

From Prompt to Professional Judgement: Designing Human-Centered AI Integration in Elementary Teacher Education

Beth Pesnell

Kansas State University

Abstract

Generative artificial intelligence (AI) is rapidly becoming part of educational practice, prompting teacher educators to reconsider how these tools should be integrated into preservice preparation. Rather than asking whether AI should be permitted or prohibited, this practitioner inquiry explores how instructional design can position AI as a catalyst for professional learning. In an undergraduate elementary STEAM methods course, preservice teachers used generative AI to produce an initial draft of an integrated instructional unit aligned to a Next Generation Science Standards (NGSS) performance expectation. They then engaged in a structured process of critique, revision, justification, and reflection designed to position AI-generated content as an object of professional analysis rather than a finished product. An iterative thematic review of students' written reflections suggested four recurring themes: AI lowered the barrier to beginning, students developed greater discernment in evaluating AI-generated content, revision became the primary site of professional judgment, and confidence emerged through ownership of instructional decisions. The findings suggest that the educational value of generative AI lies less in the technology itself than in the pedagogical structures surrounding its use. The article proposes a human-centered framework for AI integration in teacher education, arguing that thoughtfully designed learning experiences can reposition, rather than replace, the cognitive work of learning to teach.

Keywords: Artificial intelligence (AI) in education, teacher education, preservice teacher education, AI literacy

DOI: <https://doi.org/10.4148/edconsiderations.20914>

Introduction

Teacher education stands at an inflection point. Generative artificial intelligence (AI) is rapidly becoming embedded in educational practice through curriculum platforms, instructional planning tools, adaptive assessment systems, and generative writing assistants. The preservice teachers entering today's classrooms will encounter AI not as a future innovation but as a routine component of professional practice. Yet many teacher educators remain uncertain about how, or even whether, to incorporate these technologies into teacher preparation. Much of the conversation has centered on concerns about academic integrity, plagiarism, and the possibility that AI will replace the cognitive work essential to learning. Technological disruption in education is not new, and educators have continually negotiated the tension between efficiency and intellectual rigor. While these concerns are legitimate, they often obscure a more important pedagogical question: *How can teacher educators design learning experiences in which AI supports, rather than supplants, professional thinking?*

The emergence of generative AI, particularly text-generating systems powered by large language models (LLMs), has prompted educators to reconsider long-held assumptions about teaching and learning. Unlike earlier educational technologies that primarily increased access to information, large language models generate instructional materials, lesson plans, assessments, explanations, and feedback within seconds. This unprecedented capability presents both opportunity and risk. On one hand, AI has the potential to reduce routine planning tasks, generate instructional ideas, and support creative problem solving (Mollick & Mollick, 2023). On the other hand, AI-generated content can be inaccurate, biased, developmentally inappropriate, or pedagogically superficial if accepted without critical evaluation (Kasneci et al., 2023). Consequently, the central challenge for teacher education is not whether future teachers use AI, but whether they retain responsibility for interpreting, evaluating, and justifying instructional decisions. This requires designing learning experiences in which AI functions as an object of analysis rather than a substitute for professional judgment.

Recent scholarship increasingly suggests that the educational value of AI lies less in the technology itself than in the instructional design surrounding its use. Trust et al. (2025) found that AI-generated lesson plans can provide useful starting points but frequently require substantial revision to improve standards alignment, student engagement, inclusivity, and instructional quality. Similarly, emerging research on AI literacy argues that educators must develop the capacity to evaluate AI-generated content critically, recognize its limitations, and exercise informed professional judgment rather than treating AI output as authoritative (Long & Magerko, 2020). These perspectives shift the conversation from learning with AI to learning about AI through deliberate critique, revision, and reflection.

This emphasis on professional judgment is particularly important in teacher education. Lesson and unit planning are foundational practices through which preservice teachers develop pedagogical content knowledge and learn to make instructional decisions responsive to students, standards, and context. Designing instruction requires teachers to balance multiple considerations simultaneously, including learning goals, developmental appropriateness, assessment, differentiation, student engagement, and disciplinary coherence. For novice teachers, coordinating these demands can be cognitively overwhelming. The challenge is not merely producing a lesson

plan but learning to reason through the countless pedagogical decisions that shape effective instruction.

Research on teacher cognition helps explain why instructional planning is such a demanding professional task. Teachers continually draw upon their knowledge, beliefs, experiences, and contextual understanding while making instructional decisions (Borg, 2015). For preservice teachers, many of these knowledge structures are still developing, making instructional planning particularly complex. At the same time, cognitive load theory suggests that novice learners often devote substantial mental effort to managing procedural aspects of complex tasks before they can engage deeply with higher-order reasoning (Sweller, 1988). When working memory is consumed by organizing lesson structures, sequencing activities, or formatting instructional plans, fewer cognitive resources remain available for evaluating instructional coherence, considering learner variability, or strengthening disciplinary connections.

Reducing procedural demands is not synonymous with reducing intellectual rigor. Cognitive load theory distinguishes between minimizing unnecessary cognitive demands and preserving the cognitive work that contributes to meaningful learning (Sweller, 2023). In instructional design, the goal is not to eliminate productive struggle but to direct learners' attention toward the aspects of a task that matter most (Pass et al., 2003). From this perspective, generative AI may serve an important pedagogical function by reducing routine planning tasks so that novice teachers can devote greater attention to analyzing, critiquing, and refining instructional decisions.

This perspective also aligns with broader conceptions of professional learning. Effective teacher preparation involves far more than producing instructional materials; it requires developing the capacity to evaluate curricular resources, justify pedagogical choices, and adapt instruction for diverse learners (Darling-Hammond, 2017; Shulman, 1987). Brown's (2009) concept of Pedagogical Design Capacity further positions teaching as a design process in which educators perceive, select, adapt, and mobilize curricular resources to meet the needs of particular learners and contexts. From this perspective, AI-generated content is best understood as another curricular resource whose value depends on teachers' professional judgment rather than on the authority of the technology itself. Similarly, the Technological Pedagogical Content Knowledge (TPACK) framework emphasizes that technology should support, rather than drive, instructional decision making (Mishra & Koehler, 2006). Within this framework, AI should not function as an instructional authority but as one of many tools that teachers evaluate through the lenses of pedagogy, content knowledge, and learner needs.

These theoretical perspectives informed the design of an integrated Science, Technology, Engineering, Arts, and Math (STEAM) unit assignment in an undergraduate elementary STEAM methods course. Rather than asking preservice teachers either to avoid AI or to solely rely on it, the assignment intentionally positioned AI-generated content as the starting point for professional analysis. Students first generated a draft unit plan using a large language model, then systematically critiqued, revised, justified, and transformed that draft into an instructional unit aligned with standards, developmentally appropriate practices, and their own emerging teaching philosophy. The intellectual work of the assignment shifted from generating instructional materials to evaluating and improving them.

The purpose of this article is not to argue that AI inherently improves learning, nor to suggest that generative AI should replace traditional lesson design. Instead, it describes one instructional approach to integrating AI within teacher education while preserving the professional reasoning central to instructional planning (Brown, 2009). Drawing on preservice teachers' reflections, this practitioner inquiry explores how intentionally structured AI use influenced students' perceptions of lesson planning, interdisciplinary integration, and professional judgment. More broadly, it argues that the educational question is no longer whether AI belongs in teacher preparation, but how teacher educators can design learning experiences that ensure professional thinking remains at the center of AI-supported practice. Rather than treating AI as an instructional authority, this approach positions it as a curricular resource that prospective teachers must interpret, evaluate, and adapt in response to instructional goals and learner needs. This article proposes a simple design principle for AI integration in teacher education: *AI should reposition, not replace, the cognitive work of learning to teach.*

Designing for Professional Judgment

The ideas from teacher cognition, cognitive load theory, AI literacy, and TPACK informed the design of an integrated STEAM unit assignment in *Teaching STEAM PK–6*, an undergraduate methods course for preservice elementary teachers. The course is designed to help future teachers move beyond viewing STEAM as a collection of separate disciplines and instead develop their expertise in integrating pedagogical and disciplinary knowledge and skills needed to design meaningful, integrated learning experiences centered on authentic phenomena, interdisciplinary connections, and student sensemaking.

Like many teacher educators, I have traditionally asked students to design instructional units beginning with a blank document. For some, that blank page is an invitation. For others, it is paralyzing. Although this approach emphasizes original work, I increasingly questioned whether it also introduced unnecessary barriers for novice instructional designers. Developing a week-long integrated unit requires preservice teachers to simultaneously consider standards alignment, learning progressions, assessment, differentiation, engagement, interdisciplinary coherence, children's literature, instructional resources, and developmental appropriateness. For experienced educators, these considerations become increasingly automatic through years of practice. For novices, however, coordinating all these elements at once can overwhelm working memory before deeper pedagogical reasoning begins.

My goal was not to reduce the intellectual rigor of the assignment but to reposition where students invested their cognitive effort. Rather than asking students to devote substantial time generating the initial organizational structure of a unit, I wanted them to spend more time evaluating instructional quality, analyzing standards alignment, strengthening interdisciplinary connections, considering learner variability, and making deliberate pedagogical decisions. In other words, I sought to reduce procedural demands while preserving, and perhaps strengthening, the professional reasoning required of effective teachers.

Generative AI offered an opportunity to test this instructional design. Rather than functioning as the author of the assignment, AI became the starting point for analysis. Students were expected to critique AI-generated content with the same professional lens they would eventually apply to

commercially published curriculum materials, online instructional resources, or lessons shared by colleagues. The assignment was intentionally structured so that students could not simply accept the AI-generated product; instead, every phase required analysis, revision, justification, and reflection.

Phase 1: GENERATE – Generating an Initial Draft. Pre-service teachers (PST) first selected a Next Generation Science Standards (NGSS) performance expectation and prompted a generative AI tool of their choice to create a week-long integrated STEAM unit centered around their selected performance expectation and grade level. PST worked independently to generate a prompt, based on one of my examples, that requested ideas for integrating mathematics, technology, engineering, reading, writing, social studies, and art with the selected science standard in a week-long unit of study. Their AI-generated draft was required to include a central phenomenon or driving question, a lesson sequence or outline of the unit, ideas for integration in the content areas, suggested children's literature, instructional activities, and assessment ideas. They were provided with an example unit draft that I had done using ChatGPT as a model for this phase. The complete AI conversation transcript was submitted to document the interaction and promote transparency. This initial draft served one purpose: to create an instructional artifact worthy of professional critique.

Phase 2: CRITIQUE – Critiquing Through *Mark & Spark*. The second phase became the intellectual center of the assignment. Using a structured critique protocol analysis that I called *Mark & Spark*, PST carefully examined the AI-generated unit, and using their growing expertise, evaluated it through multiple pedagogical lenses. They analyzed standards alignment, developmental appropriateness, interdisciplinary coherence, cultural responsiveness, differentiation, student engagement, assessment practices, and instructional quality. They highlighted the strengths within the unit and identified opportunities for improvement and enhancement. This analysis paved the way for their revisions of the unit. They documented every revision they made and, more importantly, explained *why* each change strengthened the lesson. They were allowed to track their changes and note the rationale for their changes in a method of their choosing. Many used a T-Chart to document their analysis process, providing the changes they were making to the draft in the left column and the rationale for each change in the right column, while others noted their changes and rationale within the draft plan as comments or in a different color font.

The phrase *Mark & Spark* reflected the dual purpose of the assignment. The PST were encouraged to leave their *mark* by exercising professional judgment and to add their *spark* through creativity, instructional innovation, and responsiveness to learners. The emphasis shifted from producing instructional materials to evaluating and improving them. Their revisions were scored based on the clarity and depth of the revisions, new ideas and innovations, and the overall enhancement to unit quality.

This phase intentionally positioned AI as an object of critique rather than an instructional authority. Analyzing lessons and units is an essential practice for teachers, yet preservice teachers often get few opportunities to critique and adapt curricular materials (Beyer & Davis, 2011). Emerging research suggests that AI-generated lesson plans often provide useful organizational structures but require substantial revision to improve rigor, inclusivity, engagement, and disciplinary accuracy

(Trust et al., 2025). Rather than asking them to identify whether the AI unit was “good” or “bad,” the assignment challenged them to determine where professional expertise remained essential.

Phase 3: REVISE – Transforming the Unit. PST next revised the AI-generated draft into a complete instructional unit that reflected their own pedagogical reasoning. Rather than editing isolated components, they transformed the draft into an instructional plan they would feel confident teaching in an elementary classroom. I opted not to provide a template for this phase but allowed each to choose the format and design of their unit plan. Their final plan was scored on organization, elements of integrated STEAM instruction, alignment with NGSS and other content area standards, developmental appropriateness and cultural relevance, and quality of resources and assessments suggested in the plan. Developing integrated units is inherently complex and time-consuming. Although integrated instruction is not a new concept, it requires teachers to create meaningful connections across disciplines and concepts (Bradbury, 2014). As preservice teachers in an elementary methods course, they had not yet completed all their content-area methods courses and were therefore often asked to connect content beyond their prior preparation. Even veteran teachers struggle to design instruction that promotes meaningful interdisciplinary connections and real-world relevance (Fu & Sibert, 2017), making this challenge even greater for novice educators. The AI-generated draft served as a starting point by suggesting possible interdisciplinary connections. PST then critically evaluated those ideas, examining standards, refining content connections, and determining whether the proposed integration was appropriate, relevant, and substantive rather than superficial. In doing so, they devoted their effort to the intellectual and creative work of instructional design rather than to generating an initial draft.

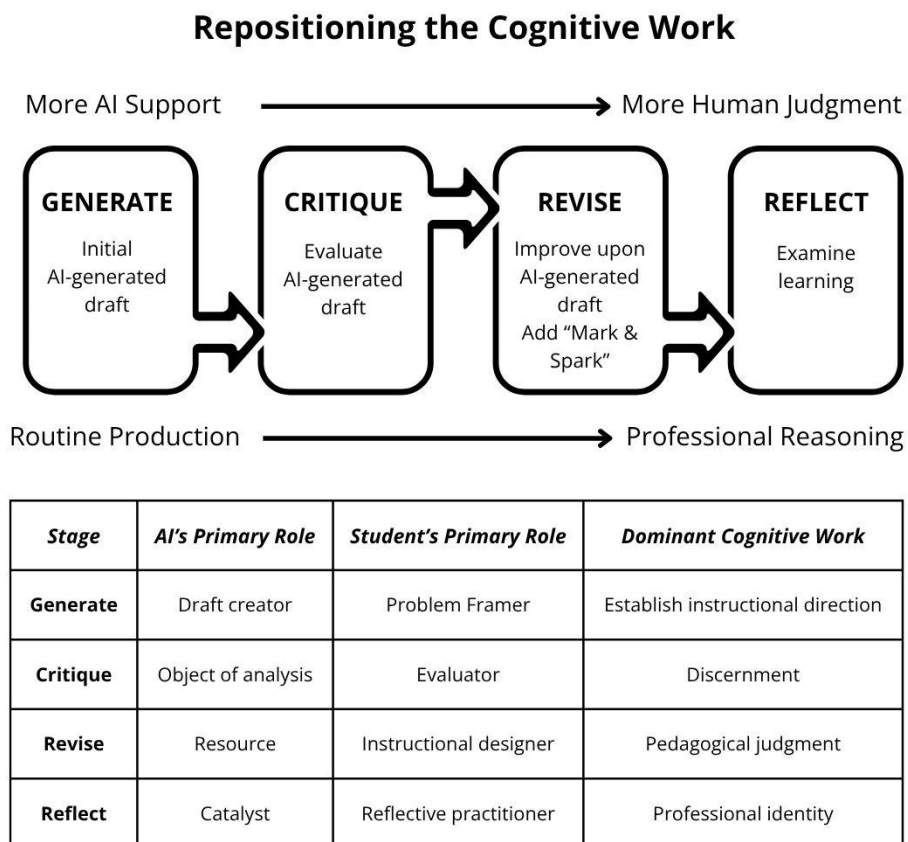
This stage represented more than revision. It required them to integrate content knowledge, pedagogical knowledge, and knowledge of learners while making dozens of instructional decisions. PST refined assessment strategies, strengthened disciplinary connections, incorporated culturally responsive resources, improved opportunities for student discourse and inquiry, and modified activities to better support elementary learners. The final product reflected not the capabilities of AI, but the emerging professional judgment of the preservice teacher.

Phase 4: REFLECT – Reflecting on Professional Learning. The final phase asked PST to submit a reflective analysis using three prompts:

1. How did designing this unit expand your understanding of integrated STEAM instruction?
2. How did the AI-generated first draft support your design process?
3. How might this experience influence your future lesson planning and instruction?

These reflections became an opportunity for students to examine not only the finished product but also their own learning process. Rather than evaluating AI itself, students were encouraged to consider how the assignment shaped their understanding of instructional design, interdisciplinary planning, and their future professional practice.

Collectively, the four phases were designed around a single instructional principle: *AI should reposition, not replace, the cognitive work of learning to teach.* (See Figure 1) By intentionally shifting students’ effort away from producing an initial draft and toward critique, revision, justification, and reflection, the assignment sought to preserve the professional reasoning that lies at the heart of effective instructional design.

Figure 1*AI-Supported Professional Judgment Framework*

Evidence of Emerging Professional Judgment

This article reports a practitioner inquiry situated within an undergraduate elementary STEAM methods course. Rather than evaluating the effectiveness of AI integration, the inquiry explored how one intentionally designed instructional experience shaped preservice teachers' developing professional reasoning. In Phase 4 of the assignment, they submitted a reflective analysis using three prompts to describe how the project influenced their understanding of integrated STEAM instruction, their experience using an AI-generated draft, and the ways the assignment might shape their future lesson planning. Thirty reflections were reviewed through multiple iterative readings to identify recurring patterns related to instructional planning, AI use, professional judgment, and confidence. Because the reflections were generated as part of normal coursework, the themes are presented as illustrative patterns in students' perceptions rather than as evidence of instructional effectiveness or causal impact.

Several themes emerged consistently across the reflections. Together, they suggest that the assignment influenced not only how students viewed AI, but also how they thought about instructional design and their role as future teachers. Rather than describing AI as replacing their thinking, students consistently described using the technology as a starting point for deeper pedagogical analysis. Four interconnected themes are presented below.

Theme One: AI Lowered the Barrier to Beginning, Not the Expectation for Thinking. The first and most prominent theme involved the challenge of getting started. Many students described the blank page as one of the most intimidating aspects of lesson planning. Rather than eliminating the intellectual work of instructional design, the AI-generated draft appeared to reduce the anxiety associated with organizing an initial structure, allowing students to focus their attention on evaluating, refining, and adapting the lesson.

Students repeatedly described the AI draft using structural metaphors. One explained that it provided “a solid base to work from,” allowing her to focus on “making it mine.” Another wrote that the draft served as “a foundation to work from,” while others referred to it as “a skeleton,” “a starting outline,” or “a head start.” One student summarized the experience succinctly: “I think the use of AI took out the skeleton and let me focus on the meat and potatoes.”

Collectively, these metaphors suggest that students understood the AI-generated lesson as a starting point rather than a finished product. The intellectual work shifted from creating an initial draft to interpreting, adapting, and strengthening the lesson through their own pedagogical reasoning.

Importantly, students did not describe the assignment as easier. Instead, they described redirecting their effort toward different aspects of instructional planning. One student reflected that “Having an initial structure helped me stay focused on the purpose of the content and making it mine.” Another explained that “Beginning with an AI-generated draft allowed more time to think deeply about standards alignment, objectives, and meaningful interdisciplinary connections.” Several students noted that they were able to devote greater attention to developmental appropriateness, learner supports, and authentic engagement rather than simply determining how to organize a week of instruction. These reflections suggest that the instructional design accomplished its intended purpose. Rather than reducing the cognitive demands of lesson planning, the assignment appeared to shift students’ attention away from generating an initial structure and toward evaluating the quality of instructional decisions.

Theme Two: Students Learned to Critique Rather Than Consume AI. Another theme was students’ willingness to critically evaluate AI-generated content rather than accept it uncritically. Instead, they consistently identified weaknesses, questioned recommendations, and described revision as an expected part of the process. Many students noted that AI-generated lessons required significant improvement. One student observed that the draft “did not account at all for cultural responsiveness.” Another explained that “suggested books were not always developmentally appropriate”, while others noted that activities lacked depth, differentiation, or meaningful engagement. One student summarized this perspective by writing, “The unit plans generated by AI absolutely need to be workshopped, modified, and carefully assembled in order to be used successfully.”

This critical stance extended beyond the assignment itself. Several students also remained cautious about future AI use. One reflected, “I do not see myself using AI often, as I believe a reliance on it could be detrimental.” Another concluded, “AI is a great stepping stone...however, it is often unreliable and needs to be double-checked for informational accuracy.” Even students who

expressed enthusiasm consistently qualified their support by emphasizing that AI should function as a brainstorming partner rather than an instructional authority.

Rather than increasing students' confidence in AI, the experience appeared to strengthen their confidence in their own professional judgment. Students did not emerge viewing AI as an expert. Instead, they described developing greater discernment about its strengths, limitations, and appropriate role in instructional planning. Learning to question AI-generated recommendations may represent one of the most valuable forms of AI literacy teacher educators can cultivate.

Theme Three: Revision Became the Site of Professional Judgment. Although students frequently appreciated the efficiency of generating an initial draft, they consistently identified revision as the most meaningful part of the assignment. Throughout the reflections, students described transforming rather than simply editing the AI-generated unit. One wrote, "I didn't just copy the AI's plan – I made it my own." Another reflected, "I was able to focus more on improving the unit versus trying to figure out how to create it." Students repeatedly described strengthening standards alignment, replacing instructional resources, increasing cultural responsiveness, improving differentiation, and designing more engaging learning experiences. Several noted that AI-generated ideas often inspired better instructional decisions precisely because they required thoughtful revision.

Interestingly, many students described this revision process as intellectually demanding. Rather than accepting the first draft, they found themselves asking whether activities genuinely supported the performance expectation, whether interdisciplinary connections were meaningful, and whether instructional choices reflected the needs of elementary learners. The assignment therefore positioned revision as an act of professional reasoning rather than mechanical editing.

This distinction is central to the design of the assignment. The goal was never to produce an acceptable AI-generated lesson plan. The goal was to develop preservice teachers who could recognize why instructional decisions mattered and justify the revisions they chose to make. Ownership emerged not from writing every word independently, but from exercising professional judgment throughout the revision process.

Theme Four: Confidence Emerged Through Ownership. The final theme centered on confidence. Nearly every student described feeling more prepared to design integrated instruction after completing the assignment. Importantly, this confidence did not appear to stem from AI itself but from the process of transforming an initial draft into an instructional unit they believed represented their own thinking.

Students repeatedly described feeling more capable of integrating content areas, designing around authentic phenomena, and developing lessons that reflected both standards and student needs. One reflected, "I feel more equipped to look at a standard and think, 'How can I bring this to life across different subjects?'" Prior to the assignment, many described STEAM in additive terms. One wrote, "I thought STEAM was about adding a creative activity to a science lesson." Another admitted, "I viewed the subjects as separate that could sometimes overlap." Integration, for them, was layering content. However, after engaging with and analyzing the AI-generated draft, their language changed. "When you plan around a real-world problem," one student reflected, "the

content areas naturally come together.” Another explained that the project helped her realize that integrated STEAM instruction was “not just about adding activities but about creating meaningful connections across disciplines.” A third observed, “One single activity could possibly have three different subject connections.” Students began describing STEAM as “an integrated mindset.” Integrated curriculum emphasizes coherence organized around authentic phenomena or driving questions. Effective STEAM instruction requires more than parallel subject inclusion; it requires conceptual interdependence (Roehrig et al., 2021). Developing a coherent interdisciplinary mindset requires scaffolding and reflection. Although the AI draft surfaced interdisciplinary possibilities, coherence emerged through students’ revision and professional decision making.

Students also consistently maintained the distinction between AI and teacher expertise. AI was described as “a planning partner,” “a co-planner,” “a brainstorming tool,” and “a starting point,” but never as a replacement for the teacher, emphasizing human authority in the process. As one student concluded, “It’s not a replacement for my knowledge as a teacher.” Another reflected that while AI could provide ideas, “It still takes a teacher to make it work.”

Taken together, these reflections suggest that students’ growing confidence was grounded less in the technology itself than in their increasing ability to evaluate, adapt, and justify instructional decisions. In other words, it strengthened their capacity to think like teachers.

Discussion

Designing AI Experiences That Cultivate Professional Judgment. The purpose of this project was never to determine whether generative AI is inherently beneficial or detrimental to learning. Instead, it explored how instructional design mediates the role AI plays in teacher preparation. The student reflections suggest that the assignment’s most significant contribution was not the use of AI itself, but the way the learning experience deliberately positioned AI-generated content as an object of professional critique. Rather than asking preservice teachers to generate instructional materials independently or accept AI-generated products uncritically, the assignment required them to evaluate, revise, justify, and ultimately transform an initial draft into instruction they believed reflected their own professional thinking.

This distinction is important because much of the current conversation surrounding AI in education remains framed as a binary choice: embrace AI as an efficiency tool or prohibit its use to preserve authentic learning. The experiences described by these preservice teachers suggest a third possibility. When instructional tasks are intentionally designed around critique, revision, and reflection, AI can serve as a catalyst for professional reasoning rather than a replacement for it. The intellectual work is not eliminated; it is repositioned.

AI Should Reposition, Not Replace, the Cognitive Work of Learning to Teach. A central implication of this inquiry is that generative AI can reposition, rather than replace, the cognitive work of learning to teach. Across the reflections, students consistently described the AI-generated draft as lowering the barrier to beginning without diminishing the intellectual demands of lesson planning. For novice teachers, coordinating standards, assessment, differentiation, engagement, and interdisciplinary integration requires the simultaneous orchestration of multiple developing

forms of professional knowledge. By providing an initial organizational framework, AI reduced procedural demands while preserving the need for pedagogical analysis.

Students did not describe thinking less; they described thinking differently. Rather than investing most of their effort in generating an initial structure, they devoted greater attention to evaluating instructional coherence, strengthening interdisciplinary connections, identifying weaknesses, improving developmental appropriateness, and considering the diverse needs of learners. These reflections align with cognitive load theory's distinction between reducing unnecessary procedural demands and preserving the cognitive processes that contribute to meaningful learning (Sweller, 1988).

This suggests an important design principle for teacher educators. AI integration should not be evaluated simply by asking whether students used AI. Instead, educators should consider where the intellectual work of the assignment resides. If AI performs the work students are intended to learn, the assignment risks diminishing opportunities for professional growth. If, however, AI creates opportunities for students to analyze, critique, justify, and improve instructional decisions, the technology may make professional reasoning more visible.

Professional Judgment Remains the Center of AI Literacy. The reflections also reinforce an increasingly important understanding of AI literacy. Competence with AI involves considerably more than writing effective prompts or generating instructional materials efficiently. Instead, preservice teachers must learn to recognize the strengths, limitations, biases, and inaccuracies of AI-generated content while exercising professional judgment about when, how, and whether specific recommendations should be adopted.

This distinction appeared repeatedly in students' reflections. Although many appreciated the organizational support AI provided, they consistently identified limitations related to cultural responsiveness, developmental appropriateness, disciplinary accuracy, learner supports, and instructional depth. Rather than describing AI as an expert, students described it as a brainstorming partner, planning assistant, or starting point. Several remained intentionally skeptical, emphasizing that AI-generated materials required careful review before classroom implementation.

These observations parallel recent scholarship suggesting that effective AI literacy requires critical evaluation rather than technological fluency alone (Long & Magerko, 2020). For teacher educators, this means that AI literacy should be conceptualized as an extension of professional pedagogical judgment. Darling-Hammond (2017) emphasizes that professional formation requires deliberate practice, feedback, and reflection. Tools that reduce logistical barriers while preserving analytic engagement may support this process and improve preservice teachers' abilities to analyze and critique pre-existing lesson plans and curricular materials (Cude & Haraway, 2017). Viewed through the lens of Pedagogical Design Capacity, AI-generated lesson plans become one more curricular resource that teachers must interpret, adapt, and mobilize in response to learners, content, and context (Brown, 2009). Evaluating AI-generated lesson plans requires many of the same competencies that teachers use when selecting commercial curriculum materials, adapting district resources, or modifying lessons shared by colleagues. In each case, the teacher, not the technology, remains responsible for determining instructional quality.

From Technology Integration to Instructional Design. This project also reinforces a broader lesson about educational technology. Historically, discussions surrounding new technologies have often focused on the capabilities of the tool rather than the design of the learning experience. Research in educational technology shares the principle that learning is influenced less by the presence of technology than by the pedagogical decisions that shape its use. Frameworks such as TPACK emphasize that technology contributes to learning only when thoughtfully integrated with pedagogy and content knowledge (Mishra & Koehler, 2006), while more recent work highlights the inseparable relationship among technology, teaching methods, purposes, and context (Fawns, 2022).

The assignment described in this article illustrates this principle. Had students been asked simply to submit an AI-generated lesson plan, the concerns frequently expressed about cognitive outsourcing would likely have been justified. Instead, the assignment deliberately embedded AI within a sequence of critique, revision, justification, and reflection. Those instructional moves created opportunities for professional learning. Throughout the assignment, AI functioned as one instructional resource among many rather than the driver of instructional design. Students repeatedly demonstrated that meaningful lesson planning required balancing technology-generated ideas with standards, disciplinary knowledge, learner needs, and professional judgment.

Design Principles for Teacher Educators. Although this project reflects one instructional context, several broader design principles may guide AI integration across teacher education programs:

- **Begin with pedagogy rather than technology.** Decisions about AI integration should emerge from learning goals rather than the capabilities of a particular platform. The instructional question should always precede the technological one.
- **Assess professional reasoning rather than AI-generated products.** When assignments emphasize critique, revision, and justification, assessment focuses on students' instructional thinking rather than the sophistication of AI output.
- **Position AI as a starting point rather than a finished product.** AI-generated materials are most valuable when they provide opportunities for students to evaluate quality, identify limitations, and strengthen instructional design.
- **Make revision visible.** Requiring students to document and justify changes encourages metacognitive awareness while making professional reasoning observable to both instructors and learners.
- **Cultivate healthy skepticism.** Rather than encouraging either unquestioning acceptance or complete rejection of AI, teacher education should help preservice teachers develop thoughtful discernment about when and how AI contributes to high-quality instruction.

Collectively, these principles suggest that AI integration is fundamentally an instructional design challenge rather than a technological one. The educational value of AI depends less on the sophistication of the tool than on the cognitive work educators intentionally ask students to perform.

Limitations and Future Directions

Several limitations should be considered when interpreting the experiences described in this article. First, the reflections represent the perceptions of preservice teachers enrolled in a single undergraduate elementary STEAM methods course at one institution. As such, the findings are intended to illustrate how one instructional design was experienced rather than to provide evidence of generalizable learning outcomes.

Second, the primary source of evidence consisted of students' written reflections completed as part of a course assignment. Although these reflections provide valuable insight into students' perceptions of their learning, they represent self-reported experiences rather than direct measures of changes in pedagogical reasoning or instructional performance. Future research might examine how structured AI integration influences lesson quality, classroom practice during student teaching, or longer-term professional decision making.

Third, the themes presented emerged through an instructor's iterative reading of student reflections rather than through a formal qualitative research methodology involving multiple coders or procedures to establish trustworthiness. The purpose of this practitioner inquiry was to identify recurring patterns that might inform instructional design rather than to make empirical claims about cognitive change or establish causal relationships.

Finally, generative AI technologies continue to evolve rapidly. The capabilities, limitations, and classroom applications of these tools will undoubtedly change over time. Consequently, the enduring contribution of this project lies less in the specific AI platform students used than in the instructional design principles that positioned AI as a catalyst for critique, revision, and professional judgment. Future research should continue exploring how these design principles translate across disciplines, teacher preparation contexts, and emerging AI technologies.

Conclusion

Generative AI has prompted teacher educators to reconsider long-standing assumptions about lesson planning, originality, and the role of technology in professional learning. Much of the public conversation has centered on whether AI should be permitted in educational settings, often framing its use as a choice between embracing innovation and protecting authentic thinking. This project suggests that such a framing may overlook a more consequential question. The issue is not simply whether preservice teachers use AI, but how teacher educators design learning experiences that shape what students do with it.

The assignment described in this article intentionally positioned AI-generated content as the beginning of instructional thinking rather than its conclusion. Students were asked not merely to generate lesson plans but to critique them, strengthen them, justify their revisions, and reflect on the professional decisions that transformed an initial draft into meaningful instruction. Their reflections suggest that this process encouraged greater attention to standards alignment, interdisciplinary coherence, developmental appropriateness, cultural responsiveness, and learner needs. Just as importantly, many students described developing a more discerning view of AI itself, seeing it as a tool whose value depended upon thoughtful human judgment.

Although these observations arise from one instructional context and are based on students' perceptions rather than direct measures of learning, they point toward an important design principle for teacher education. The educational value of AI does not reside in the sophistication of the technology but in the pedagogy that surrounds its use. Assignments that require critique, revision, justification, and reflection preserve the intellectual work of learning to teach while helping future educators develop the professional discernment that they will need in AI-rich classrooms.

Perhaps the most significant lesson from this experience is that professional judgment became visible. Students did not simply accept AI-generated lesson plans; they questioned assumptions, identified omissions, revised instructional decisions, and articulated why those revisions mattered. In doing so, they practiced the very work that defines effective teaching. The assignment did not ask students to demonstrate whether AI could plan instruction. It asked them to demonstrate whether they could.

Teacher education has always been about more than producing lesson plans. It is about cultivating educators who can make thoughtful instructional decisions in response to learners, contexts, and emerging challenges. That responsibility remains unchanged in an era of generative AI. What has changed are the tools available to support that work.

The challenge before teacher educators, therefore, is not to decide whether AI belongs in teacher preparation. It is to design learning experiences that ensure AI repositions, rather than replaces, the cognitive work of becoming a teacher. When thoughtfully integrated, generative AI does not diminish the need for professional judgment. Instead, it can create new opportunities to make that judgment visible, discussable, and ultimately teachable.

AI Disclosure Statement

Artificial intelligence tools were used in the development of this article as part of an iterative writing, revision, and design process. Specifically, generative AI was used as a support for brainstorming, organizing ideas, refining language for clarity and coherence, and assisting with the synthesis of existing scholarship and student reflection data.

All substantive decisions, including conceptual framing, selection and interpretation of sources, and final wording were made by the author. Citations included in this work refer to real, verifiable publications, and responsibility for accuracy, attribution, and interpretation rests solely with the author.

AI tools did not replace scholarly judgment, original thinking, or professional expertise. Instead, they functioned as a cognitive and compositional aid within a human-in-the-loop process, supporting reflection, revision, and clarity while preserving the intellectual, ethical, and pedagogical integrity of the work.

References

- Beyer, C. J. & Davis, E.A. (2011). Learning to critique and adapt science curriculum materials: Examining the development of preservice elementary teachers' pedagogical content knowledge. *Science Education*, 96(1), 130-157. <https://doi.org/10.1002/sce.20466>
- Borg, S. (2015). *Teacher cognition and language education: Research and practice*. Bloomsbury. <https://doi.org/10.5040/9781474219983>
- Bradbury, L. U. (2014). Linking science and language arts: A review of the literature which compares integrated versus non-integrated approaches. *Journal of Science Teacher Education*, 25(4), 465-488. <https://doi.org/10.1007/s10972-013-9368-6>
- Brown, M. W. (2009). The teacher-tool relationship: Theorizing the design and use of curriculum materials. In J. T. Remillard, B. A. Herbel-Eisenmann, & G. M. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction*. (1st ed.). 17-36. Routledge. <https://doi.org/10.4324/9780203884645>
- Cude, M.D., & Haraway, D.L. (2017). Documenting preservice teacher growth through critical assessment of online lesson plans. <https://api.semanticscholar.org/CorpusID:149203638>
- Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, 40(3), 291–309. <https://doi.org/10.1080/02619768.2017.1315399>
- Fawns, T. (2022). An entangled pedagogy: Looking beyond the pedagogy-technology dichotomy. *Postdigit Science and Education*, 4, 711-728. <https://doi.org/10.1007/s42438-022-00302-7>
- Fu, J., & Sibert, E. (2017). Teachers' perspectives: Factors that impact implementation of integrated curriculum in K-3 classrooms. *International Journal of Instruction*, 10(1), 169–186. <https://doi.org/10.12973/iji.2017.10111a>
- Kasneci, E., et al. (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>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- 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>
- Mollick, E., & Mollick, L. (2023). Assigning AI: Seven approaches for students, with prompts. *The Wharton School Research Paper*. SSRN Electronic Journal. <https://dx.doi.org/10.2139/ssrn.4475995>
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1
- Roehrig, G., Dare, E., Ellis, J., & Ring-Whalen, E. (2021). Beyond the basics: A detailed conceptual framework of integrated STEM. *Disciplinary and Interdisciplinary Science Education Research*, 3(1), 1-18. <https://doi.org/10.1186/s43031-021-00041-y>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>

- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (2023). The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review*, 35, 95. <https://doi.org/10.1007/s10648-023-09817-2>
- Trust, T., Maloy, R. W., Xu, C., & Pelletier, K. (2025). Civic education in the age of AI: Should we trust AI-generated lesson plans? *Contemporary Issues in Technology and Teacher Education*, 25(3). <https://citejournal.org/volume-25/issue-3-25/social-studies/civic-education-in-the-age-of-ai-should-we-trust-ai-generated-lesson-plans/>

Beth Pesnell (bpesnell@ksu.edu) is a teaching assistant professor of the Department of Curriculum & Instruction, College of Education, at Kansas State University in Manhattan, KS.