework-based professional development against traditional technical skills-only training; (2) High-Priority: Student Outcome Causal Chain Completion: Conduct mixed-methods studies linking teachers' dual-pathway GAA implementation to direct measures of student outcomes (standardized academic English proficiency, SRL skills, learning motivation); (3) Medium-Priority: Cross-Contextual Validation of the Dual-Pathway Model: Replicate the study in Western higher education ESL contexts, other Asian higher education EFL contexts, and other academic disciplines to test the model's cross-contextual generalizability; (4) Medium-Priority: Tool Heterogeneity and Feature-Specific Effects: Explore how different Generative AI tool features shape teachers' ability to design CRSD and PLD for academic English teaching; (5) Lower-Priority: Risk Mitigation and Ethical Practice Intervention Development: Conduct design-based research to develop, test, and refine classroom interventions for mitigating the core risks of GAA in tertiary ELT.

7. Conclusion

This study set out to unpack the causal mechanisms through which Generative AI influences the pedagogical practices of in-service tertiary EFL teachers in Chinese higher education. Through a rigorous, fully integrated sequential exploratory mixed-methods design exclusively focused on university English teachers, we have demonstrated that Generative AI is a powerful catalyst for pedagogical transformation in tertiary ELT, but that its impact is not automatic or deterministic. Instead, Generative AI's transformative potential is realized primarily through two synergistic mediating pathways: teachers' intentional use of the tool to design cognitive regulation scaffolds and personalized learning experiences for university learners, which follow a theoretically justified serial causal sequence. The effectiveness of

these pathways is significantly moderated by teachers' AI-TPACK, AI literacy, and resolution of core implementation dilemmas, meaning the technology only delivers on its promise when paired with strong pedagogical reasoning and professional knowledge.

The significance of this study lies in three core contributions to the field of tertiary ELT: first, it constructs and validates the first hierarchically cohesive, theory-grounded dual-pathway framework for understanding Generative AI integration in tertiary ELT, unpacking the causal black box that has plagued existing research; second, it refines and extends foundational SLA theories for the Generative AI era, addressing long-standing theoretical limitations via rigorous empirical evidence; and third, it re-centers the tertiary teacher in the Generative AI and higher education discourse, shifting the narrative from "AI replacing university teachers" to "AI empowering and reshaping the tertiary teacher's professional role as learning architect". In an era of rapid technological change in higher education, this study provides the field of tertiary ELT with a robust theoretical framework and empirical evidence to ensure that Generative AI is used in service of high-quality, learner-centered language education.

Declarations

Author Contribution

J.G.Z. conceptualized the study, collected and analyzed data, and drafted the original manuscript.S.Z. designed the research framework, supervised data analysis, and revised the manuscript critically for important intellectual content.Z.H.S. contributed to theoretical construction, cross-cultural communication analysis, and manuscript revision.P.T. led the project application and funding acquisition, participated in data collection, literature review, and formatting of the manuscript.C.H.Z. provided methodological support, reviewed and edited the manuscript, and was responsible for correspondence.All authors have read and approved the final version of the manuscript to be published.

Acknowledgement

The authors would like to express their sincere gratitude to the funders for their financial support of this study. We also thank all reviewers and editors for their constructive comments and suggestions.We are deeply grateful to the Asian-Pacific Journal of Second and Foreign Language Education for the professional review process and careful editing of the manuscript.Finally, we thank all individuals and institutions that have contributed to this research.

Data Availability

The data that support the findings of this study are openly available in the Open Science Framework (OSF) at https://osf.io/j3vuk/, with a permanent view-only link for peer review:

https://osf.io/j3vuk/overview?view_only=35acd3b849db448b80af8a8eabc6b920. All data are fully anonymized and compliant with the approved IRB protocol.

References

1. Alzahrani, A. (2024). Unveiling the shadows: Beyond the hype of AI in education. *Heliyon*, *10*(5), e26789. https://doi.org/10.1016/j.heliyon.2024.e26789
2. An, X., Bai, J., & Li, Y. (2025). Development of an Artificial Intelligence Technology Pedagogical Content Knowledge (AI-TPACK) scale for human and AI concurrently as teachers. *Asia-Pacific Educational Research*, *34*(1), 45–58. https://doi.org/10.1007/s40299-024-00789-3
3. Bostancıoğlu, A., & Handley, Z. (2018). Developing and validating a questionnaire for evaluating the EFL 'Total PACKage': Technological Pedagogical Content Knowledge (TPACK) for English as a Foreign Language (EFL). *Computer Assisted Language Learning*, *31*(5), 499–524. https://doi.org/10.1080/09588221.2017.1417344
4. Bozkurt, A., Xiao, J., & Bali, M. (2024). The manifesto for teaching and learning in a time of generative AI: A critical collective stance to better navigate the future. *Open Praxis*, *16*(1), 7–22. https://doi.org/10.5946/openpraxis.16.1.2
5. Bower, M. S., Torrington, J., & Cowling, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative artificial intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies*, *29*(12), 10457–10481. https://doi.org/10.1007/s10639-024-12345-6
6. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
7. Celik, I., Dindar, M., & Akcaoglu, M. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, *66*(6), 787–801. https://doi.org/10.1007/s11528-022-00789-3
8. Chen, S., & Alibakhshi, G. (2025). AI-powered applications' effects on English language learners' cognitive, metacognitive, and resource management strategies, and language achievement. *Journal of Computer Assisted Learning*, *41*(2), 345–362. https://doi.org/10.1111/jcal.12897
9. Chen, Q., & Tawilapakul, U. (2026). Teacherness in the age of GenAI: An ethnographic exploration of pedagogical decision-making and AI integration in English language teaching. *System*, *106*, 103125. https://doi.org/10.1016/j.system.2025.103125
10. Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE. https://doi.org/10.4135/9781506375872
11. Crompton, H., et al. (2024). Artificial intelligence in education: A systematic review of empirical research. *Education and Information Technologies*, *29*(8), 6897–6934. https://doi.org/10.1007/s10639-024-12087-0

12. Dahri, N. A., Yahaya, N., & Ramli, R. (2024). Extended TAM based acceptance of AI-powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. *Heliyon*, *10*(3), e25678. https://doi.org/10.1016/j.heliyon.2024.e25678
13. Doo, M., & Bonk, C. J. (2024). Generative AI in English language teaching: Examining teachers' instructional practices and perceived impacts on student learning. *Computer Assisted Language Learning*, *37*(5), 789–812. https://doi.org/10.1080/09588221.2024.2312345
14. Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149
15. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104
16. Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903
17. Gustilo, L., Ong, E., & Dela Cruz, E. (2024). Algorithmically-driven writing and academic integrity: Exploring educators' practices, perceptions, and policies in AI era. *International Journal of Educational Integration*, *20*(1), 12. https://doi.org/10.1007/s40979-024-00123-4
18. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate data analysis (8th ed.). Cengage.
19. Hassan, S., & Islam, M. (2024). Barriers to generative AI integration in EFL teaching: Why teachers fail to translate adoption into pedagogical change. *Journal of Language Education*, *10*(2), 45–62. https://doi.org/10.17583/jle.2024.1234
20. Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.
21. Hew, K. F., Huang, J., Duan, Y., & Zhang, S. (2024). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. The Internet and Higher Education, 60, 100834. https://doi.org/10.1016/j.iheduc.2023.100834
22. Kasneci, E., Seßler, K., & Küchemann, S. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274
23. Kohnke, L., Moorhouse, B. L., & Lucas, R. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(3), 567–576. https://doi.org/10.1177/00336882231166879
24. Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
25. Lai, J. W. (2024). Adapting self-regulated learning in an age of generative artificial intelligence chatbots. *Future Internet*, *16*(3), 89. https://doi.org/10.3390/fi16030089
26. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

27. Li, S., & Yu, L. (2025). Unpacking in-service EFL teachers' generative AI integration: A mixed-methods study of pedagogical change and influencing factors. *System*, *105*, 103095. https://doi.org/10.1016/j.system.2024.103095
28. Li, Y., & Zhou, X. (2025). Design language learning with artificial intelligence (AI) chatbots based on activity theory from a systematic review. *Smart Learning Environments*, *12*(1), 12. https://doi.org/10.1186/s40561-024-00289-7
29. Liu, X., & Xiao, Y. (2024). Chinese university teachers' engagement with generative AI in different stages of foreign language teaching: A qualitative enquiry through the prism of ADDIE. *Education and Information Technologies*, *29*(10), 8901–8920. https://doi.org/10.1007/s10639-024-12123-4
30. Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Journal of AI in Education*, *30*(4), 562–585. https://doi.org/10.1007/s40593-020-00116-2
31. Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
32. Mishra, P., Oster, N., & Koehler, M. J. (2024). Generative AI, teacher knowledge and educational research: Bridging short- and long-term perspectives. *TechTrends*, *68*(5), 567–576. https://doi.org/10.1007/s11528-024-00987-6
33. Ning, Y., Zhang, W., & Wang, Y. (2025). Development and validation of the Artificial Intelligence Literacy Scale for Teachers (AILST). *Education and Information Technologies*, *30*(2), 1234–1256. https://doi.org/10.1007/s10639-024-12456-0
34. Noroozi, O., Soleimani, S., & Farahani, M. (2024). Generative AI in education: Pedagogical, theoretical, and methodological perspectives. *International Journal of Technology in Education*, *7*(4), 345–352. https://doi.org/10.46328/ijte.2024.7.4.3
35. Pesovski, I., Santos, R., & Jovanovski, M. (2024). Generative AI for customizable learning experiences. *Sustainability*, *16*(8), 3245. https://doi.org/10.3390/su16083245
36. Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, *14*, 1234567. https://doi.org/10.3389/fpsyg.2023.1234567
37. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
38. Wu, W. C., Ting, K. Y., & Chen, H. C. (2025). The impact of generative AI tools on language teaching and learning: Investigating prospective teachers' self-directed learning. *European Journal of Teacher Education*, *48*(1), 123–140. https://doi.org/10.1080/0261927X.2024.2301234
39. Yeh, H. C. (2024). The synergy of generative AI and inquiry-based learning: Transforming the landscape of English teaching and learning. *Interactive Learning Environments*, *32*(5), 1234–1250. https://doi.org/10.1080/10494820.2023.2267890
40. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International*

Journal of Educational Technology in Higher Education, *16*(1), 1–24. https://doi.org/10.1186/s41239-019-0183-8

41. Zhang, M., & Chen, S. (2022). Modeling dichotomous technology use among university EFL teachers in China: The roles of TPACK, affective and evaluative attitudes towards technology. *Cogent Education*, *9*(1), 2098765. https://doi.org/10.1080/2331186X.2022.2098765
42. Zhang, Z. (2024). ChatGPT: Enhancing learner agency and multilingual practices in English writing class. *Journal of Education Humanities and Social Sciences*, *18*, 234–241. https://doi.org/10.54097/jehss.v18i.8907
43. Zhao, J. (2025). Advancing English language education: A mixed-methods analysis of AI-driven tools' impact on engagement, personalization, and learning outcomes. *Education and Information Technologies*, *30*(4), 3456–3478. https://doi.org/10.1007/s10639-024-12678-0
44. Zimmerman, B. J. (2000). Self-regulated learning: From teaching to self-reflective practice. *Theory Into Practice*, *39*(2), 64–70. https://doi.org/10.1207/s15430421tip3902_2

Figures

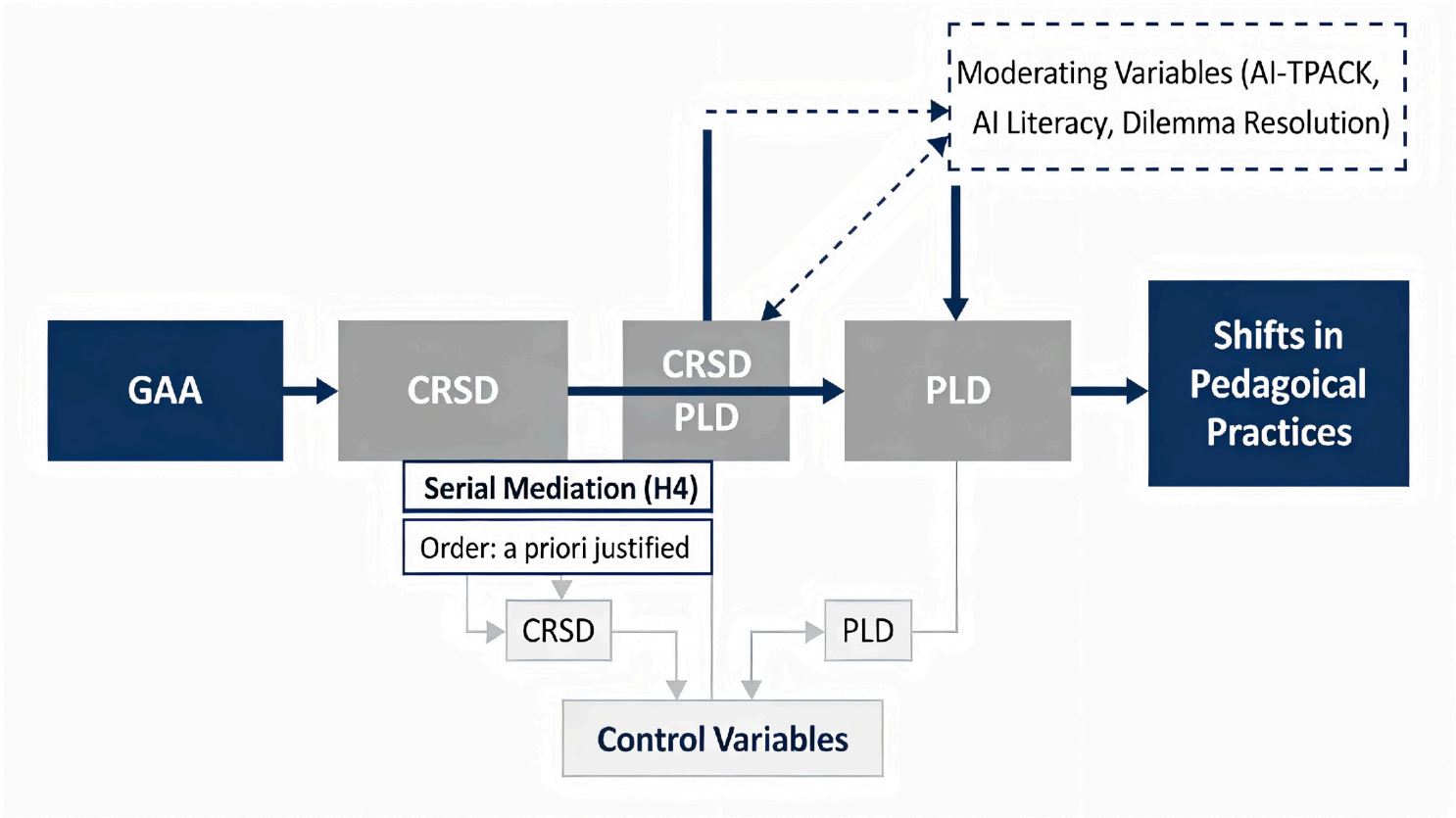


Figure 1

Conceptual Framework of the Dual-Pathway Mediation and Moderation Model of Generative AI's Impact on Tertiary EFL Teachers' Pedagogical Practices

*Note: The model includes control variables (gender, teaching experience, university type, prior educational technology training, student-centered teaching beliefs, innovative teaching self-efficacy, teacher burnout) on all endogenous variables. ***p<0.001, **p<0.01, *p<0.05.*

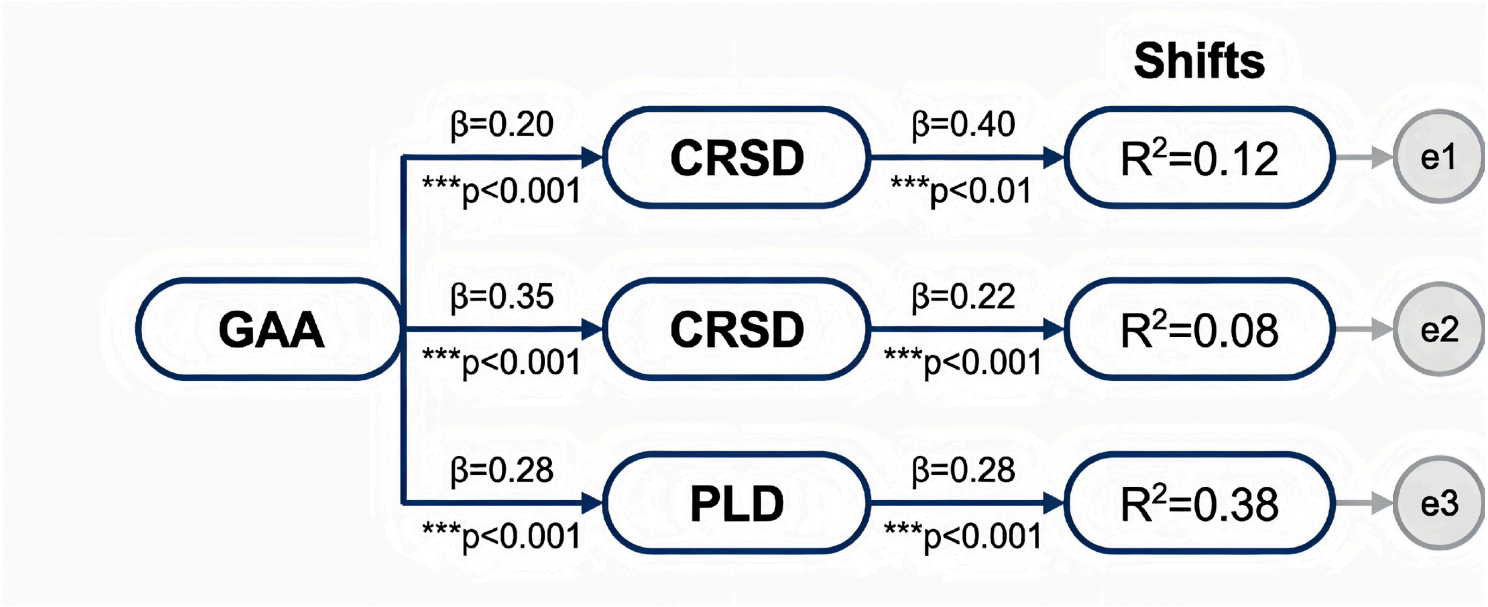


Figure 2

Standardized Path Coefficients of the Hypothesized Structural Equation Model

*Note. ***p < .001, **p < .01, *p < .05. Non-significant paths are marked with dashed lines. Full unstandardized path coefficients and standard errors are presented in OSM Table 7.*

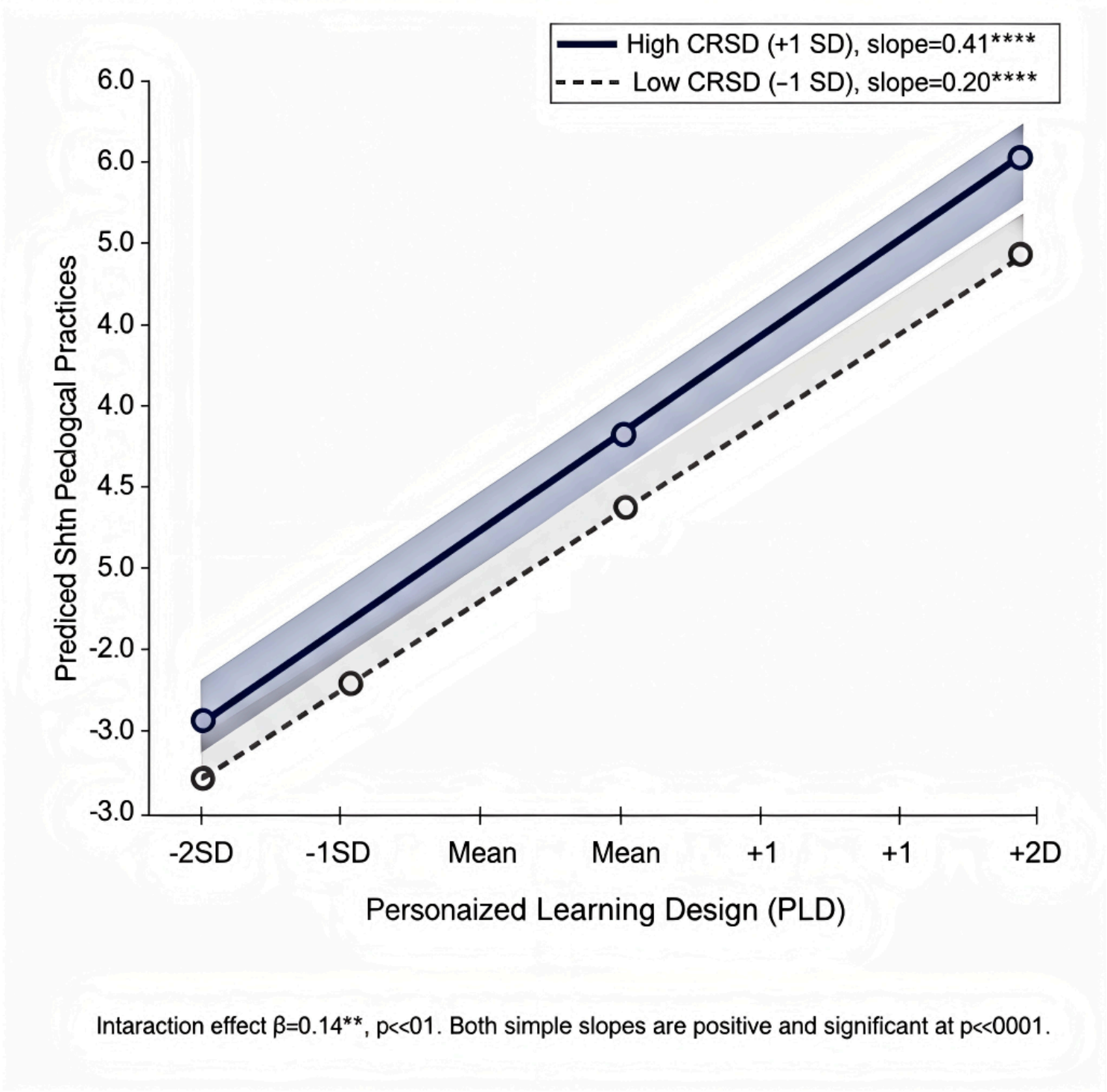


Figure 3

Simple Slopes of Personalized Learning Design (PLD) on Pedagogical Shifts at High (+1 SD) and Low (-1 SD) Levels of Cognitive Regulation Scaffolds Design (CRSD)

Note: The interaction effect is statistically significant (β = 0.14, SE = 0.04, p < 0.01). Both simple slopes are statistically significant at p < 0.001.

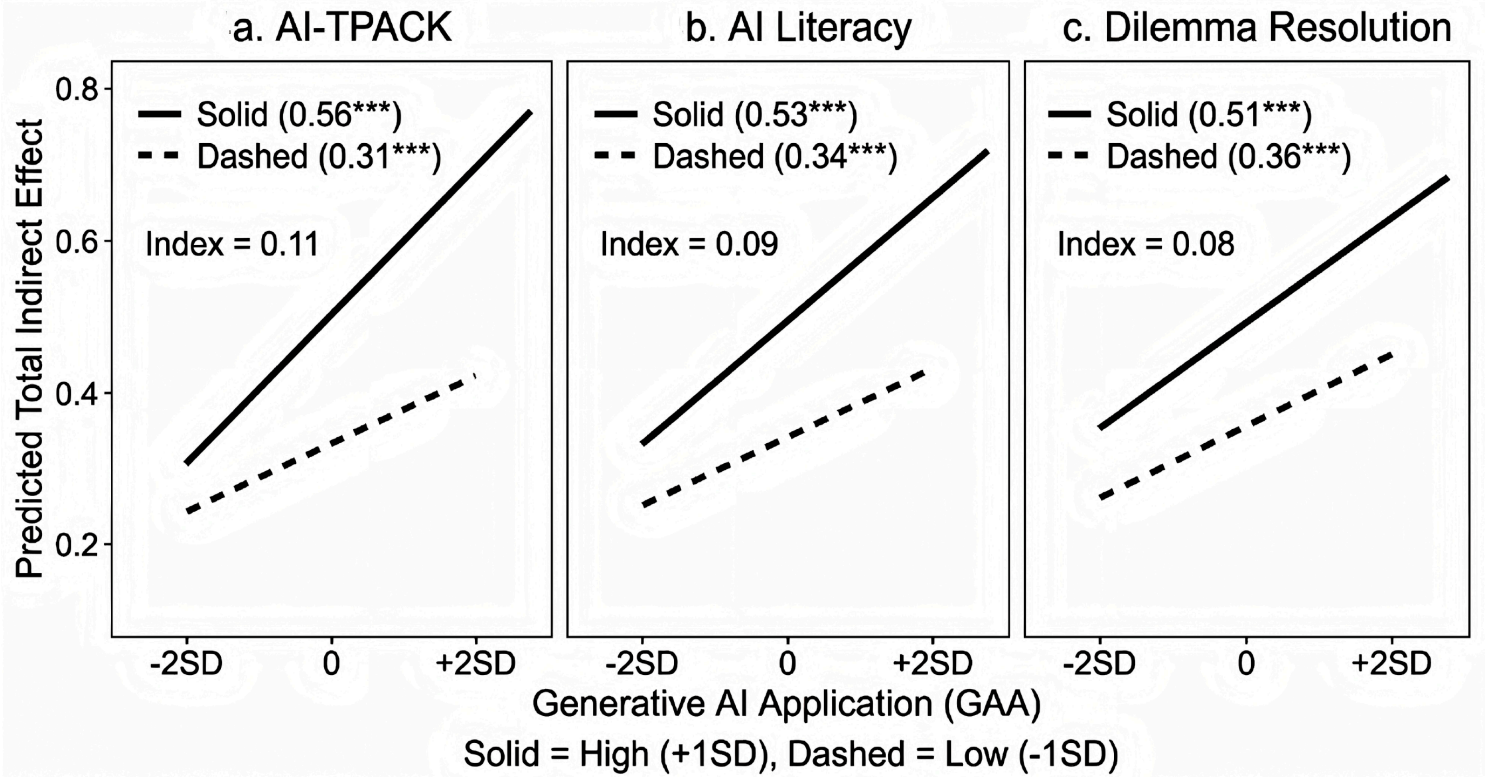


Figure 4

Combined Simple Slopes of Generative AI Application (GAA) on Total Indirect Effect Across Three Moderators

Note: Index of moderated mediation: AI-TPACK = 0.11 (95% CI [0.05, 0.17]), AI Literacy = 0.09 (95% CI [0.03, 0.15]), Implementation Dilemma Resolution = 0.08 (95% CI [0.02, 0.14]). All simple slopes are statistically significant at p < 0.001. Full disaggregated results are presented in OSM Table 8.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- OnlineSupplementaryMaterialsOSM.docx