

Beyond the First Prompt: Preservice Teachers' Intentions and Practices in AI-assisted Lesson Design

Şule Yılmaz Özden ^{1,*}, Nahide İrem Azizoğlu ², Elife Nur Saydam ³, and Huijing Wen ⁴

¹ Department of Instructional Technologies, Sakarya University, Sakarya, Türkiye

² Department of Turkish Language Education, Sakarya University, Sakarya, Türkiye

³ Department of Elementary Education, Sakarya University, Sakarya, Türkiye

⁴ Department of Education, Moravian University, Bethlehem, PA, USA

Email: sule@sakarya.edu.tr (Ş.Y.Ö.); azizoglu@sakarya.edu.tr (N.İ.A.); elifenursaydam@sakarya.edu.tr (E.N.S.); wenh@moravian.edu (H.W.)

*Corresponding author

Manuscript received April 7, 2026; revised May 5, 2026; accepted June 18, 2026; published October 22, 2026

Abstract—This study examines Preservice Teachers' (PSTs) traditional lesson-planning practices, their proficiency levels, their interaction patterns with a specialized Artificial Intelligence (AI) powered lesson planning tool called Madlen, and their experiences and reflections throughout this integration process using a qualitative case study design. Data were collected through semi-structured interviews, an assessment rubric for AI supported lesson planning, and chat transcripts documenting PSTs' interactions with the AI tool. Ten PSTs from primary education and Turkish language education programs participated in the study, each designing three lesson plans with Madlen following structured AI literacy training. Baseline analyses revealed that PSTs possessed structured, outcome-oriented planning competencies. However, interaction transcript analyses showed that most participants employed single turn, product-oriented prompts with minimal iterative refinement or critical evaluation of AI outputs. Although participants acknowledged the potential of AI tools to enhance time efficiency, content diversity, and creative lesson design, their actual engagement patterns indicated that these positive perceptions did not translate into deep, critically reflective AI use. This perception-action gap, in which declarative knowledge about effective AI use coexists with superficial interaction behaviors, carries significant implications for assessing AI literacy in teacher education contexts. These findings suggest that AI literacy training alone is insufficient; PSTs require scaffolded opportunities to develop skills in iterative prompting, critical evaluation of outputs, and pedagogical adaptation. This study contributes to the growing literature on AI in teacher education by providing direct evidence from human AI interaction data, moving beyond perception-based approaches that dominate current research.

Keywords—Artificial Intelligence (AI) in education, lesson planning, preservice teachers, AI-assisted instructional design, human-AI interaction, AI literacy

I. INTRODUCTION

The integration of Artificial Intelligence (AI) into educational contexts is reshaping how instructional experiences are conceived, designed, and delivered [1, 2]. Within this evolving landscape, generative AI tools and intelligent platforms designed for educational settings have opened new possibilities for pedagogical innovation, from developing content aligned with specific curricula to designing assessments and providing real-time feedback [3]. These developments indicate that AI can assist with routine content creation while also enabling more personalized and adaptive forms of instruction [4].

However, despite the growing ubiquity of AI tools in educational settings, a critical research gap remains in teacher

preparation and AI integration. While substantial scholarly attention has been devoted to how school-aged students use AI technologies to enhance learning [5, 6], significantly less empirical research has examined how preservice teachers integrate AI into lesson planning. This research gap is particularly concerning given the mounting evidence that many Preservice Teachers (PSTs) demonstrate limited AI literacy and lack the capacity to critically assess AI generated outputs [7, 8]. Without such competencies, AI tools risk producing generic, culturally insensitive, or pedagogically inappropriate content [9, 10].

This gap matters for more than practical reasons and also shapes how preservice teachers form their professional identity and exercise pedagogical judgment as AI becomes part of teaching. Early experiences with AI during teacher preparation programs can shape preservice teachers' technology adoption patterns, instructional design approaches, their comfort with digital innovation, and capacity for critical engagement with emerging educational technologies [11]. Moreover, these formative experiences shape preservice teachers' understanding of the appropriate balance between technological assistance and teacher pedagogical judgment, a balance that will prove increasingly crucial as AI tools become more sophisticated and pervasive in educational contexts.

The urgency of this research is further underscored by the growing disconnect between the ambitious AI-integration goals set by national and international policy frameworks and the limited empirical evidence available on how these goals are being realized in teacher education programs [12, 13]. Understanding how PSTs balance technological possibilities with pedagogical responsibilities and develop critical evaluation skills in AI-mediated contexts, is essential for designing effective teacher preparation curricula [14, 15]. Without such understanding, teacher education programs risk preparing graduates who can use AI tools at a surface level but lack the critical competencies needed to integrate them in pedagogically meaningful ways.

This study addresses this gap by combining traditional qualitative data sources with systematic analysis of PST-AI chat transcripts. By examining both what PSTs report and what they demonstrably do when interacting with a specialized lesson planning AI tool, the study provides a more complete and empirically grounded account of AI integration in teacher education.

This study contributes to the field in three ways. First, it

moves beyond perception-based approaches by providing direct empirical evidence from human-AI interaction transcripts, offering a window into what preservice teachers actually do when working with AI tools rather than what they report doing. Second, it documents a perception-action gap in which PSTs articulate the importance of critical evaluation and teacher oversight yet demonstrate no such behaviors in their actual interactions with the AI tool. Third, drawing on this gap, the study proposes the concept of “procedural AI literacy” to capture the distinction between declarative understanding of effective AI use and the capacity to enact that understanding in practice, extending current frameworks such as the UNESCO AI Competency Framework.

The present study conducts a comprehensive investigation into how PSTs utilize a specialized AI tool for lesson planning purposes. Through systematic examination of their interaction patterns, experiences, and perceptions, this research seeks to illuminate the dynamics between established lesson planning approaches and AI-assisted instructional design. The study is guided by four interrelated research questions:

- (1) How do preservice teachers prepare lesson plans using traditional, non-AI-assisted methods?
- (2) What levels of proficiency do preservice teachers demonstrate in AI-supported lesson planning?
- (3) To what extent and in what ways do preservice teachers interact with the AI tool during the lesson planning process?
- (4) What are preservice teachers’ experiences, perceptions, and reflections regarding AI-assisted lesson planning, and how do these align with their observed interaction patterns?

II. LITERATURE REVIEW

A. AI Literacy and Preparedness of Preservice Teachers

AI literacy has become an increasingly important competency as AI tools become more embedded in everyday educational practice [16]. As outlined in the theoretical framework presented below, effective AI literacy goes beyond familiarity with AI tools; it requires knowledge, skills, and dispositions that work together to support competent practice in professional and social contexts [14, 16]. For AI to be meaningfully integrated into teaching, educators need a working understanding of what these tools can and cannot do, and where they fit within pedagogical practice. Yet research consistently shows that preservice teachers fall short in this regard.

Research consistently reveals substantial weaknesses in PSTs’ AI literacy. Pei *et al.* [2], using Bloom’s Taxonomy and TAM with 200 US preservice teachers, found significant gaps in advanced cognitive skills such as analyzing AI outputs and evaluating their pedagogical suitability. Similarly, Guan *et al.* [7] observed that despite acknowledging AI’s instructional potential, PSTs’ interactions remained predominantly superficial, characterized by uncritical acceptance of outputs and limited consideration of pedagogical appropriateness. Hava and Babayigit [12] extended these findings by demonstrating that, although general digital literacy predicts AI-TPACK development, AI-specific competencies remain significantly

below levels required for effective pedagogical integration. Collectively, these findings underscore a concerning disparity between the accessibility of AI tools and PSTs’ readiness to employ them effectively.

B. Adoption Patterns and Attitudes toward AI Integration

Comprehending the motivations behind preservice teachers’ use of AI technologies offers essential insights into the elements that promote or hinder effective integration. Investigations in this domain often use established theoretical frameworks from the technology adoption literature, such as the Diffusion of Innovations (DOI) theory and various versions of the Technology Acceptance Model (TAM), to elucidate and forecast adoption patterns.

Research on AI adoption patterns among PSTs reveals a complex interplay of individual, social, and institutional factors. Nissim and Simon [8] identified distinct adopter profiles, finding that early adopters demonstrated greater willingness to experiment and more critical evaluation of AI outputs. Acquah *et al.* [17] identified social influence and performance expectancy as primary determinants of AI adoption intentions, underscoring the role of peer modeling and institutional culture. Gamlem *et al.* [5] found that despite broad awareness and frequent use of GenAI tools, Norwegian PSTs consistently reported low self-assessed expertise and concerns about accuracy and trustworthiness. These findings suggest that while basic adoption levels are generally high, deep and intentional integration remains heavily influenced by institutional support and perceived pedagogical utility.

C. AI as a Collaborative Partner in Lesson Planning

An increasingly influential body of research frames AI not as a replacement for pedagogical expertise but as a collaborative partner that enhances instructional design [18]. This study directly examines how PSTs navigate the balance between leveraging AI’s technical affordances and exercising pedagogical judgment.

Recent empirical studies demonstrate that GenAI tools have considerable potential to support instructional design by enhancing the quality of lesson plans produced by PSTs when compared to those developed through traditional instructional approaches [18]. Within this emerging literature, AI is frequently conceptualized as a collaborative pedagogical partner or a “second mind” that supports idea generation, suggests instructional strategies, and facilitates interdisciplinary connections [18]. By investigating the extent of GenAI utilization and analyzing PST experiences and reflections, this study directly examines how future educators navigate the complex balance between leveraging AI’s technical affordances and exercising pedagogical judgment.

In line with this perspective, Fan *et al.* [1] developed LessonPlanner, a specialized LLM aligned with Gagné’s Nine Events of Instruction. Their controlled comparison showed that LessonPlanner produced plans with enhanced pedagogical coherence and improved alignment with learning outcomes, while reducing users’ cognitive load. This suggests that domain-specific AI tools may better support instructional design than general-purpose systems. Similarly, Cai *et al.* [19] demonstrated that AI-driven critique and feedback systems can markedly improve preservice teachers’ lesson-planning skills, particularly in specialized areas such

as inclusive education. Building upon the notion of AI as an instructional collaborator, Sun and Huang [20] explored AI's collaborative potential in group-based lesson planning, finding that GenAI effectively scaffolded collaborative discussions as an idea generator, discussion facilitator, and resource locator, though human oversight remained critical for pedagogical coherence. Sarkar *et al.* [21] found significant contextual variations in preferences for AI generated lesson plans: elementary teachers preferred human-generated plans for their contextual depth and cultural sensitivity, while secondary educators appreciated AI plans for structural clarity and adaptability, suggesting that optimal AI-human integration varies across educational levels.

Discipline-specific investigations further reveal critical gaps in evaluation. Gurl *et al.* [10] found that while PSTs could identify pedagogical weaknesses in AI-generated mathematics content, they showed significantly limited ability to detect mathematical inaccuracies. Similarly, Clark and Kessel [9] demonstrated concerning patterns of cultural bias and oversimplified narratives in AI-generated social studies lesson plans. Together, these findings highlight the need for discipline-specific AI evaluation training that addresses both content accuracy and representational adequacy.

Despite this collaborative potential, research indicates a persistent gap between basic tool use and pedagogically grounded application. PSTs frequently treat AI generated content as finished products rather than starting points for pedagogical reflection and adaptation [2, 7]. Frøsig and Romero [4] argued that when teachers rely on AI to autonomously generate outputs such as lesson plans, their professional agency diminishes, as they lose the power to act, affect matters, make decisions, and take a stance in the design process. To address this, the authors propose a Hybrid Intelligence approach in which human and artificial intelligence operate collaboratively, moving teachers from passive users of AI generated content to co-creators who actively shape and are shaped by these technologies. Successful integration depends not only on individual competencies but also on supportive institutional environments that encourage experimentation and model critical evaluation practices [8]. Together, these studies suggest that meaningful AI integration in teacher education requires both strengthening PSTs' pedagogical decision making skills and establishing institutional ecosystems that promote informed, contextually responsive, and ethically grounded AI use in lesson planning.

III. THEORETICAL FRAMEWORK

This study is grounded in the AI Literacy framework as articulated in recent educational technology literature [14, 16] and applies it using UNESCO's AI competency framework for teachers [22].

AI literacy refers to a set of interrelated competencies that enable individuals to effectively interact with AI systems, critically evaluate AI generated outputs, and engage in ethical and responsible use of AI technologies [14, 16]. These competencies involve knowledge, skills, and dispositions that work together and are all necessary for effective professional practice. This capability is important because improper use of AI can lead to significant ethical concerns,

including misinformation, algorithmic bias, and privacy violations [16].

UNESCO's AI competency framework for teachers [22] provides the structure for examining AI literacy in this study. The framework identifies 15 competencies organized under five areas: human-centered mindset, ethics of AI, AI foundations and applications, AI pedagogy, and AI for professional development. These competencies develop across three progression levels, including Acquire, Deepen, and Create, with the Acquire level specifically intended to foster the basic AI literacy required by all teachers. In the context of this study, these levels offer a useful lens for interpreting PSTs' interaction behaviors: when a participant issues a single prompt and accepts the output without question, this aligns with pre-Acquire or early Acquire engagement, whereas iterative refinement and critical evaluation of outputs would signal Deepen level competency.

From a teacher education perspective, the UNESCO framework is particularly relevant because it emphasizes competencies that extend beyond technical proficiency, namely reflection, judgment, and ethical reasoning as core professional capacities. Recent studies building on this framework have argued that preservice teachers need structured, hands-on opportunities to engage with AI tools in ways that build their ability to critically evaluate AI generated content, recognize potential biases, and ensure alignment with local curricula and cultural contexts [14, 16]. In this way, the UNESCO framework adds a normative and ethical dimension to AI literacy research, grounding the discussion in how AI should be used in instructional planning and teaching practice, particularly as generative AI tools play a growing role in lesson design.

The TPACK framework [23], extended into AI-TPACK [24, 25], informed the design of the assessment rubric by integrating technological, pedagogical, and content knowledge dimensions. However, TPACK is not employed as a primary analytical framework; rather, the study grounds its analysis in AI literacy theory as defined through UNESCO's progression levels, while drawing on TPACK's integrative logic for instrumentation.

IV. MATERIALS AND METHODS

A. Research Design

This study employed a qualitative case study design [26–28]. The bounded case is defined as PSTs' engagement with a specialized AI lesson-planning tool within a structured teacher-education context at a Turkish public university. This case was selected as theoretically informative because it represents a common yet understudied scenario: PSTs who have received explicit AI literacy training interacting with a purpose-built educational AI tool (rather than general-purpose tools such as ChatGPT) within an authentic practicum setting governed by national curriculum requirements.

B. Ethical Approval and Informed Consent

All research procedures adhered to the ethical principles of scientific research and publication. The study received ethical approval from the Ethics Committee of Sakarya University Rectorate (Decision No: 15; Record No: 49).

Given the inherent power dynamic between researchers and PST participants (who are students in the program), participation was entirely voluntary. Consent forms explicitly stated that non-participation would have no academic consequences. All data were anonymized using participant codes (S1–S10), and interview transcripts were stored on password-protected institutional servers.

C. The AI-Supported Lesson Planning Tool: Madlen

The study employed the AI-powered tool “Madlen,” which supports lesson planning aligned with the Turkish national curriculum framework. Unlike general-purpose AI tools such as ChatGPT, Madlen provides a structured and curriculum-centered approach to lesson planning with several distinguishing features. The tool offers a significant potential for exploring AI-enhanced instructional design in localized, context-specific educational environments.

Madlen operates as a conversational AI interface where users interact through natural language prompts. The tool is designed to generate lesson plan components, including attention-grabbing activities, motivational segments, instructional sequences, assessment items, and summary activities that align with specific learning outcomes from the Turkish Ministry of National Education curriculum. The platform’s predefined instructional tools (e.g.,

attention-getting, motivation, review, transition to the lesson, processing, and summary) align with the standard lesson plan template used in public schools in Turkey, ensuring structural compatibility with practicum requirements.

The tool’s design philosophy positions AI as an assistive partner rather than a replacement for teacher decision making. Users are expected to provide context (grade level, subject area, learning outcomes, desired methods, and duration) and then iteratively refine the AI generated output through follow-up prompts. This design intention reflects an iterative and dialogue-based refinement process and makes the analysis of actual interaction patterns particularly informative, as it allows researchers to examine the gap between intended and observed tool use.

The Lesson Planning Tools module, the primary focus of this study, allows teachers to generate structured lesson plans through a conversational, prompt-guided interface. The AI assistant scaffolds this process by eliciting contextual parameters (e.g., grade level, subject, learning objectives, curriculum alignment), then producing outputs aligned with the Turkish Ministry of National Education (MEB) curriculum alongside international frameworks including IB, AP, UbD, Common Core, A-Level, IGCSE, and SAT (see Fig. 1).

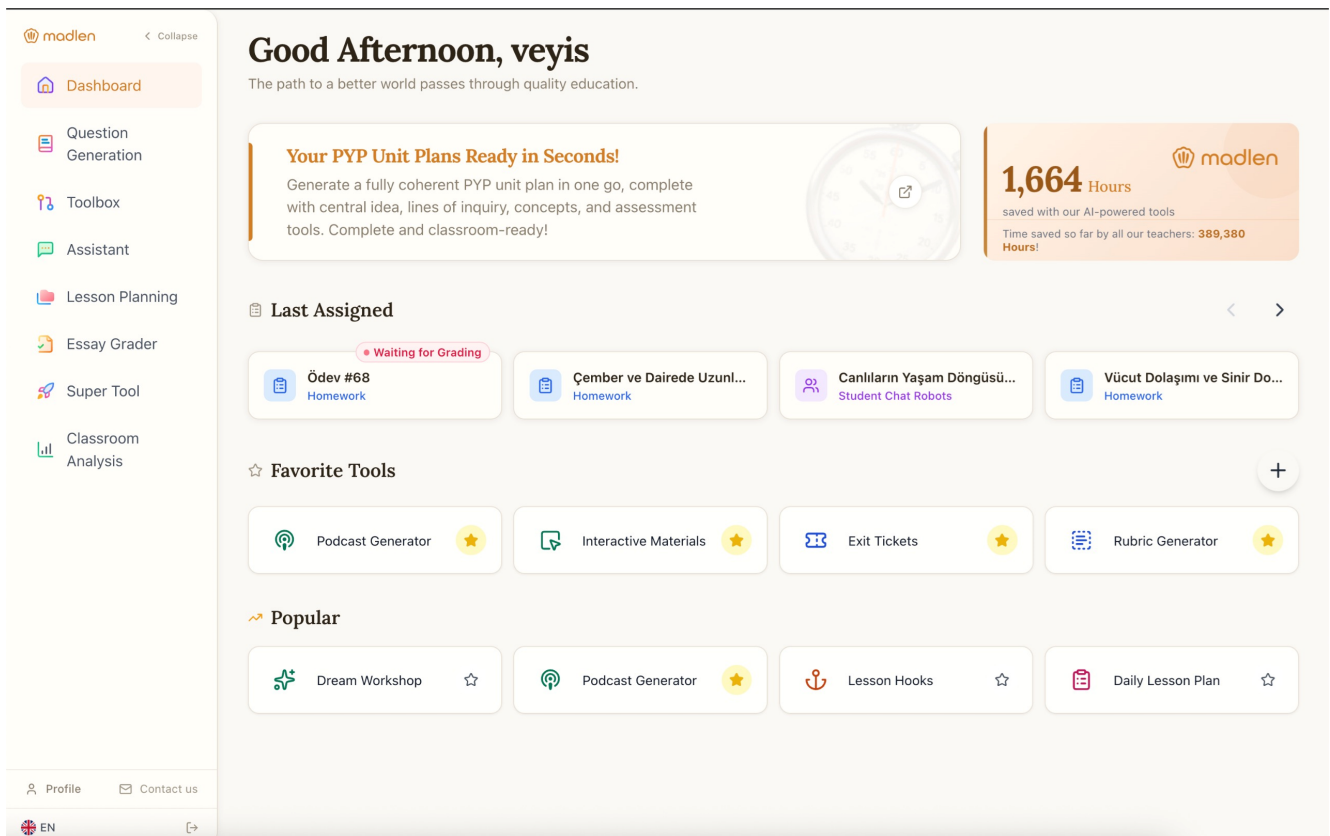


Fig. 1. Interface of Madlen.

D. Participants

The study’s participants included 10 female PSTs in their last year at a public university’s school of education, pursuing undergraduate degrees in Primary Education ($n = 4$) and Turkish Language Education ($n = 6$). According to privacy rules, the preservice teachers who took part have been assigned codes such as S1, S2, and so on.

E. Data Collection Tools

Data were collected using three complementary instruments designed to triangulate between self-reported and observed AI engagement:

First, the pre-interview form consisted of five semi-structured questions to elicit participants’ prior experiences with artificial intelligence, their typical

lesson-planning workflows, and their initial attitudes toward AI-supported planning.

Second, the post-interview form included four semi-structured questions focusing on participants' experiences and perceived changes after preparing lesson plans with Madlen, the AI-assisted planning environment.

Third, AI-supported lesson-planning artifacts were collected, including participants' final lesson plans and written process reflections that described, in a step-by-step manner, how they engaged with Madlen during their planning. Also, the records of PSTs' conversations with Madlen, capturing the actual prompts written by PSTs, Madlen's responses, and any iterative exchanges, were collected as screenshots.

The assessment rubric for AI-supported lesson planning was developed to evaluate participants' use of Madlen and the quality of the resulting lesson plans. The rubric employed a single-dimensional, single-item scale assessing participants' depth of AI engagement on a four-point scale (see Table 1). The rubric assesses a single construct across four levels: No Use (0 points: no benefit derived from AI interaction), Basic Use (1 point: single-step, superficial use with no revisiting), Developing Use (2 points: multi-step, feedback-driven use with limited reflection), and Proficient Use (3 points: strategic use across all planning stages with critical assessment of AI feedback).

Table 1. The assessment rubric for AI-supported lesson planning

Proficiency Level	Description
No Use: 0 point	No benefit was derived from interaction with the AI tool.
Basic Use: 1 point	The AI tool was used in a single step and in a superficial manner to generate an initial output; it was not revisited during the process.
Developing Use: 2 points	The AI tool was used as a guiding resource across multiple feedback-driven steps during the production process; the student revised the generated output but engaged in limited reflection.
Proficient Use: 3 points	The AI tool was strategically utilized throughout all stages of design, production, revision, and evaluation; the student systematically improved the product by critically assessing and responding to the AI generated feedback.

The use of a single-item rubric in this study was a deliberate methodological choice. The primary aim of the rubric was not to measure PSTs' AI competencies in a multidimensional, psychometrically comprehensive manner, but rather to determine the extent to which they integrated AI support into the lesson-planning process. The literature indicates that single-item measures can be appropriate and functional for assessing specific, narrowly defined constructs [29, 30]. However, the study's findings were not based solely on the rubric evaluation. Interviews with PSTs, process reflections, and lesson plan documents were examined together, thereby employing data triangulation. The use of multiple data sources in qualitative research is considered an important element that enhances the study's credibility and the consistency of its interpretations [31]. Accordingly, the limitations of a single-item evaluation were offset by findings from multiple data sources. To strengthen content validity, the rubric was reviewed by two experts in educational technology and revised based on their feedback regarding clarity, relevance, and alignment with AI literacy

competencies.

The rubric was developed drawing on three AI literacy competency dimensions, namely critical evaluation, reflective use, and responsible pedagogical decision making as described in the UNESCO AI competency framework [22]. It was refined following feedback from two domain experts in educational technology to enhance content validity. Each PST's rubric score was determined by two independent raters who jointly evaluated both the complete chat transcript and the corresponding final lesson plan for each of the three planning tasks. The final proficiency score for each participant represents the mean of three rubric ratings (one per lesson plan). Inter-rater agreement for the rubric was calculated using Miles and Huberman's [32] formula, yielding a coefficient of 0.93.

Proficiency levels were classified as follows: Low (mean score 0–1.49), Medium (mean score 1.50–2.49), and High (mean score 2.50–3.00). These thresholds were established a priori based on the rubric's conceptual structure: Low corresponds to no use or basic single-step engagement, Medium to developing multi-step use, and High to proficient strategic use.

The two interview forms were initially drafted by the researchers and refined through expert review by three academics. Two of these experts, who specialized in information and educational technologies, contributed, while one expert in Turkish language education provided feedback on linguistic clarity.

F. Data Collection Process

Data were collected during Spring 2025. After recruitment and informed consent, PSTs were asked to select three distinct learning outcomes from the national curriculum. Based on these learning outcomes, three different lesson plans would be designed using Madlen, which was made ready for practicing in a state school. Prior to initiating the process with preservice teachers, researchers conducted a pre-application interview using a "Preliminary Interview Form". Following this phase, a researcher specializing in instructional technologies delivered two online training sessions to prepare preservice teachers for AI-supported lesson planning. The first session lasted two hours and focused on AI literacy, with particular emphasis on prompt engineering as a foundational skill for effective interaction with AI systems. One week later, a second two-hour session was conducted to introduce Madlen, the AI-assisted lesson planning environment. This session provided hands-on guidance on using Madlen to design lesson plans and instructional activities aligned with national curriculum standards, supported by practical examples. Following these trainings, PSTs designed three distinct lesson plans over three consecutive weeks. Each lesson plan was developed within a one-week timeframe. The planning activities were systematically monitored by the research team. After completing all three plans and implementing them during the field experience, post-interviews were conducted, and weekly reflections were collected. Critically, all chat transcripts from interactions with Madlen were collected for subsequent analysis. The data collection process is visualized in Fig. 2.

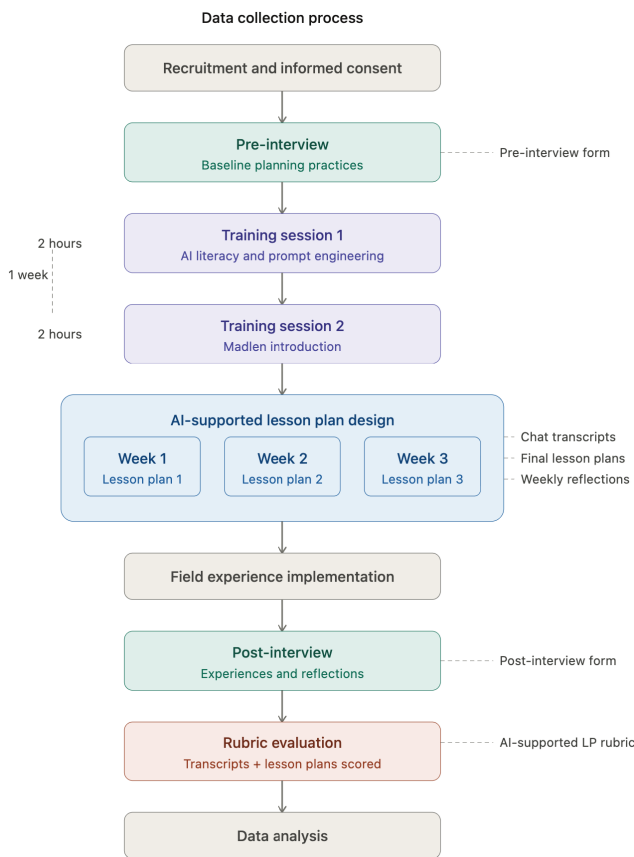


Fig. 2. Data collection process.

G. Data Analysis

Interview recordings and reflections were analyzed following the conventional content analysis approach proposed by Hsieh and Shannon [33]. During the analysis, the lead researcher repeatedly read the data from a holistic perspective, and initial codes were generated from participants' statements. No predetermined coding scheme or set of theoretical categories was used; codes were derived directly from the dataset. Accordingly, an inductive coding approach was adopted throughout the data analysis process.

In the initial stage of the coding process, pilot coding was conducted on a subset of the dataset. The codes generated by the researchers during the pilot phase were compared; codes that overlapped in meaning, were redundant, or lacked clarity in scope were reviewed and revised accordingly. Through this process, code definitions were clarified, and the coding framework was finalized. For example, statements such as "The lesson plans I prepared with AI tools were richer in content, more comprehensive, and supported by a wider variety of sources" (S2) were coded under "preparing rich and comprehensive lesson plans", while statements such as "I was able to design more varied activities in all three lesson plans I prepared with artificial intelligence" (S9) were coded under "increasing activity diversity". Similarly, the statement "I was able to create materials using digital tools" (S4) was coded under "designing digital/technology-supported materials". Following the completion of the pilot coding process, the entire dataset was analyzed using the finalized coding framework.

During the analysis, codes sharing similar characteristics were grouped into categories, which were then organized

under higher-level themes reflecting PSTs' perceptions and experiences regarding AI-supported lesson planning. To ensure consistency and research reliability, the dataset was independently re-coded by a second researcher with expertise in educational sciences. Inter-coder reliability was calculated using Miles and Huberman's [32] formula, yielding high levels of agreement across all datasets (pre-interview: 0.97; post-interview: 0.97; process reflections: 0.98).

To enhance the study's validity, various strategies were employed in line with the criteria for credibility and dependability proposed by Lincoln and Guba [34] for qualitative research. To ensure data triangulation, different data sources were used together: interviews, process reflections, and lesson plans prepared by preservice teachers were analyzed collectively. Comparing findings from different data sources enhanced the consistency and credibility of the research findings [34]. To ensure transparency in the coding process, the stages of data analysis were systematically documented, and detailed explanations of the research process were provided. The independent recoding by a second researcher, as described above, further supported the dependability of the findings.

V. RESULTS

A. Findings from Pre-interviews: Traditional Lesson Planning Practices (RQ1)

To answer the first research question, which aims to characterize PSTs' lesson-planning practices under traditional (non-AI) conditions, a qualitative content analysis of semi-structured interviews was conducted. The resulting codebook and frequency distribution (Table 2) reveal the typical steps and decision-making patterns guiding participants' planning practices. The analysis revealed that PSTs' traditional lesson planning was predominantly outcomes oriented. This is followed by the selection of instructional methods and techniques, the consideration of lesson duration and grade level, the reliance on textbooks, and the inclusion of in-lesson assessment activities. Less frequently, participants referred to the design of customized instructional materials, the accommodation of individual differences, and the use of the national digital education platform or other online resources. Notably, a small number of participants reported incidental use of AI even within accounts of traditional planning, suggesting a gradual normalization of digital supports. Table 2 summarizes these patterns, which are further elaborated in the narrative that follows.

Preservice teachers described a broadly outcome-driven but variably resourced approach to traditional lesson planning, anchoring their processes in stated objectives, method selection, timing, materials, and assessment, while some also incorporated AI support as a drafting aid. One participant described a process that began with reviewing school textbooks and then turned to AI to draft specific lesson plan components, selectively incorporating elements from the AI's suggestions into the final plan:

"When designing a lesson plan, I first examine the school textbooks. Then I ask the AI to write the attention-grabbing, motivational, review, and implementation sections for a lesson plan, one by one, and I tell it to make adjustments to

the lesson plan by specifying the number of lesson hours and the class level. If I don't like the feedback it gives me, I ask for a different lesson plan. I usually determine the techniques suitable for the learning outcomes I will use in the lesson myself and ask them to design the lesson plan accordingly. I write down the parts I like in each feedback on a piece of paper and organise them to create my lesson plan." (S1).

Another participant described a more traditional planning sequence that prioritizes learning outcomes, content, pedagogy, learner engagement, and evaluation, while also attending to time constraints, learner-centeredness, and

diverse learning styles:

"First, I determine the goal and learning outcomes. Then I plan the content and select the methods and techniques to be used. I use learning materials to capture the students' attention. Finally, I complete the process by evaluating. In this process, I pay attention to time management, student-centered content, and different learning styles." (S8).

Together, these accounts illustrate a shared commitment to alignment with learning outcomes, coupled with heterogeneous strategies for sourcing, structuring, and refining lesson components.

Table 2. Preservice teachers' lesson planning practices under traditional (non-AI) conditions

Code	f	Preservice Teacher
Considering learning outcomes	9	S1 S2 S3 S4 S5 S6 S8 S9 S10
Determining teaching methods and techniques	7	S1 S2 S3 S5 S7 S8 S9
Utilizing artificial intelligence tools	3	S1 S2 S4
Considering course duration	5	S1 S5 S6 S8 S10
Acquiring knowledge on the subject	5	S3 S5 S6 S7 S9
Considering traditional lesson planning	4	S3 S4 S7 S10
Making use of the textbook	4	S1 S2 S4 S7
Considering grade level	4	S1 S2 S5 S6
Conducting an assessment during the lesson	4	S5 S6 S8 S9
Focusing on designing teaching materials	3	S5 S9 S10
Considering individual differences	2	S7 S8
Making use of EBA*	1	S4
Observing the teacher's instructional practices	1	S4
Focusing on the activity design process	1	S5
Making use of the internet	1	S7

Note: f = number of participants who reported the code. *EBA: National Education Information Network.

B. Proficiency Levels in AI-supported Lesson Planning (RQ2)

To address the second research question, preservice teachers' proficiency in using the AI tool (Madlen) during lesson planning was evaluated using the Assessment Rubric for AI-Supported Lesson Planning (see Table 1). The results are summarized in Table 3, which presents the distribution of proficiency levels among participants.

The analysis revealed that most preservice teachers demonstrated low to moderate proficiency in employing AI tools for lesson design. Although participants engaged with Madlen to generate ideas, organize lesson components, and align plans with curriculum goals, their use of AI remained primarily assistive rather than transformative. No participant achieved a high level of proficiency, suggesting that the integration of AI tools into pedagogical planning is still in its early stages for preservice teachers.

Table 3. Proficiency levels in AI-supported lesson planning

PST	LP1 Score	LP2 Score	LP3 Score	Mean Score	Proficiency Level
S1	1	1	1	1.00	Low
S2	1	1	1	1.00	Low
S3	1	1	1	1.00	Low
S4	1	1	0	0.67	Low
S5	1	1	1.5	1.17	Low
S6	1	1	0	0.67	Low
S7	0	1	0	0.33	Low
S8	1	2	1.5	1.50	Medium
S9	1	1	1	1.00	Low
S10	1	2	1	1.30	Low

C. PST-AI Interaction Patterns: Evidence from Chat Transcripts (RQ3)

To address the third research question, all chat transcripts were systematically analyzed to characterize PSTs' actual interaction behaviors with Madlen. The analysis revealed three distinct interaction patterns, summarized in Table 4.

Table 4. Interaction patterns observed across PST-Madlen chat transcripts

Interaction Pattern	Description	PSTs	Percentage
Single-Turn, Product-Oriented	One prompt issued; AI output accepted without any follow-up	S1, S2, S6, S7	40%
Minimal Iteration, Logistically Driven	One to two follow-up prompts focused on logistical adjustments (e.g., time, format) rather than content evaluation	S3, S4, S9	30%
Multiple Prompts, Specification-Oriented	Multiple prompts across separate lesson plans with greater contextual detail, but no critical evaluation of AI output quality	S5, S8, S10	30%

Note: The predominance of single-turn, product-oriented interactions (Pattern 1) and logistically driven minimal iteration (Pattern 2) corresponds to Basic Use (1 point) on the rubric scale.

1) Pattern 1: Single-turn, product-oriented interaction

Nearly half of the participants (S1, S2, S6, S7) exhibited a single-turn interaction pattern in which one prompt was issued to Madlen and the resulting output was accepted without any follow-up, revision request, or critical evaluation. This pattern was consistent across all three lesson plans for each participant. For example, S1's entire

interaction for one lesson plan consisted of:

"You are a Turkish language teacher. You will teach the topic of postpositions, conjunctions, and interjections with your 6th-grade students. Students will encounter this topic for the first time, so you will deliver a content presentation before the lesson. After the content presentation is completed, prepare a lesson plan for the second lesson." (S1, Lesson Plan 1).

Following this single prompt, Madlen generated a complete lesson plan, and S1 engaged in no further interaction, including no requests for revision, no questions about pedagogical appropriateness, and no evaluation of the suggested activities. Similarly, S2 issued a single directive prompt specifying a traditional game-based method and a specific learning outcome, accepted the resulting plan without modification, and did not engage further with the tool.

This pattern is noteworthy because these PSTs had received explicit training in prompt engineering and in the iterative use of AI. The gap between training content and observed behavior suggests that knowing about iterative prompting does not automatically translate into practicing iterative prompting.

2) Pattern 2: Minimal iteration, logistically driven

Three participants (S3, S4, S9) demonstrated limited follow-up interaction, but their additional prompts focused on logistical adjustments rather than pedagogical evaluation of AI generated content. S3, for instance, issued an initial prompt requesting a mathematics lesson plan with specific structural components (attention-getting, motivation, review, transition, and processing sections). Her only follow up was a question asking how she could have the students use a ruler for this topic. This question sought a practical implementation detail rather than evaluating whether the AI generated activities were pedagogically appropriate for the stated learning outcome.

S9's interaction pattern illustrates logistically driven iteration most clearly. After receiving an initial lesson plan, S9 noted: "The lesson plan provided will be productive for students, but it will be too long for one class period. Can you update the duration to 40 minutes?" When the revised plan differed from the first version, S9 requested that the activities in the first lesson plan, which she found instructive and thought provoking, be edited to fit 40 minutes. These exchanges, while technically iterative, focused exclusively on time management rather than on evaluating the pedagogical quality, content accuracy, or contextual appropriateness of the AI output.

3) Pattern 3: Multiple prompts, specification-oriented

Three participants (S5, S8, S10) issued multiple prompts with greater contextual detail, yet their interactions still lacked a critical evaluation of the AI output's quality. S5, for example, provided rich contextual information across three separate lesson plans that specified debate topics, traditional game methods, and rhetorical devices, demonstrating stronger prompt-engineering skills. However, the interaction pattern remained unidirectional, with detailed input followed by uncritical acceptance of the output. S10's prompts were among the most detailed in the sample, specifying activity types (i.e., tombala games, card-matching activities), pedagogical structures, and desired lesson plan formats. Despite this specificity at the input stage, S10 did not assess whether the AI generated activities aligned with the stated learning outcomes or were developmentally appropriate for fourth-grade students.

D. Cross-cutting Observations

Several findings were consistent across all three

interaction patterns. First, none of the participants, regardless of which pattern they followed, critically evaluated the outputs they received from Madlen. No one questioned the accuracy or cultural relevance of the content, asked the AI to explain its reasoning, or requested alternative suggestions for comparison.

Second, on the rare occasions when participants did send follow-up prompts, these focused on surface-level adjustments such as timing, formatting, or structural changes. None of the follow-up prompts addressed substantive pedagogical concerns such as learning-outcome alignment, differentiation, assessment quality, or developmental appropriateness.

Third, most prompts followed a formulaic structure that closely mirrored the training content: participants would assign a role ("You are a teacher"), specify the context (grade level, subject, duration), and request a product ("prepare a lesson plan"). Notably absent from this structure were the iterative refinement and critical evaluation strategies covered during the training sessions.

Fourth, participants consistently requested lesson plans that matched the standard Turkish lesson plan template, which includes components such as attention-getting, motivation, review, transition to the lesson, processing, and summary. This suggests that compliance with institutional formats took precedence over deeper engagement with the pedagogical possibilities the AI tool could offer.

E. PSTs Experiences, Perceptions, and Reflections on AI-supported Lesson Planning (RQ4)

Preservice teachers' reflections on their reasons for using the AI tool (Madlen) during lesson planning are presented in Table 5. The table summarizes the key themes derived from participants' responses and illustrates the purposes for which AI was employed in lesson design. These themes reveal how preservice teachers positioned AI as a supportive resource in relation to their established, traditional lesson planning practices.

Participants emphasized that the AI-supported process enriched their lesson plans in terms of content diversity, material variety, and technological integration. As one preservice teacher noted:

"The lesson plans I prepared with AI tools were richer in content, more comprehensive, and supported by a wider variety of sources. The plans I prepared using traditional methods relied more on my own knowledge and experience, which sometimes led to repetitive content or limited materials. The AI-assisted plans, on the other hand, offered much greater variety in both content and activities." (S2).

Another participant highlighted the opportunity to incorporate digital tools and technology-enhanced materials, stating:

"While working with traditional methods, I designed concrete materials myself, but in the AI-supported process, I was able to create materials using digital tools. This allowed me to integrate technology more actively into my lesson plans." (S4).

Similarly, preservice teachers reported that the AI tool facilitated greater flexibility and creative autonomy in developing instructional activities:

"I was able to design more varied activities in all three

lesson plans I prepared with artificial intelligence. If there were points I didn't like, I could easily modify them by giving commands such as, 'Can you do this differently?' It is usually difficult to find distinct activities for each learning outcome, but this wasn't the case with the lesson plans I prepared with AI. That's why I think AI-supported lesson plans provide more alternatives." (S9).

Others contrasted the constrained creativity of traditional planning with the expansive opportunities afforded by the AI environment:

"In the lesson plans I prepared outside of Madlen, the options were more limited. Consequently, I felt more constrained in terms of creativity and developing different approaches." (S10).

Overall, these reflections indicate that preservice teachers perceived AI-assisted lesson planning as a process that broadened their pedagogical repertoire, enabled experimentation with diverse instructional strategies, and enhanced the adaptability of their lesson designs.

Table 5. Purposes of PSTs for using an AI tool in lesson planning

Code	f	Preservice teacher
Preparing rich and comprehensive lesson plans	8	S1 S2 S3 S4 S5 S8 S9 S10
Generating creative and original plan alternatives	6	S1 S2 S3 S4 S9 S10
Increasing the variety of activities	4	S1 S2 S8 S9
Designing digital/technology-supported learning materials	4	S1 S3 S4 S8
Adapting instruction to individual learner needs	3	S1 S3 S10
Increasing the effectiveness of classroom activities	2	S7 S10
Designing activities tailored to learning outcomes	1	S4
Developing diverse assessment tools	1	S8
Developing interactive and engaging content	1	S10

Note: f = number of participants who reported the code.

Table 6 summarizes preservice teachers' experiences and perceptions related to preparing lesson plans using the AI tool called Madlen. The table presents the key codes and themes derived from participants' reflections, highlighting both the perceived advantages and the challenges associated with integrating AI into instructional design. Together, these findings provide insight into how preservice teachers evaluated the affordances and limitations of AI-supported lesson planning.

Most participants expressed favorable views of AI-supported lesson planning and emphasized the practicality, efficiency, creativity, and adaptability that the tool afforded. They noted that the AI tool streamlined lesson preparation, expanded the diversity of instructional ideas, and facilitated differentiated instruction. As one preservice teacher explained:

"First of all, it helped me in terms of practicality and saving time. It helped me to use rich content. It helped me to generate new ideas." (S5).

Another participant underscored the time-saving and idea-generating benefits of AI-assisted planning:

"Artificial intelligence tools primarily save me time when preparing lesson plans. They help me gain new ideas and see different activity suggestions. And by seeing these different activity suggestions, I can think of other things myself." (S3).

Others described AI as a creative catalyst, supporting experimentation with novel pedagogical approaches:

"Thanks to artificial intelligence, I think I can actually experiment with different methods and techniques on different topics. At this point, I think it increases my creativity and contributes to me." (S6).

Table 6. Summary of coded categories on preservice teachers' experiences with AI-supported lesson planning

Theme	Code	f	Participants
Preservice teachers' positive views on the use of AI tools in lesson planning	Saving time	9	S1 S2 S3 S4 S5 S7 S8 S9 S10
	Generating original and creative ideas	9	S1 S2 S3 S4 S5 S6 S7 S8 S10
	Enhancing the effectiveness and content richness of lessons	9	S1 S2 S3 S4 S5 S7 S8 S9 S10
	Providing practicality in instruction	5	S3 S8 S4 S5 S10
	Expanding methodological repertoire	4	S2 S6 S9 S10
	Supporting the content creation process	3	S1 S4 S10
	Supporting student motivation	3	S1 S2 S3
	Suitability for personalised teaching	2	S8 S9
	Providing an opportunity to integrate prior and new knowledge	1	S3
	Promoting personal growth	3	S1 S3 S10
	Ease of access	1	S10
	Providing detailed guidance and recommendations	1	S10
Preservice teachers' negative views on the use of AI tools in lesson planning	Requires teacher supervision	6	S2 S3 S4 S7 S6 S8
	Difficult to use	4	S5 S7 S9 S10
	Creating a lesson plan identical to a traditional one	2	S6 S10
	The implementation of some suggestions may not be feasible	2	S1 S3
	Requiring prompting skills	3	S3 S5 S9
	Lacks direct contribution to the teaching profession	1	S3
	Has a negative impact on the teacher's creativity	1	S4
	Requires time to master	1	S7
	Poor visual representation	1	S6
	Insufficient options	1	S6
	The pilot products failed to meet expectations	1	S9
	The recommendations lack cultural relevance	1	S8

Note: f = number of participants who reported the code.

Participants also valued AI's contribution to inclusive and differentiated learning design, particularly in adapting

content to diverse learner needs:

"I definitely think that designing a lesson plan with artificial intelligence has contributed to me. As I mentioned earlier, it saved time and made it easier to create activities for students with different needs and learning styles in the classroom. When I specify this in the commands, it actually shows them richer, more varied content. Therefore, since I used it during this internship period, I believe it has contributed to me because it enabled me to prepare applications, activities, and plans that would appeal to almost all students in the classroom in a shorter time." (S9).

Despite these positive experiences, several preservice teachers articulated negative views on the pedagogical and practical limitations of AI integration. Concerns centered on the appropriateness of AI generated materials, the tool's initial complexity, and the potential risk of overreliance on AI at the expense of teacher creativity and autonomy.

"In some cases, because artificial intelligence cannot fully understand the pedagogical background of games, the activities it suggests may not be entirely appropriate for the class level and target learning outcomes. At this point, the teacher's support is again very important. The teacher should filter what it says. It presents these things to us, but in my opinion, the teacher should act as a filter to ensure they are appropriate for their class level. And again, the teacher's support becomes very critical." (S2).

The initial learning curve for PSTs was also mentioned as a challenge:

"Madlen is quite rich in content, so using it was daunting at first. I mean, I was a bit confused about what to use and how to use it." (S10).

"At first, I found artificial intelligence very complex to use. But as I started using Madlen, I realized it wasn't complex at all. When we gave the commands correctly, the responses it provided were very positive and useful for us." (S5).

Other participants noted contextual mismatches and expressed reservations about excessive automation:

"On the negative side, some activities did not meet class requirements. My class was crowded." (S1).

"It is quite helpful in the idea generation process, but I don't think it's right to have artificial intelligence prepare the lesson plan directly. Because this can get in the way of the teacher's creativity." (S4).

Overall, the findings suggest that while PSTs recognize AI as a valuable cognitive and creative partner, they also underscore the continued importance of teacher agency and pedagogical judgment in mediating AI generated outputs.

VI. DISCUSSION

The findings of this study point to three key insights. First, most participants used the AI tool by issuing a single prompt and accepting the output without any follow-up, and none of them critically evaluated the responses they received, regardless of their interaction profile. Second, there was a clear gap between what PSTs said and what they actually did: they spoke in their interviews about the importance of teacher oversight and critical evaluation, yet their chat transcripts showed no evidence of these behaviors. Third, training that focused on prompt engineering, while necessary, was not enough on its own to help PSTs develop the kind of critical, reflective engagement that effective AI use requires.

When planning lessons through traditional methods, PSTs demonstrated an emerging set of lesson-planning competencies by aligning learning outcomes, instructional methods, materials, and learner characteristics. These findings are consistent with previous research emphasizing the importance of pedagogical reasoning and reflective practice in lesson design [35–37]. Notably, seven of the ten participants reported already using AI informally during lesson planning. This suggests that the challenge for teacher education is not simply promoting AI adoption but supporting PSTs in developing more critical and reflective forms of AI use.

The second major finding concerns the extent to which Madlen was utilized in lesson planning. Preservice teachers engaged with the tool at low to moderate levels, and none demonstrated sustained or advanced use across planning stages. This finding aligns with previous research showing that PSTs often hold positive attitudes toward AI while possessing only moderate levels of AI literacy and self-efficacy, which limits meaningful pedagogical integration [2, 12, 17, 38–41].

The analysis of chat transcripts offers a window into what PSTs did when working with the AI tool, rather than what they reported doing. Three interaction patterns emerged. In the first pattern, which accounted for the largest share of participants, a single prompt was issued, and the AI's response was accepted without revision. In the second pattern, one or two follow-up prompts were used, but these addressed logistical matters such as timing or formatting rather than the quality or appropriateness of the content. In the third pattern, participants issued more detailed and contextually rich prompts yet still did not evaluate whether the AI generated activities were pedagogically sound or aligned with their intended learning outcomes. This stands in sharp contrast to the Hybrid Intelligence approach proposed by Frøsig and Romero [4], in which teachers and AI operate as co-creators, with teachers retaining the power to act, make decisions, and take a stance throughout the design process. In the present study, even the most engaged participants functioned as specification providers rather than critical collaborators. They directed the AI's input parameters but did not evaluate its output quality. The majority of participants tended toward single-exchange, unidirectional generation followed by passive acceptance of AI outputs. This pattern closely resembles what Tran *et al.* [42] have described as "Linear Extraction" behavior in their process mining study of GenAI use among PSTs. The present findings align with the Linear Extraction profile and suggest that prompt-engineering instruction alone may be insufficient. PSTs also need structured opportunities to critically evaluate AI outputs, revise their prompts, and iteratively refine AI generated content during authentic lesson-planning tasks.

Despite their limited proficiency, participants generally viewed AI-assisted lesson planning positively, emphasizing its potential to enhance efficiency, creativity, instructional diversity, and individualized learning support. These findings are consistent with previous research showing that PSTs recognize AI's potential for enhancing instructional efficiency and pedagogical creativity [20, 43]. At the same time, participants acknowledged the need for teacher oversight to ensure pedagogical quality and contextual

appropriateness, while also expressing concerns about the feasibility of some AI generated suggestions, limited cultural sensitivity, and the risk of overreliance on AI. Similar concerns have been reported in previous studies highlighting both the benefits and limitations of AI-supported lesson planning [5, 10]. In interviews, participants spoke thoughtfully about the need to filter AI outputs, adapt them to specific classroom contexts, and maintain their own pedagogical judgment. S2, for example, described the teacher's role as that of a filter who must evaluate and adjust AI suggestions. S9 expressed appreciation for how AI could support differentiated design. Yet when we examined their chat transcripts, we found virtually no evidence of these behaviors in practice. No participant asked Madlen to revise content for pedagogical quality, questioned whether the activities were developmentally appropriate, or requested alternatives for comparison. This gap between stated beliefs and observed behavior has important implications for how we understand and assess AI literacy. The UNESCO framework [22] outlines a progression from Acquire to Deepen to Create, but this sequence does not account for what our data show. PSTs in this study held the right ideas about AI use and still failed to act on them. Knowing that AI outputs should be evaluated is not the same as actually evaluating them. We use the term "procedural AI literacy" to name this gap between what PSTs understand and what they do. Kerr and Kim [44] reported a similar pattern among pre-service EFL teachers, who interacted with AI in directive, single-turn ways despite holding thoughtful views about its pedagogical role. Chua *et al.* [45] likewise found that PSTs recognized challenges such as the need for precise prompting but did not act on that recognition in practice. From a practical standpoint, this finding carries significant consequences for teacher education. If perception-based measures such as interviews and surveys consistently show that PSTs understand the importance of critical AI use, but interaction data reveal that they do not practice it, then current assessment approaches may systematically overestimate PSTs' readiness to work with AI in meaningful ways. Designing teacher education programs that close this gap will require not only technical instruction but also scaffolded, practice-based opportunities for PSTs to develop iterative prompting, critical evaluation of outputs, and pedagogical adaptation as enacted skills rather than abstract principles.

In summary, while AI tools hold considerable potential to enhance efficiency, creativity, and differentiation in lesson planning, their educational value ultimately depends on the quality of human oversight that accompanies their use. Effective AI integration in teacher education calls for more than technical proficiency; it requires the sustained development of critical, ethical, and reflective competencies that enable PSTs to engage with AI as informed pedagogical partners rather than passive consumers of generated content [46].

Despite these contributions, several limitations should be acknowledged. First, the sample consisted of ten female preservice teachers enrolled in two teacher education programs at a single university. This small and homogeneous sample limits the generalizability of the findings, and the results should therefore be interpreted as exploratory rather

than confirmatory. The gender homogeneity of the sample prevents any claims regarding gender-related differences in AI interaction patterns or perceptions. Future research should employ larger and more diverse samples across multiple institutions, programs, and demographic backgrounds to enhance the representativeness and external validity of the findings.

Second, the use of a single-item rubric to assess AI integration levels, while justified by the study's specific focus and supported by methodological precedent [29, 30], represents a limitation in terms of measurement granularity. Future studies should consider developing multidimensional rubrics that incorporate dimensions such as interaction depth, critical evaluation quality, pedagogical adaptation, and iterative refinement sophistication.

Third, the study was conducted using a single AI-powered lesson planning tool (Madlen) aligned with the Turkish national curriculum. This platform- and curriculum-specific context may have constrained participants' interaction behaviors in ways that differ from those observed with general-purpose AI tools or tools designed for other curricular frameworks. Cross-platform comparative studies would help determine the extent to which the observed patterns are tool-dependent or generalizable across AI-supported lesson-planning environments.

Fourth, chat transcript data were collected through participant-submitted screenshots rather than automated interaction logging. This method, while practical, may have resulted in incomplete data capture. Future research should employ automated logging systems to ensure comprehensive and uninterrupted documentation of human-AI interactions.

Finally, the findings point to a need for longitudinal research designs that can track the development of PSTs' AI interaction skills over time, particularly in response to scaffolded training interventions that combine prompt engineering with metacognitive reflection strategies.

VII. CONCLUSION

This study investigated preservice teachers' engagement with a specialized AI tool (Madlen) for lesson planning by examining their established planning practices, the nature and extent of AI-supported use, their actual interaction patterns with the AI tool, and their perceptions and reflections on AI-assisted lesson design. Four principal conclusions emerge.

First, preservice teachers demonstrated structured, outcomes-oriented planning competencies under traditional conditions, prioritizing learning outcomes, curriculum alignment, grade level, and instructional resources. These results confirm that lesson planning is an inherently complex process that requires the coordination of multiple pedagogical, contextual, and cognitive elements, and that PSTs possess a foundation on which AI-assisted practices could build.

Second, the predominance of single-turn, product-oriented interaction patterns and the complete absence of critical evaluation across all interaction profiles indicate that AI literacy training focused on prompt engineering is necessary but insufficient. PSTs require scaffolded opportunities to practice iterative AI dialogue, critical evaluation of outputs, and pedagogical adaptation as enacted skills within authentic

planning tasks, rather than merely as abstract principles discussed in training sessions.

Third, the documented perception-action gap, in which PSTs articulated the importance of teacher oversight and critical evaluation in interviews while demonstrating no such behaviors in their actual AI interactions, reveals that positive attitudes and declarative knowledge about appropriate AI use do not automatically translate into critically reflective practice. This finding challenges the assumption, implicit in much of the current literature, that perception-based measures adequately capture AI literacy.

Fourth, preservice teachers expressed positive perceptions of AI-assisted lesson planning, highlighting its efficiency, time-saving capacity, and potential to enhance instructional design. Nevertheless, they consistently emphasized the importance of teacher oversight to ensure the pedagogical quality and contextual appropriateness of AI generated outputs. Collectively, these results highlight a balanced perspective in which AI serves as a supportive tool that boosts creativity and efficiency, but the teacher retains an indispensable role in ensuring instructional relevance and integrity.

Future research should employ larger and more diverse samples across multiple institutions, disciplines, and cultural contexts, utilizing mixed methods approaches that integrate qualitative depth with quantitative generalizability. Comparative studies across various AI platforms would help elucidate the impact of tool design and interface features on instructional creativity and pedagogical decision making. Longitudinal designs are needed to track the development of PSTs' AI interaction skills over time, particularly in response to scaffolded training interventions that combine prompt engineering with metacognitive reflection strategies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and design of the study. Material preparation, data collection, and analysis were performed by S.Y.O., N.İ.A., and E.N.S. The first draft of the manuscript was written by S.Y.O., N.İ.A., E.N.S., and H.W. H.W. also contributed critical revision of the manuscript. All authors commented on previous versions of the manuscript. All authors read and approved the final version of the manuscript and agreed to its publication.

REFERENCES

- [1] H. Fan, G. Chen, X. Wang, and Z. Peng, "LessonPlanner: Assisting novice teachers to prepare pedagogy-driven lesson plans with large language models," in *Proc. 37th Annual ACM Symposium on User Interface Software and Technology*, 2024, pp. 1–20. doi: 10.1145/3654777.3676390
- [2] B. Pei, J. Lu, and X. Jing, "Empowering preservice teachers' AI literacy: Current understanding, influential factors, and strategies for improvement," *Computers and Education: Artificial Intelligence*, vol. 8, pp. 1–17, 2025. doi: 10.1016/j.caeai.2025.100406
- [3] S. Doğan, U. Y. Nalbantoğlu, İ. Çelik, and N. A. Doğan, "Artificial intelligence professional development: A systematic review of TPACK, designs, and effects for teacher learning," *Professional Development in Education*, vol. 51, no. 3, pp. 519–546, 2025. doi: 10.1080/19415257.2025.2454457
- [4] T. B. Frøsig and M. Romero, "Teacher agency in the age of generative AI: Towards a framework of hybrid intelligence for learning design," arXiv preprint, arXiv.2407.06655, 2024.
- [5] S. M. Gamlem, J. McGrane, C. Brandmo, S. Moltudal, S. Z. Sun, and T. N. Hopfenbeck, "Exploring pre-service teachers' attitudes and experiences with generative AI: A mixed methods study in Norwegian teacher education," *Educational Psychology*, pp. 27–51, 2025. doi: 10.1080/01443410.2025.2528663
- [6] B. E. Smith, A. Y. Shimizu, S. K. Burriss, M. Hundley, and E. Pendergrass, "Multimodal composing with generative AI: Examining preservice teachers' processes and perspectives," *Computers and Composition*, vol. 75, 102896, 2025. doi: 10.1016/j.compcom.2024.102896
- [7] L. Guan, Y. Zhang, and M. M. Gu, "Pre-service teachers' preparedness for AI-integrated education: An investigation from perceptions, capabilities, and teachers' identity changes," *Computers and Education: Artificial Intelligence*, vol. 8, pp. 1–10, 2025. doi: 10.1016/j.caeai.2024.100341
- [8] Y. Nissim and E. J. Simon, "The diffusion of artificial intelligence innovation: Perspectives of preservice teachers on the integration of ChatGPT in education," *Journal of Education for Teaching*, vol. 51, no. 2, pp. 381–401, 2025. doi: 10.1080/02607476.2025.2469128
- [9] C. Clark and C. Kessel, "I for one welcome our new computer overlords: Using artificial intelligence as a lesson planning resource for social studies," *Contemporary Issues in Technology and Teacher Education*, vol. 24, no. 2, pp. 196–222, 2024.
- [10] T. J. Gurl, M. P. Markinson, and A. F. Artzt, "Using ChatGPT as a lesson planning assistant with preservice secondary mathematics teachers," *Digital Experiences in Mathematics Education*, vol. 11, no. 1, pp. 114–139, 2025. doi: 10.1007/s40751-024-00162-9
- [11] P. Panday-Shukla, "Exploring generative artificial intelligence in teacher education," *Teaching and Teacher Education*, vol. 165, 105088, 2025. doi: 10.1016/j.tate.2025.105088
- [12] K. Hava and Ö. Babayigit, "Exploring the relationship between teachers' competencies in AI-TPACK and digital proficiency," *Education and Information Technologies*, vol. 30, pp. 3491–3508, 2025. doi: 10.1007/s10639-024-12939-x
- [13] A. M. Alzahrani, "Unveiling the shadows: Beyond the hype of AI in education," *Heliyon*, vol. 10, no. 9, e30696, 2024. doi: 10.1016/j.heliyon.2024.e30696
- [14] K. Sperling, C. J. Stenberg, C. McGrath, A. Åkerfeldt, F. Heintz, and L. Stenliden, "In search of artificial intelligence (AI) literacy in teacher education: A scoping review," *Computers and Education Open*, vol. 6, 100169, 2024. doi: 10.1016/j.caeo.2024.100169
- [15] T. T. Wijaya, M. Su, Y. Cao, R. Weinhandl, and T. Houghton, "Examining Chinese preservice mathematics teachers' adoption of AI chatbots for learning: Unpacking perspectives through the UTAUT2 model," *Education and Information Technologies*, vol. 30, pp. 1387–1415, 2025. doi: 10.1007/s10639-024-12837-2
- [16] D. Tagare, T. Karki, and W. Yu, "K-12 teachers' ethical competencies for AI literacy: Insights from a systematic literature review," *Computers & Education*, 105435, 2025. doi: 10.1016/j.compedu.2025.105435
- [17] B. Y. S. Acquah, F. Arthur, I. Salifu, E. Quayson, and S. A. Nortey, "Preservice teachers' behavioural intention to use artificial intelligence in lesson planning: A dual-staged PLS-SEM-ANN approach," *Computers and Education: Artificial Intelligence*, vol. 7, pp. 1–18, 2024. doi: 10.1016/j.caeai.2024.100307
- [18] A. Yorulmaz, H. Z. Okulu, N. Muslu-Komurcu, and H. Cokcaliskan, "Enhancing STEM lesson plans: Preservice primary teachers' collaboration with ChatGPT," *Education and Information Technologies*, pp. 1–31, 2025. doi: 10.1007/s10639-025-13736-w
- [19] H. Cai, B. Han, J. Sun, X. Li, and L.-H. Wong, "Harnessing AI for teacher education to promote inclusive education: Investigating the effects of ChatGPT-supported lesson plan critiques on the development of pre-service teachers' lesson planning skills," *The Internet and Higher Education*, vol. 67, 101022, 2025. doi: 10.1016/j.iheduc.2025.101022
- [20] Y. Sun, and Y. Huang, "Exploring the role of generative AI in collaborative lesson planning for pre-service teachers," *Journal of Computer Assisted Learning*, vol. 41, no. 4, e70098, 2025. doi: 10.1111/jcal.70098
- [21] S. Sarkar, M. Sun, A. Liu, Z. Tian, L. Esbenshade, J. He, and Z. Zhang, "Connecting feedback to choice: Understanding educator preferences in GenAI vs. human-created lesson plans in K-12 education—A comparative analysis," arXiv preprint, arXiv.2504.05449, 2025.
- [22] UNESCO, *AI Competency Framework for Teachers*, Paris: UNESCO, 2023. doi: 10.54675/ZJTE2084
- [23] P. Mishra and M. J. Koehler, "Too cool for school? No way! Using the TPACK framework: You can have your hot tools and teach with them,

- too," *Learning & Leading with Technology*, vol. 36, no. 7, pp. 14–18, 2009.
- [24] F. Karataş and B. A. Ataç, "When TPACK meets artificial intelligence: Analyzing TPACK and AI-TPACK components through structural equation modelling," *Education and Information Technologies*, vol. 30, no. 7, pp. 8979–9004, 2025. doi: 10.1007/s10639-024-13164-2
- [25] Y. C. Fan, B. C. Kuo, P. C. Wu, and C. H. Liao, "A model for developing AI pedagogical and content knowledge in inservice and preservice non-STEM elementary teachers," *Education and Information Technologies*, pp. 1–21, 2025. doi: 10.1007/s10639-025-13813-0
- [26] J. W. Creswell, *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, 3rd ed. Thousand Oaks, CA: Sage Publications, 2013.
- [27] S. B. Merriam, *Qualitative Research: A Guide to Design and Implementation*, Wiley, 2009.
- [28] R. K. Yin, *Case Study Research and Applications: Design and Methods*, 6th ed. Thousand Oaks, CA: Sage Publications, 2018.
- [29] C. Fuchs and A. Diamantopoulos, "Using single-item measures for construct measurement in management research: Conceptual issues and application guidelines," *Die Betriebswirtschaft*, vol. 69, no. 2, pp. 195–210, 2009.
- [30] J. P. Wanous, A. E. Reichers, and M. J. Hudy, "Overall job satisfaction: How good are single-item measures?" *Journal of Applied Psychology*, vol. 82, no. 2, pp. 247–252, 1997. doi: 10.1037/0021-9010.82.2.247
- [31] M. Q. Patton, *Qualitative Research & Evaluation Methods*, 4th ed. Sage, 2015.
- [32] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.
- [33] H. F. Hsieh and S. E. Shannon, "Three approaches to qualitative content analysis," *Qualitative Health Research*, vol. 15, no. 9, pp. 1277–1288, 2005. doi: 10.1177/1049732305276687
- [34] Y. S. Lincoln and E. G. Guba, *Naturalistic Inquiry*, Newbury Park, CA: Sage, 1985.
- [35] J. König, M. Krepf, A. Bremerich-Vos, and C. Buchholtz, "Meeting cognitive demands of lesson planning: Introducing the CODE-PLAN model to describe and analyze teachers' planning competence," *The Teacher Educator*, vol. 56, no. 4, pp. 466–487, 2021. doi: 10.1080/08878730.2021.1938324
- [36] M. Estaiteyeh and I. DeCoito, "Technology-enhanced differentiated instruction in STEM education: Teacher candidates' development and curation of learning resources," *Contemporary Issues in Technology and Teacher Education*, vol. 24, no. 3, pp. 291–312, 2024.
- [37] B. I. F. Abdulla, "Analyzing the structure and effectiveness of lesson plans for pre-service English language teachers," *International Journal of English Learning and Applied Linguistics*, vol. 4, no. 2, pp. 143–162, 2024. doi: 10.21111/ijelal.v4i2.11633
- [38] M. Lucas, P. Bem-haja, Y. Zhang, C. Llorente-Cejudo, and A. Palacios-Rodríguez, "A comparative analysis of pre-service teachers' readiness for AI integration," *Computers and Education: Artificial Intelligence*, vol. 8, 100396, 2025. doi: 10.1016/j.caeai.2025.100396
- [39] H. K. Park and J. Park, "Lesson plan development for pre-service elementary school teachers through collaboration with AI: Focused on ChatGPT," *New Physics: Sae Mulli*, vol. 74, no. 2, pp. 215–226, 2024. doi: 10.3938/NPSM.74.215
- [40] A. Ismail, A. Aliu, M. Ibrahim, and A. Sulaiman, "Preparing teachers of the future in the era of artificial intelligence," *Journal of Artificial Intelligence, Machine Learning and Neural Network*, vol. 44, pp. 31–41, 2024. doi: 10.55529/jaimlenn.44.31.41
- [41] N. Sămărescu, R. Bumbac, A. Zamfiroiu, and M. C. Iorgulescu, "Artificial intelligence in education: Next-gen teacher perspectives," *Amfiteatru Economic*, vol. 26, no. 65, pp. 145–161, 2024.
- [42] P. T. H. Tran, L. Huynh, T.-A. Bien, B. Dang, and A. Nguyen, "Preparing preservice teachers for generative AI in lesson planning: A process mining study of AI mindset and tool-only training," *Journal of Digital Learning in Teacher Education*, pp. 16–32, 2026. doi: 10.1080/21532974.2025.2583516
- [43] L. Kohnke, D. Zou, A. W. Ou, and M. M. Gu, "Preparing future educators for AI-enhanced classrooms: Insights into AI literacy and integration," *Computers and Education: Artificial Intelligence*, vol. 8, 100398, 2025. doi: 10.1016/j.caeai.2025.100398
- [44] R. C. Kerr and H. Kim, "From prompts to plans: A case study of pre-service EFL teachers' use of generative AI for lesson planning," *English Teaching*, vol. 80, no. 1, pp. 95–118, 2025. doi: 10.15858/engtea.80.1.202503.95
- [45] C. K. Heng, R. V. Sathasivam, N. Mafarja, and S. S. A. Rahim, "Exploring the benefits and challenges of AI-driven lesson planning among preservice science teachers," *STEM Education*, vol. 6, no. 1, pp. 1–20, 2026. doi: 10.3934/steme.2026001
- [46] J. Tondeur, O. Trevisan, S. K. Howard, and J. van Braak, "Preparing preservice teachers to teach with digital technologies: An update of effective SQD-strategies," *Computers & Education*, 105262, 2025. doi: 10.1016/j.compedu.2025.105262

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