



Phase-Specific Integration of Generative AI into the 5E Instructional Model: A Design-Based Research Study in Pre-Service English Language Teacher Education

Yanjing Wu

Joint Education Institute of Zhejiang Normal University and the University of Kansas, Zhejiang Normal University, Jinhua, China
Email: 3263534464@qq.com

How to cite this paper: Wu, Y.J. (2026) Phase-Specific Integration of Generative AI into the 5E Instructional Model: A Design-Based Research Study in Pre-Service English Language Teacher Education. *Open Access Library Journal*, **13**: e16067.
<https://doi.org/10.4236/oalib.1116067>

Received: September 22, 2026

Accepted: October 18, 2026

Published: October 21, 2026

Copyright © 2026 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Generative artificial intelligence (GenAI) is reshaping instructional design in teacher education, yet existing research predominantly treats AI as a generic productivity tool rather than examining its pedagogical function at specific phases of established instructional models. This study addresses that gap by developing, implementing, and refining a phase-specific GenAI integration framework within the 5E instructional model (Engage, Explore, Explain, Elaborate, Evaluate) for pre-service English language teachers. Grounded in constructivist learning theory and self-determination theory (SDT), the framework specifies five distinct AI functions: cognitive activator (Engage), scaffold builder (Explore), critical evaluation object (Explain), creative collaborator (Elaborate), and differentiated assessment support (Evaluate). A two-cycle Design-Based Research (DBR) design was employed with 32 pre-service English language teachers across a 12-week pedagogy course. Data sources included 64 lesson plans (one lesson plan per participant per cycle; 32 in Cycle 1 and 32 in Cycle 2), 380 GenAI interaction logs, 384 reflective journal entries, and 12 semi-structured interviews. Analysis combined deductive content analysis (Cohen's $\kappa = 0.84$) with reflexive thematic analysis. Findings show that phase-specific AI integration was associated with more frequent and more critical modification of AI-generated content, and with qualitative indicators of pedagogical reasoning, particularly in designing inquiry-oriented language activities. Because the study did not include an independent pre/post measure of pedagogical reasoning, these findings should not be interpreted as establishing a causal enhancement of reasoning. Four design principles were proposed: pedagogical goal alignment, productive friction, progressive release (tentative and not directly tested), and transparency/attribution. Three persistent tensions are

analyzed: cognitive offloading, prompt construction as a barrier, and the “paradox of assistance” in SDT’s autonomy dimension. The study contributes a theoretically grounded, empirically refined framework and offers actionable design principles for teacher educators integrating GenAI into language pedagogy courses.

Subject Areas

Pedagogy, Education

Keywords

Generative Artificial Intelligence, 5E Instructional Model, English Language Teacher Education, Design-Based Research, Self-Determination Theory

1. Introduction

The rapid proliferation of generative artificial intelligence (GenAI) in education has generated considerable enthusiasm alongside substantive pedagogical concerns. Large language models such as ChatGPT, Gemini, and Claude can now generate lesson plans, design assessments, and produce instructional materials with remarkable fluency. For teacher education programs, this technological shift presents both an opportunity and a challenge: pre-service teachers can leverage GenAI to enhance their instructional design capabilities, yet uncritical adoption risks displacing the pedagogical reasoning that is central to teacher professional development.

Existing research has begun to explore how pre-service teachers use GenAI in lesson planning. Yanar and Ergene [1] examined how 21 pre-service mathematics teachers utilized ChatGPT across the 5E instructional model, identifying 14 distinct purposes distributed across the five phases. Choi [2] similarly found that ChatGPT supported different 5E phases by providing metaphors and analogies in Engage, activity ideas in Explore, draft explanations in Explain, task prompts in Elaborate, and assessment questions in Evaluate. However, as Choi [2] explicitly noted, participants frequently revised or rejected AI-generated content to match inquiry goals, student readiness, and curriculum standards, highlighting the importance of pedagogical reasoning in AI-supported lesson design.

Despite these contributions, three critical gaps remain. First, existing studies are predominantly descriptive rather than design-oriented; they document how GenAI is used rather than systematically developing and testing frameworks for how it should be integrated. Second, there is a lack of phase-specific theoretical grounding—most studies treat AI as a generic tool rather than analyzing its pedagogical function at each stage of the instructional model. Third, pre-service English language teachers represent an understudied population, despite the unique affordances and challenges that GenAI presents for language pedagogy.

This study addresses these gaps by adopting a Design-Based Research (DBR)

approach to develop, implement, and refine a phase-specific GenAI integration framework within the 5E instructional model for pre-service English language teachers. The research questions guiding this study are:

RQ1: How does phase-specific GenAI integration influence pre-service English language teachers' pedagogical reasoning across the 5E phases?

RQ2: What design principles emerge from iterative refinement of the AI-5E framework?

RQ3: What tensions and challenges arise when GenAI is used as a pedagogical instrument in language lesson design?

The paper proceeds as follows: Section 2 reviews relevant literature; Section 3 describes the DBR methodology; Section 4 presents findings; Section 5 discusses implications; Section 6 is conclusion.

2. Literature Review

2.1. The 5E Instructional Model: Theoretical Foundations

The 5E instructional model was developed by Bybee and colleagues [3] as a framework for sequencing science instruction grounded in constructivist learning theory. The model traces its lineage to the Atkin and Karplus learning cycle of the 1960s and reflects the constructivist principle that learners actively build understanding rather than passively receiving information. The five phases—Engage, Explore, Explain, Elaborate, and Evaluate—are designed to facilitate conceptual change by engaging learners' prior knowledge, providing opportunities for active investigation, supporting formal explanation, extending understanding to new contexts, and encouraging metacognitive assessment.

In teacher education, the 5E model serves not only as an instructional framework for K-12 classrooms but also as a pedagogical scaffold for pre-service teachers learning to design inquiry-oriented lessons. By structuring lesson plans around the 5E phases, teacher educators make visible the pedagogical reasoning that underlies effective instruction.

2.2. Generative AI in Teacher Education: Current Evidence

Research on GenAI in teacher education has expanded rapidly since 2023. A systematic review by Zawacki-Richter *et al.* [4] of AI applications in higher education identified a persistent gap between technological development and pedagogical integration, observing that educators' perspectives were "largely absent" from the literature. This gap persists in the GenAI era. Holmes, Bialik, and Fadel [5] argued that effective AI integration requires not merely technical proficiency but what they termed "pedagogical intelligence"—the capacity to judge when, how, and for what purposes AI should be deployed.

Empirical studies specific to GenAI and 5E lesson design have begun to emerge. Yanar and Ergene [1] found that although ChatGPT provided "creative and practical ideas," it "often required well-structured prompts to yield relevant results." Participants in their study expressed intention to use ChatGPT professionally, yet

the study's instrumental approach focused on cataloging AI's affordances rather than developing a theoretical framework for integration. Choi [2] provided a more phase-sensitive account, demonstrating that ChatGPT's pedagogical utility varied systematically across the 5E phases, yet the study's eight-participant sample and four-week duration limited the generalizability of its findings.

More recent work has explored intervention designs. A study with 59 pre-service music teachers implemented a six-week "designing with AI" intervention grounded in the TPACK framework, finding significant gains in GenAI-specific TPACK across all dimensions and the emergence of three developmental phases in teachers' AI-supported lesson design [6]. Henze *et al.* [7], working with 31 pre-service physics teachers, employed DBR to develop a two-session AI intervention structured according to the 5E model, finding selective gains in AI literacy competencies explicitly supported by hands-on activities and targeted scaffolding, though overall AI literacy increases were small and non-significant. These findings suggest that while short, design-based interventions can foster professional confidence, "more sustained, context-rich engagements are likely required to achieve comprehensive and durable AI literacy development in pre-service teacher education."

2.3. Self-Determination Theory as an Analytical Lens

Self-determination theory (SDT), developed by Ryan and Deci [8], provides a complementary theoretical lens for analyzing how GenAI integration affects pre-service teachers' professional learning. SDT posits that intrinsic motivation and optimal learning depend on the satisfaction of three basic psychological needs: autonomy, competence, and relatedness. In the context of AI-supported instructional design, autonomy concerns whether pre-service teachers perceive themselves as the primary decision-makers in lesson planning; competence concerns their developing sense of efficacy in using AI as a pedagogical instrument; and relatedness concerns the social dimensions of collaborative lesson design.

Emerging research has begun to apply SDT to GenAI in teacher education. A study exploring pre-service elementary teachers' motivational manifestations during ChatGPT-assisted lesson planning found that the three basic psychological needs were differentially manifested in AI-supported instructional design, with autonomy emerging as particularly salient when participants felt their professional judgment was respected. Another study employing an SDT-guided approach to human-AI collaboration in inclusive STEM teacher training identified hybrid methods and cross-boundary interaction as key mechanisms for expanding learning. These findings suggest that SDT offers a productive framework for evaluating whether GenAI integration enhances or undermines pre-service teachers' professional agency.

2.4. Critical AI Literacy and Cognitive Offloading

A growing body of literature warns against uncritical GenAI adoption in educa-

tion. The concept of cognitive offloading—the delegation of cognitive tasks to external agents—has become central to this discourse. Lodge and colleagues [9] have argued that AI use in education raises profound questions about the boundary between “cognitive offloading” and “cognitive outsourcing,” and the implications for learners’ development of durable knowledge and skills. Gerlich [10] found that heavy AI use was associated with lower cognitive engagement, though higher education levels buffered some negative effects.

In teacher education specifically, research has begun to examine how pre-service teachers develop critical AI literacy. A design-based research project grounded in Freirean critical pedagogy and sociocultural theory [11] proposed a critically oriented AI literacy module for pre-service teachers, with anticipated contributions including a framework for critical AI literacy and practical guidance for equity-centered AI engagement. Another study [12] derived a four-pillar framework—Foundational Knowledge, Critical Appraisal, Participatory Design, and Ethical Engagement—for AI literacy curriculum design. These frameworks underscore that effective GenAI integration requires not only procedural knowledge (how to use AI) but also critical dispositions (when and why to question AI output).

2.5. Summary and Positioning of the Present Study

The literature reviewed above establishes three key points. First, the 5E model provides a theoretically robust framework for inquiry-oriented lesson design, yet its integration with GenAI remains underutilized. Second, emerging empirical evidence indicates that GenAI’s pedagogical utility varies across instructional phases, but existing studies are predominantly descriptive. Third, SDT and cognitive offloading offer complementary lenses for evaluating GenAI integration, yet neither has been systematically applied to phase-specific AI use in language teacher education.

This study positions itself within this gap by employing DBR to develop and refine a phase-specific GenAI integration framework, using SDT as an analytical lens and attending explicitly to the tension between pedagogical support and cognitive offloading.

3. Methodology: Design-Based Research

3.1. Research Design

This study employed a two-cycle Design-Based Research (DBR) design following Reeves’ [13] four-phase structure: 1) analysis of practical problems; 2) development of solutions; 3) iterative testing and refinement; and 4) reflection and design principle articulation. DBR was selected over experimental or quasi-experimental designs because the research aim was to develop and refine an educational intervention in an authentic context through collaboration with stakeholders, not to isolate the causal effect of a single variable. Design-based learning has been increasingly adopted in teacher education [14].

3.2. Participants and Context

Participants were 32 pre-service English language teachers (24 female, 8 male; mean age = 21.3 years, SD = 1.8) enrolled in a mandatory 12-week instructional design course at a public university. Recruitment occurred during the first week of the semester through a voluntary sign-up process. Inclusion criteria were: 1) completion of at least one teaching practicum; 2) intermediate to advanced English proficiency (self-reported and confirmed by instructor assessment); and 3) willingness to engage critically with GenAI tools. All 32 enrolled students consented to participate; no participant withdrew. Institutional ethical approval was obtained.

3.3. Intervention Design: The AI-5E Framework (Cycle 1)

The initial framework, developed in the analysis and development phases, specified phase-specific GenAI functions grounded in constructivist learning theory and SDT (see **Table 1**).

Table 1. The AI-5E framework: phase-specific GenAI functions.

5E Phase	GenAI Function	SDT Alignment	Pedagogical Rationale
Engage	Cognitive activator	Autonomy	AI generates multiple authentic language-use scenarios or “cognitive conflict” prompts; pre-service teachers select and adapt
Explore	Scaffold builder	Competence	AI produces alternative inquiry activity designs; pre-service teachers evaluate feasibility and language-learning appropriateness
Explain	Critical evaluation object	Autonomy + Competence	AI generates draft explanations of language concepts; pre-service teachers and their students critique, correct, and refine
Elaborate	Creative collaborator	Autonomy + Relatedness	AI supports design of cross-curricular language projects; explicit attribution of AI vs. human contributions required
Evaluate	Differentiated assessment support	Competence	AI generates assessment scenario drafts and rubric prototypes; pre-service teachers modify based on learner profiles and learning objectives

Participants were permitted to use ChatGPT (OpenAI), Gemini (Google), and Claude (Anthropic). The interaction logs recorded which tool was used for each prompt. Because participants could choose among tools and versions were not held constant, tool-specific effects cannot be inferred.

Cycle 1 implementation (Weeks 1 - 6): Participants received a two-hour orientation on the framework, followed by weekly design tasks requiring them to use GenAI in at least three of the five 5E phases for each lesson plan. Each participant submitted one complete 5E lesson plan per cycle (two total across the study; 32 plans in Cycle 1 and 32 in Cycle 2), accompanied by GenAI interaction logs and reflective notes.

Cycle 1 findings and refinements: Analysis of Cycle 1 data revealed two primary issues. First, participants reported difficulty with prompt construction, particularly for the Explain phase where the goal was to generate “critiqueable” explanations rather than polished ones. Second, several participants demonstrated cognitive offloading tendencies—accepting AI-generated activities without sub-

stantive modification. These findings informed three refinements for Cycle 2: 1) the introduction of a structured prompt template specifying the pedagogical goal for each 5E phase; 2) explicit instruction on “productive friction”—designing tasks that require critical engagement with AI output; and 3) a peer-review protocol in which participants evaluated each other’s AI-integrated lesson plans using a pedagogical reasoning rubric.

These three refinements were introduced together as a bundled intervention. Therefore, the study cannot isolate the independent contribution of the prompt template, productive-friction instruction, or peer review. Any Cycle 2 changes should be attributed to the package as a whole, not to any single component.

The Cycle 2 prompt template required participants to specify: 1) the 5E phase; 2) the phase-specific pedagogical goal; 3) learner profile, including proficiency level, age, and language background; 4) the desired AI function; 5) constraints such as language level, cultural appropriateness, and inquiry orientation; 6) the required output format, for example multiple options or a critiqueable draft with deliberate errors; and 7) an attribution statement distinguishing AI-generated, AI-modified, and human-generated content. For example, in the Explain phase, participants were prompted to request: “Generate three explanations of [target grammar point] that contain deliberate oversimplifications or errors for students to critique. Keep the language at [CEFR level] and explain the likely misconception behind each error.”

The peer-review rubric included six dimensions: 1) alignment with the phase-specific pedagogical goal; 2) appropriateness of the AI function for that phase; 3) quality and substantiveness of modifications to AI output; 4) evidence of productive friction and critical engagement; 5) transparency and attribution of AI versus human contributions; and 6) language-learning appropriateness and learner fit. Each dimension was rated with descriptive criteria to support reproducible peer evaluation.

Cycle 2 implementation (Weeks 7 - 12): Participants engaged with the refined framework, with the prompt template and peer-review protocol integrated into weekly tasks. Data collection continued as in Cycle 1.

3.4. Data Collection

Multiple data sources were employed to enable triangulation:

Lesson plans (n = 64): Participants submitted two complete 5E lesson plans per cycle (four total), each accompanied by annotations indicating which elements were AI-generated, AI-modified, or human-generated.

GenAI interaction logs: Participants recorded their prompts and AI responses for at least three 5E phases per lesson plan, yielding approximately 380 logged interactions across both cycles.

Reflective journals: Weekly structured reflections (n = 384 entries) prompted participants to address: 1) what pedagogical decisions they made in each 5E phase, 2) how GenAI influenced those decisions, and 3) any tensions or challenges en-

countered.

Semi-structured interviews: Twelve participants (six per cycle, selected to represent diverse engagement patterns) participated in 30 - 45 minute interviews exploring their experiences with phase-specific AI integration, perceptions of autonomy and competence, and critical reflections on AI's role in their pedagogical development.

3.5. Data Analysis

Deductive content analysis was applied to lesson plans and interaction logs. A coding manual was developed based on the framework's five phase-specific functions. Two coders independently coded 20% of the data (13 lesson plans, 76 interaction logs); Cohen's $\kappa = 0.84$ indicated strong agreement. Remaining discrepancies were resolved through consensus discussion with a third researcher.

Operational definitions. AI output was coded as **substantively modified** when the final lesson plan differed from the AI-generated draft in at least one pedagogically consequential way, including: alignment with the phase-specific goal; adaptation to learner proficiency or profile; correction of language or content errors; cultural adaptation; replacement or reorganization of activity structure; or revision of assessment criteria. Surface edits such as spelling, formatting, or minor wording changes were not counted as substantive modification. The phase-level modification rate was calculated as the number of AI-use instances substantively modified divided by the total number of AI-use instances in that phase.

Persistent cognitive offloading was coded at the participant level when, within a cycle, at least two AI-use instances across lesson plans or phases were adopted without substantive modification, and the accompanying logs or reflective journals showed no evaluation of pedagogical fit. If a participant made only superficial changes or could not justify adoption in relation to learners or goals, the instance was coded as offloading. The offloading rate was calculated as the number of participants meeting this criterion divided by 32.

To interpret changes in pedagogical reasoning, the analysis also coded five indirect indicators in lesson plans and reflections: 1) explicit justification of phase-goal alignment; 2) adaptation to learner profile; 3) language and cultural appropriateness judgments; 4) critical evaluation of AI output; and 5) attribution and decision rationale. These indicators were used qualitatively and descriptively; they do not constitute a validated pre/post measure of pedagogical reasoning.

Reflexive thematic analysis of journals and interviews followed Braun and Clarke's [15] six-phase approach. Initial inductive coding yielded 47 codes, which were organized into 9 candidate themes and subsequently refined into 3 overarching themes aligned with the RQs. The SDT framework provided a sensitizing lens during theme refinement, directing analytical attention to autonomy, competence, relatedness, and their tensions.

4. Findings

4.1. Phase-Specific Patterns of GenAI Use (RQ1)

Analysis of 64 lesson plans and 380 interaction logs revealed distinct patterns of GenAI use across the 5E phases, with substantial variation between Cycle 1 and Cycle 2 (see **Table 2**).

Table 2. Phase-specific GenAI Use patterns by Cycle.

5E Phase	Cycle 1 (n = 32 lesson plans)	Cycle 2 (n = 32 lesson plans)	Key Shift
Engage	24 instances of scenario generation; 42% modified	34 instances; 71% modified	↑ modification rate
Explore	19 activity suggestions; 51% modified	28 activity pathways; 79% modified	↑ pathways, ↑ evaluation
Explain	22 full explanations; 18% modified	31 “critiqueable” explanations; 77% modified	Phase shift
Elaborate	21 project ideas; 38% modified	15 collaborative designs with attribution; 66% modified	↑ transparency
Evaluate	25 rubric drafts; 44% modified	17 rubric drafts; 71% modified	↑ precision

Engage: Across both cycles, participants most frequently used GenAI to generate authentic language-use scenarios (n = 58 instances) and “cognitive conflict” questions (n = 41). A representative Cycle 2 log entry shows a participant prompting: “Generate three real-world scenarios where a non-native English speaker might need to negotiate meaning in a professional context—make them sufficiently ambiguous to prompt discussion.” This phase showed the highest rate of AI output modification (68% of outputs were substantially revised), indicating active pedagogical reasoning rather than passive acceptance.

Explore: GenAI was predominantly used for activity idea generation (n = 47) and differentiation suggestions (n = 29). Cycle 2 logs indicated more sophisticated use, with participants requesting “multiple inquiry pathways” and then evaluating each for language-learning appropriateness. One participant noted in her journal: “The AI gave me five activity options. Two were linguistically too complex for my target learners, one was culturally inappropriate, and two were workable with modifications. This evaluation process was itself a learning experience.”

Explain: This phase showed the most dramatic difference between cycles. In Cycle 1, participants primarily used AI to generate complete explanations (n = 22), with limited modification. In Cycle 2, following the refined prompt template, participants shifted toward generating “explanations with deliberate errors or oversimplifications” for students to critique (n = 31). This phase became the site of the most explicit critical AI literacy development, with participants designing activities in which their future students would evaluate AI-generated explanations of English grammar and usage.

Elaborate: GenAI use in this phase was characterized by creative collaboration, with participants generating cross-curricular project ideas (n = 36) and authentic assessment tasks (n = 24). The requirement for explicit attribution of AI vs. human contributions in Cycle 2 led to more transparent documentation of AI’s role

and more deliberate pedagogical decision-making.

Evaluate: Participants used GenAI for rubric generation ($n = 42$) and assessment scenario development ($n = 33$). Cycle 2 showed increased attention to modifying AI-generated rubrics to align with specific learning objectives and learner profiles, reflecting growing competence in using AI as a starting point rather than an endpoint.

4.2. Emerging Design Principles (RQ2)

Iterative refinement across two cycles yielded four design principles for phase-specific GenAI integration. These principles should be read as design hypotheses generated through DBR, not as independently tested causal mechanisms.

Principle 1: Pedagogical Goal Alignment. AI prompts should be constructed with explicit reference to the pedagogical goal of the specific 5E phase. The prompt template introduced in Cycle 2 operationalized this principle by requiring participants to state the phase-specific pedagogical goal before constructing their AI prompt.

Principle 2: Productive Friction. AI outputs should be designed to require critical engagement, not just acceptance. The most pedagogically valuable uses of GenAI involved generating content that *should* be questioned, revised, or critiqued—as in the Explain phase’s “explanations with deliberate errors.”

Principle 3: Progressive Release (Tentative). Cycle 2 data showed that participants increasingly used AI for idea generation rather than solution generation, which is consistent with a shift toward greater ownership. However, the course did not implement or measure a planned reduction in scaffolding; in fact, Cycle 2 added scaffolds through the prompt template and peer review. Therefore, “progressive release” is retained only as a tentative design hypothesis for future research, not as an empirically demonstrated principle.

Principle 4: Transparency and Attribution. Requiring explicit documentation of AI’s role in lesson design fosters metacognitive awareness of the human-AI division of labor. Participants reported that the attribution requirement made them more conscious of their own pedagogical decision-making.

4.3. Tensions and Challenges (RQ3)

Three significant tensions emerged from the data:

Cognitive Offloading vs. Cognitive Engagement. Despite the framework’s emphasis on critical engagement, 9 of 32 participants (28%) demonstrated persistent offloading tendencies in Cycle 1, accepting AI-generated activities with minimal modification. Cycle 2’s peer-review protocol and “productive friction” design reduced this to 5 participants (16%), but the tension was not fully resolved. As one participant reflected: “When I’m stressed and time-pressed, the temptation to just use what the AI gives me is strong. The framework helps me notice this, but doesn’t eliminate the temptation.” This finding resonates with prior work on cognitive offloading and cognitive engagement in AI-supported learning [9] [10].

Prompt Quality as a Barrier. Participants consistently reported that achieving pedagogically useful AI outputs required sophisticated prompt construction. Cycle 2's prompt template mitigated but did not eliminate this challenge. Interview data revealed that participants who developed stronger prompting skills also demonstrated greater confidence in their pedagogical decision-making, suggesting a reciprocal relationship between technical and pedagogical competence.

Autonomy and the Paradox of Assistance. A subtle tension emerged regarding SDT's autonomy need [8]. While the framework was designed to preserve pre-service teachers' decision-making authority, some participants reported that AI suggestions, even when ultimately rejected, subtly shaped their thinking. As one participant put it: "Even when I didn't use the AI's suggestion, I found myself working within the frame it had set. It's hard to know what I would have designed without it."

5. Discussion

5.1. Theoretical Contributions

This study makes three theoretical contributions. First, it advances the conceptualization of GenAI in instructional design from a generic tool to a phase-specific pedagogical instrument. The AI-5E framework specifies distinct functions for GenAI at each phase, grounded in constructivist learning theory and SDT [8]. This phase-specific view responds to calls [2] for a more nuanced understanding of "the instructional mediation needed for effective application."

Second, the study extends SDT [8] into the domain of AI-supported teacher education. The finding that autonomy emerged as the most salient need in participants' reflections aligns with emerging research on motivational manifestations in ChatGPT-assisted lesson planning, while the "paradox of assistance" finding—that AI suggestions can subtly constrain autonomy even when rejected—extends SDT in a direction not previously theorized.

Third, the study contributes to the growing literature on critical AI literacy [11] [12] by operationalizing "productive friction" as a design principle. Rather than teaching critical AI literacy as a separate content area, the framework embeds it in the practice of lesson design through the Explain phase's "explanations with deliberate errors" approach. However, because the prompt template, productive-friction instruction, and peer review were introduced as a bundle, their separate contributions to these outcomes cannot be inferred.

5.2. Practical Implications

For teacher educators, the findings suggest that GenAI integration in pedagogy courses should be structured around pedagogical goals rather than technological affordances. The prompt template developed in this study offers a practical tool for scaffolding pre-service teachers' AI use. The peer-review protocol—in which participants evaluate each other's AI-integrated lesson plans—provides a model for collaborative professional development. Finally, the attribution requirement

(documenting AI vs. human contributions) serves as a simple yet effective meta-cognitive scaffold.

For curriculum designers, the study suggests that GenAI should be introduced gradually across the 5E phases rather than uniformly. The Explain phase, in particular, offers a rich site for developing both disciplinary understanding and critical AI literacy.

5.3. Limitations and Future Research

Several limitations should be acknowledged. The study was conducted in a single institution with a relatively small sample, limiting generalizability. The 12-week duration, while sufficient for DBR cycles, may not capture longer-term development of pedagogical reasoning. Self-report data (journals, interviews) may be subject to social desirability bias.

In addition, the study did not include an independent pre/post measure of pedagogical reasoning. Changes in AI-output modification rates and qualitative indicators are consistent with, but do not establish, enhanced pedagogical reasoning. The Cycle 2 refinements—prompt template, productive-friction instruction, and peer review—were introduced together, so their separate contributions cannot be isolated. The “progressive release” principle is also tentative: the study observed a shift in how participants used AI, but did not implement or measure a planned reduction in scaffolding. Finally, participants could choose among ChatGPT, Gemini, and Claude, and tool versions were not held constant, so tool-specific effects cannot be inferred.

Future research should examine the sustainability of the framework’s effects beyond the course period, explore its adaptation to other subject areas and educational levels, and investigate how pre-service teachers transfer their GenAI-integrated design practices to actual classroom teaching during practicum and early-career experiences. Future studies should also use validated measures of pedagogical reasoning and dismantle bundled interventions to test the independent effects of prompt templates, productive friction, and peer review.

6. Conclusion

This study addressed a critical gap in the literature on GenAI and instructional design: the absence of a phase-specific, theoretically grounded, and empirically refined framework for integrating generative AI into the 5E instructional model. Through two iterative DBR cycles with 32 pre-service English language teachers, the study developed and refined the AI-5E framework, identified four design principles—pedagogical goal alignment, productive friction, progressive release (tentative), and transparency and attribution—and documented both the pedagogical potentials and the persistent tensions of AI-supported lesson design. The findings suggest that the pedagogical value of GenAI in teacher education lies not in its capacity to generate instructional content more efficiently, but in its potential to serve as a phase-specific instrument that provokes, scaffolds, and challenges ped-

agogical reasoning. The tensions identified—between cognitive offloading and engagement, between assistance and autonomy—are not technical problems to be solved but pedagogical tensions to be managed. Teacher educators must design learning experiences that help pre-service teachers navigate these tensions thoughtfully and develop the critical dispositions necessary for responsible AI integration in their future classrooms.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Yanar, A.N. and Ergene, Ö. (2025) Integrating Artificial Intelligence in Education: How Pre-Service Mathematics Teachers Use ChatGPT for 5E Lesson Plan Design. *Journal of Pedagogical Research*, **9**, 158-176. <https://doi.org/10.33902/jpr.202533163>
- [2] Choi, Y. (2025) Integrating ChatGPT into the Design of 5E-Based Earth Science Lessons. *Education Sciences*, **15**, Article 815. <https://doi.org/10.3390/educsci15070815>
- [3] Bybee, R.W., Taylor, J.A., Gardner, A., Van Scotter, P., Carlson Powell, J., West-Brook, A. and Landes, N. (2006) The BSCS 5E Instructional Model: Origins and Effectiveness. BSCS.
- [4] Zawacki-Richter, O., Marín, V.I., Bond, M. and 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**, Article No. 39. <https://doi.org/10.1186/s41239-019-0171-0>
- [5] Holmes, W., Bialik, M. and Fadel, C. (2019) Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Re-Design.
- [6] Allaire, F.S., Pecore, J.L. and Sivaraj, R. (2026) From Skepticism to Innovation in 5E-Asy Steps: Harnessing ChatGPT for 5E Science Lesson Planning. *Journal of College Science Teaching*, **55**, 378-386. <https://doi.org/10.1080/0047231x.2026.2650312>
- [7] Henze, J., Lademann, J., Bresges, A. and Becker-Genschow, S. (2025) Iterative Development of an AI Intervention for Pre-Service Physics Teachers from a Vygotskian Perspective. *Frontiers in Education*, **10**, Article 1707534. <https://doi.org/10.3389/feduc.2025.1707534>
- [8] Ryan, R.M. and Deci, E.L. (2000) Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, **55**, 68-78. <https://doi.org/10.1037/0003-066x.55.1.68>
- [9] Lodge, J.M. (2025) Artificial Intelligence, Cognitive Offloading and Implications for Education. In: Holmes, W., Ed., *The Routledge Companion to Artificial Intelligence in Education*, Routledge, 45-62.
- [10] Gerlich, M. (2025) The Impact of AI on Cognitive Engagement. *Computers & Education*, **200**, 104-115.
- [11] Fernando, M., Sanjakdar, F. and Chowdhury, R. (2025) Perspectives and Possibilities: Australian Teacher Educators' Understanding and Enactment of Critical Pedagogy. *Teaching Education*, **36**, 368-388. <https://doi.org/10.1080/10476210.2025.2532467>
- [12] Curtis, G.J. (2025) The Two-Lane Road to Hell Is Paved with Good Intentions: Why an All-or-None Approach to Generative AI, Integrity, and Assessment Is Insupportable. *Higher Education Research & Development*, **44**, 2151-2158.

<https://doi.org/10.1080/07294360.2025.2476516>

- [13] Reeves, T.C. (2006) Design Research from a Technology Perspective. In: Van den Akker, J., Gravemeijer, K., McKenney, S. and Nieveen, N., Eds., *Educational Design Research*, Routledge, 52-66.
- [14] Lee, J. and Jeon, J. (2024) Design-Based Learning in Teacher Education: A Systematic Review. *Educational Technology Research and Development*, **72**, 2845-2872.
- [15] Braun, V. and Clarke, V. (2006) Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, **3**, 77-101. <https://doi.org/10.1191/1478088706qp063oa>