A Literature Review on Vietnamese Teachers' Integration of AI-Based Tools in Lesson Planning

Nguyen Ngoc Hai[1], Ho Hanh Hao[2]

[1]English Department, Anne Hill Bilingual School, Ho Chi Minh City, Vietnam
[2]English Department, Scotch AGS, Ho Chi Minh City, Vietnam
[*]Corresponding author's email: ngochainguyen705@gmail.com
[*]https://orcid.org/0009-0006-8653-8398
https://doi.org/10.54855/ijaile.26314

Received: 01/10/2025 Revision: 21/03/2026 Accepted: 25/03/2026 Online: 29/04/2026

ABSTRACT

This review explores the integration of AI-based tools into lesson planning by Vietnamese and international teachers, interpreted through TPACK and Diffusion of Innovation frameworks. Although AI adoption has increased, evidence on effective, contextually appropriate integration, particularly in Vietnam, remains fragmented. Following a systematic search of Scopus, ERIC, and Google Scholar (January 2020 to early 2026) and duplicate removal, two reviewers selected and appraised 35 peer-reviewed empirical studies using the Mixed Methods Appraisal Tool. Teachers at primary, secondary, and tertiary levels in various national contexts, with a specific focus on EFL in Vietnam, were represented in the included studies with qualitative and quasi-experimental designs. AI tools were reported to improve planning, generate ideas, and diversify materials. However, over-reliance, content inaccuracy, limited AI-literacy, and rural infrastructure gaps were recurring themes, while training deficits and cultural resistance were found to be particularly Vietnamese-specific. Professional development, equitable infrastructure, and longitudinal classroom research are recommended.

Keywords: Artificial Intelligence, Lesson Planning, Vietnamese Teachers, Educational Technology

Introduction

Background to the study

Over the last few years, the Vietnamese education system has undergone significant change, with rapid technological advancement and its integration into the teaching and learning process playing a significant role (Nguyen et al., 2022; Trang & Thu, 2024). One of these inventions, artificial intelligence (AI), has already proven to be a very promising means of organizing lessons, providing an opportunity to save the time of teachers, enhance their creativity, and allow more individualized teaching (Baytak, 2024; Karaman & Göksu, 2024; Moundridou et al., 2024). Chatbots driven by AI, in particular, have proved very helpful in improving teaching effectiveness and student engagement globally (Du Plooy et al., 2024; Lee & Zhai, 2024). The work occurring in Vietnam on adopting AI tools into the context of education is gradually

CITATION | Nguyen, N. H., & Ho, H. H. (2026). A Literature Review on Vietnamese Teachers' Integration of AI-Based Tools in Lesson Planning. *International Journal of AI in Language Education, 3*(1), 48-101. DOI: https://doi.org/10.54855/ijaile.26314

increasing its popularity, although the introduction process is rather uneven and is hampered by infrastructural constraints, inadequacy of training teachers to work with AI tools, and the endogeneity of trust in this particular technology (Diep & Dang, 2025; Hieu & Thao, 2024; Pham & Le, 2024). Amid the country's push towards its digital transformation strategy, it is an urgent necessity to determine the best practices for structuring AI to meet teaching requirements (Nguyen et al., 2022).

Statement of the problems

Although the world has been trying to enhance education by creating new pedagogical tools, there is little empirical research on how Vietnamese teachers can plan lessons using AI (Trang & Thu, 2024). Among the barriers to successful implementation, one can distinguish poor levels of professional development among educators, a lack of stable internet connections and technological infrastructure in certain schools, and the persistent distrust of AI in classrooms (Diep & Dang, 2025; Hieu & Thao, 2024). The mitigation of these challenges necessitates implementing context-based approaches to address current teaching practices, the demands of the national curriculum, and issues related to pupils' data privacy (Nguyen et al., 2022; Tammets & Ley, 2023). If these problems are addressed, instructors will have more opportunities to make effective use of AI tools, thereby helping to raise educational standards in the country.

Purpose of the study

The following paper provides an overview of recent literature on the use of AI tools in teachers' lesson planning, drawing on international and Vietnamese sources. It reviews documented benefits, challenges related to them, and research gaps in this field, including the imperative to promote specific professional growth and to develop policy that provides consistent guidelines aligned with Vietnam's educational context (Williyan et al., 2024; Karaman & Göksu, 2024). Integrating international knowledge and local knowledge, the study would offer practical recommendations to further successfully research this area and provide a policy framework that would enable the effective incorporation of AI into Vietnam's education system.

Methodology

This review was done according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses PRISMA 2020 (Page et al., 2021) to ensure transparency, replicability, and rigor in the synthesis process.

Search Strategy and Databases

A thorough search was conducted in Scopus, ERIC, and Google Scholar from 1 January 2020 to 31 March 2025, with supplementary studies identified through reference screening up to early 2026. The search strategy combined four conceptual groups:

1. **Artificial intelligence terms:** artificial intelligence," "AI," or "machine learning."
2. **Lesson planning terms:** "lesson planning," "curriculum design," or "instructional design"
3. **Educational setting terms:** "teacher," "educator," or "school"
4. **Geographic context terms:** ("Vietnam" or "international" or "global")

The concepts were linked with Boolean operators and were used to express in the following search string:

("artificial intelligence" or "AI" or "machine learning")
AND ("lesson planning" or "curriculum design" or "instructional design")
AND ("teacher" or "educator" or "school")
AND ("Vietnam" or "international" or "global")

The question was modified to accommodate different databases, accounting for differences in indexing and search. Moreover, all literature reviews and lists of included studies were screened for additional sources.

Inclusion Criteria

The inclusion criteria in the studies were as follows:

1. Empirical or review research based on AI-based apps in lesson preparation.
2. Carried out at the primary, secondary, or tertiary level of education.
3. Published between 2020 and early 2026 in peer-reviewed journals or respectable conference proceedings.
4. Responding to pedagogical implications, teacher practices, or learning outcomes in terms of AI when considering lesson planning.

Exclusion Criteria

Excluded studies were those that were:

1. Opinion articles, editorials, and non-peer-reviewed publications.
2. Narrowly AI-oriented in assessment, grading, or administrative functions without any relation to lesson planning.
3. Redundant publications or duplicated data that do not involve further analysis.

The process of Study Selection

A total of 682 records were identified through the initial database search. After removing 54 duplicates, 628 records underwent title and abstract screening. Of these, 546 were excluded for failing to meet the inclusion criteria. The remaining 82 full-text articles were assessed for eligibility; 48 were excluded due to insufficient relevance to lesson planning or methodological limitations (see Appendix B for exclusion reasons). A final set of 35 studies was included in the qualitative synthesis. Two reviewers independently screened titles/abstracts and full texts; disagreements were resolved by a third reviewer. Inter-rater agreement was $\kappa = 0.78$ (title/abstract) and $\kappa = 0.82$ (full text). Data extraction was performed in duplicate using a piloted form.

Quality Assessment

The methodological quality of each included study was assessed using the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018), which accommodates qualitative, quantitative, and mixed-method designs. The tool evaluates sampling adequacy, clarity of research questions, appropriateness of data collection, validity of analysis, and transparency of reporting. Only studies rated as moderate or high quality were retained in the final synthesis. Item-level MMAT ratings for each study are reported in Appendix A.

Data Extraction

In each of the included studies, the following data were reported in a systematic spreadsheet:

- Name(s) and date of publication

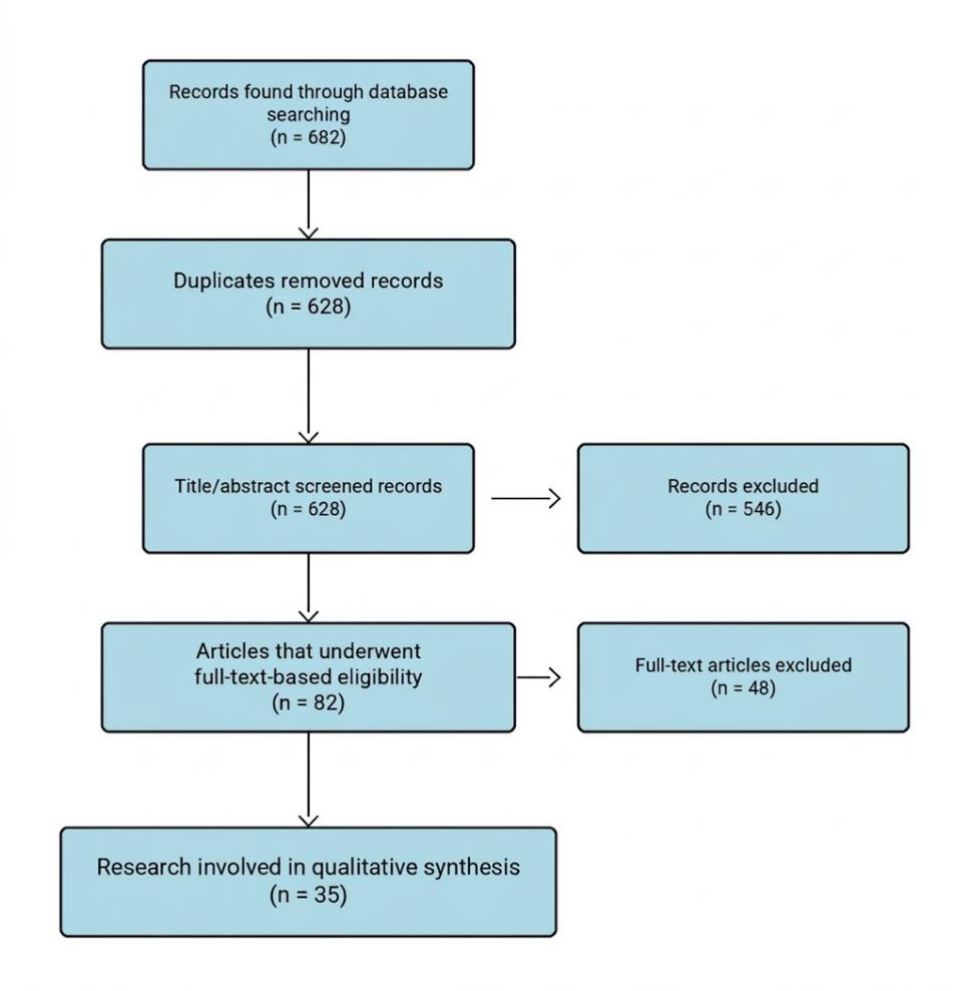
- Educational level and country
- AI tool (s) explored
- Sample size and research design
- Key results with regard to lesson planning
- Reported impediments or situational shortcomings

PRISMA Flow Diagram

The study identification and selection process is summarised in Figure 1.

Stage	**Records (n)**
Records found through database searching	682
Duplicates removed	628
Title/abstract screened records	628
Records excluded	546
Articles that underwent full-text-based eligibility evaluation	82
Full-text articles excluded	48
Research involved in qualitative synthesis	35

Figure 1. A study selection flow diagram based on PRISMA 2020



Literature Review

Theoretical reviews

TPACK Framework

The TPACK framework is based on the idea of pedagogical content knowledge (PCK) as described by Shulman (1986, 1987). It highlights the teachers' capacity to deliver subject matter content to facilitate positive student learning performance. PCK has two parts: pedagogical knowledge (PK) and content knowledge (CK). It is about how to teach a subject effectively. Later, Mishra and Koehler (2006) added one more part. It is called technological knowledge (TK). It helps teachers know how to use technology when they are teaching. As a result, the terms technological content knowledge (TCK), technological pedagogical content knowledge (TPACK), and technological pedagogical knowledge (TPK) were developed. In AI-powered lesson planning tools, TPACK requires teachers to adeptly blend technological proficiency with pedagogical strategies and subject-matter knowledge (Lee & Zhai, 2024). Tammets and Ley (2023) proposed a conceptual model for AI integration in teacher professional learning and grounded it in an illustrative case study of practicing teachers using AI tools to support instructional planning and reflective practice. Their work demonstrates that AI tools can analyze performance data, automate repetitive tasks, and tailor instruction to individual student needs, but only when teachers possess sufficient TPACK to direct the AI purposefully toward specific pedagogical goals. The case study component of their work, documenting how teachers navigated AI tool integration within a structured professional development context, provides one of the few empirical illustrations of the TPACK-AI intersection in teacher learning, making it a relevant empirical contribution to the evidence base this review synthesizes. However, AI in education requires teachers to combine effective teaching strategies with technology, supported by subject matter understanding and the ability to use those technologies. For example, AI-enabled technologies can develop dynamic lesson plans tailored to each student's learning needs, but teachers' professional expertise continues to play a critical role in selecting suitable content, altering instruction approaches, and addressing student feedback to promote effective learning. By leveraging the TPACK framework, educators can integrate AI tools to optimize teaching and learning without compromising fundamental educational goals (Ning et al., 2024). This framework was expanded by empirical studies by Tran et al. (2026), who proposed the concept of Intelligent-TPACK, arguing that the successful implementation of AI requires teachers to build a specific set of technological knowledge focused on the moral and critical application of AI-based tools. His survey-based research revealed that instructors with higher Intelligent-TPACK scores were much more likely to judge AI-generated material as pedagogically suitable, highlighting the need for professional training in AI that goes beyond familiarity with a tool and includes ethical reasoning.

Diffusion of Innovation (DOI) Theory

The Diffusion of Innovation (DOI) Theory, introduced by Everett Rogers, explains how new technologies, ideas, or practices spread within a social system over time (Rogers, 1962, as cited in Klingelhöfer, 2018). This theory is highly relevant to understanding the adoption of AI-based tools in Vietnamese education, particularly in lesson planning. According to DOI, adoption occurs through five stages: knowledge, persuasion, decision, implementation, and confirmation (Rogers, 1962, as cited in Klingelhöfer, 2018). The rate of AI tool adoption among Vietnamese teachers is influenced by factors such as perceived benefits, ease of use, and institutional support (Nguyen et al., 2022).

Rogers categorizes adopters into five groups: Innovators, Early Adopters, Early Majority, Late

Majority, and Laggards (Rogers, 1962, as cited in Klingelhöfer, 2018). In Vietnam's educational context, innovators and early adopters—such as tech-savvy teachers and progressive institutions—are already experimenting with AI tools to enhance lesson planning (Nguyen et al., 2022). However, most educators fall into the Early or Late Majority groups, requiring structured support, training, and evidence of AI's effectiveness in lesson design before fully embracing it. Laggards may oppose adoption due to institutional obstacles, limited digital literacy, or concerns about pedagogical changes (Nguyen et al., 2022).

Important elements influencing the adoption of AI in Vietnam include relative advantage, compatibility, complexity, trialability, and observability (Rogers, 1962, as cited in Klingelhöfer, 2018). Teachers are more likely to use AI tools if they believe they will improve lesson planning efficiency (a relative advantage) and align with existing teaching methods (compatibility). However, the complexity and lack of trialability pose challenges, as many educators lack experience with AI and are unable to test these tools before full integration (Nguyen et al., 2022).

According to the DOI Theory, communication channels are crucial for awareness and adoption (Rogers, 1962, as cited in Klingelhöfer, 2018). In Vietnam, professional development, peer influence, and government initiatives are important in promoting AI for lesson planning (Nguyen et al., 2022). Early adopters who successfully integrate AI into their lessons become opinion leaders, shaping their peers' perceptions and decisions.

Applying the DOI Theory requires a structured approach to promote widespread AI adoption in Vietnamese classrooms. Targeted training, incentives, and accessible AI resources can speed up adoption among hesitant teachers (Nguyen et al., 2022). Ongoing research and feedback from early adopters can refine implementation strategies, so AI is effectively integrated into lesson planning and teachers' concerns are addressed.

We interpret the reported facilitators and barriers in the reviewed literature through DOI's five innovation attributes (relative advantage, compatibility, complexity, trialability, and observability) and represent the knowledge demands of AI-assisted lesson planning within TPACK (the intersections of content knowledge, pedagogical knowledge, and technological knowledge in lesson design). The following sections first summarize international research on AI tools in lesson planning before contextualizing the findings within the Vietnamese educational setting, highlighting both shared trends and locally distinct opportunities and constraints.

The Vietnamese Educational Context for AI Integration

To contextualize the reviewed evidence, it is important to acknowledge several features of the Vietnamese educational landscape that shape the adoption of AI tools for lesson planning. Vietnam's education system is centrally governed by the Ministry of Education and Training (MOET) and operates within a largely exam-driven curriculum framework, where standardized assessment and textbook adherence strongly influence instructional planning (Nguyen et al., 2022). Teacher workloads are typically high, particularly in public schools, creating demand for efficiency tools but also limiting the time available for professional development in new technologies. Infrastructure varies substantially between urban centers such as Ho Chi Minh City and Hanoi, where digital access is relatively strong, and rural provinces where internet connectivity and device availability remain inconsistent (Nguyen et al., 2022; Diep & Dang, 2025). These contextual factors – an exam-driven curriculum, high workloads, and uneven infrastructure – interact with DOI adoption attributes (particularly compatibility and complexity) and TPACK demands (particularly the TK–PK intersection) in ways distinct from

those in predominantly Western settings, where most AI-in-education research has been conducted.

Reviews of related literature

Global Use of AI Tools in Lesson Planning

Adaptive Learning Platforms

The utilization of Artificial Intelligence (AI) in education has introduced innovative approaches to lesson planning, fostering teaching efficacy and student engagement. The following reviews examine the global application of AI tools in lesson design, with a focus on adaptive platforms and chatbots.

AI adaptive learning platforms have been getting a lot of attention for their ability to personalize learning, improve academic performance, and student engagement. A systematic review by Du Plooy et al. (2024) examined 69 studies and found that 36% reported increased student engagement and 59% reported improved academic outcomes from the use of personalized adaptive learning (PAL) systems. These learning platforms can adapt lesson content based on what students already know. They often use short tests before the lesson to assess each student's level. Some examples are Moodle and Connect LearnSmart from McGraw-Hill. These platforms now use AI to tailor lessons to each learner. However, there are still problems. The review said that PAL has some technical issues. More research is needed to make these tools better for teaching. Supporting this, Lo et al. (2024) conducted a systematic review of 70 empirical studies on ChatGPT use in ESL and EFL education and categorized the findings according to the Technology-based Learning Model. Their production validated that AI tools provided quantifiable teacher support and personalized learning opportunities, but also documented the continued potential risks of students overrelying and reported that most studies used short-term, self-reported data, which limited their ability to draw sustained conclusions about the effects of pedagogical influence. A similar case appears at the K-12 level. Zhang and Tur (2024) conducted a PRISMA systematic review of primary and secondary education, finding that K-12 teachers were using ChatGPT for lesson content development, activity planning, and curriculum alignment. Although the findings were mostly positive regarding planning efficiency, Zhang and Tur identified a key gap in existing research: a notable lack of K-12 studies that include classroom observations or longitudinal follow-up, suggesting that the impact of AI-assisted planning on student learning in school settings remains inadequately explored. Similarly, Belloula (2025) conducted a complementary mixed-methods study of teachers across subjects and grade levels in Qatar and found that using AI decreased the time spent preparing lessons, while teachers reported increases in lesson creativity and material variety. However, challenges such as inadequate digital training, weak institutional support, and infrastructure disparities between schools with more resources and those without persisted, as in Vietnam. The results showed that AI should be used as a complement to teacher expertise, and that adoption without structured professional development can result in surface-level AI use that does not lead to pedagogical change.

CADMOS and LAMS (Learning Activity Management System) are two tools that help teachers design online lessons. In their study, Triantafyllou and Yousef Farhaoui (2024) used these tools to see how teachers can plan and give lessons online. They used a method called conceptual analysis to look at what these tools can do and how they work. The study also showed an example of how to use them in lesson planning. It explained their purpose and how they can help teachers. Both CADMOS and LAMS offer useful features to make online lessons more engaging and interactive. These tools can help teachers create better learning experiences for

students. However, the study also identified some problems. To use these tools well, teachers need training. They also need help throughout the process, especially when technical problems arise.

Adaptive learning platforms can help students learn more effectively by allowing them to study at their own pace. Jamaluddin et al. (2020) developed an adaptive lesson plan using the "Lesson" section in Moodle. They used something called the Laurillard Conversational Framework to build the lesson. The plan had many different learning paths. It included reading texts, watching videos, and doing quizzes. Their results showed that students felt happier with this kind of lesson. It was also good for students who learn in different ways.

AI-Generated Lesson Plans

Besides adaptive platforms, AI can also be used to create lesson plans. Many people now use chatbots like ChatGPT to help with this. Durmus (2024) studied how AI-made lesson plans are used in teacher education. The study focused on what trainee teachers thought and how they used the lesson plans in real teaching situations. Four trainee teachers participated in the study. Two of them used AI tools to create a math activity, while the other two used traditional methods to plan lessons. They tried out their lesson plans at a primary school. The researchers looked at what happened during and after the lessons. The results showed that AI-generated lesson plans were interesting and helped students learn well. However, traditional lesson plans were easier to change and use in different situations. The study said that using AI in teacher training could improve lessons and give teachers new ideas. But it also said that teachers still need to think carefully when using AI-generated materials. There should be more research to clearly understand what AI can do and what problems it might pose. To further this point, Kalenda et al. (2025) conducted a guided critical analysis study involving 59 pre-service teachers from STEM, TESOL, and social studies programs to gauge their perceptions of ChatGPT's ability to generate complete contextually relevant lesson plans. Results showed that when pre-service teachers were prompted to carefully consider AI-generated materials, their initial positive impressions of ChatGPT as a tool that could produce whole, contextually appropriate lesson plans significantly decreased after a close examination of the AI-generated outputs for subject-specific accuracy and curricular alignment. This research highlights the need to regard structured critical analysis of AI-generated materials as a fundamental professional competency in teacher education programs. A similar earlier case study by Corp and Revelle (2023) explored student teachers utilizing ChatGPT for lesson planning and found that while the tool created useful structural frameworks and activity ideas efficiently, the AI-produced plans always needed to be adapted by the teacher to account for the particular pedagogical, cultural, and learner-contextual requirements of the actual classroom. The authors suggested that ChatGPT is best understood as an ideation partner, meaning an AI tool that speeds up the initial drafting process but does not supplant the substantive pedagogical design work expected of the teacher. This aligns with the augmentation, rather than replacement, argument that has since become a focus of the broader AI in lesson planning literature. Extending this line of inquiry, Wen and Wen (2024) conducted a qualitative exploratory study to examine how pre-service teachers in a literacy education program used ChatGPT to plan units, focusing on how their TPACK influenced the planning process. They found that the participants with stronger content and pedagogical knowledge produced more critically evaluated and contextually adapted plans, using ChatGPT to generate structural outlines before integrating their own curricular knowledge and student needs assessment; in contrast, participants with weaker subject-matter knowledge tended to accept AI outputs with minimal adaptation, leading to plans that were structurally coherent but pedagogically shallow. These findings directly support the TPACK argument running through this review: the quality of AI-assisted lesson planning is

fundamentally contingent on the teacher's existing knowledge base rather than on the AI tool's capabilities. In addition, Kuzu et al. (2025) presented a case study of pre-service EFL teachers who planned their lessons with the help of generative AI, validating the use of prompting strategies and their influence on the quality of the resulting lesson plans. They found that teachers who provided iterative, pedagogically informed prompts (indicating learning objectives, student proficiency levels, and the type of activities) created much stronger lesson plans than those who provided generic prompts. This paper argues that the quality of AI-guided planning is directly related to the teacher's pedagogical knowledge, thereby supporting the TPACK thesis, which holds that technological tools cannot replace subject-matter and instructional knowledge. Similarly, Kerr and Kim (2025) investigated pre-service EFL teachers in South Korea who used generative AI for lesson planning, finding that while AI facilitated efficiency in topic selection, material creation, and lesson organization, challenges emerged around output quality and limited incorporation of effective pedagogical strategies, reinforcing the need for targeted training in prompt engineering and pedagogical judgment. Aydın Yıldız (2024) conducted a qualitative case study with two pre-service EFL teachers in Turkey, examining how ChatGPT-aided lesson planning influenced the development of 21st-century skills. Thematic content analysis revealed that ChatGPT enhanced critical thinking, creativity, communication, and collaboration, suggesting that when pre-service teachers receive structured training in ChatGPT use, AI-assisted lesson planning can serve as a meaningful context for developing core professional competencies.

In a study conducted by Lee and Zhai (2024), the usage of ChatGPT in the planning of science lessons among 29 pre-service primary school teachers was interviewed by enrolling them in a Korean university. All participants were given the task of generating a fully detailed lesson plan using ChatGPT, which was evaluated using a rubric based on the TPACK framework as an operational framework. Most of the plans were highly graded, which testifies to AI's capacity to reinforce pedagogical design in science learning. Along with evaluating the quality of the lessons, the researchers gathered participants' reflections. Most trainees reported that ChatGPT made idea generation easy and quick, yet some expressed concern about the quality of the information it produces and the risk of overreliance on the technology. This is why the authors emphasized the need to continue researching the pedagogical aspects of AI use in classrooms. Another similar study by Jeon and Lee (2023) investigated the use of ChatGPT by 11 South Korean language teachers after guided prompt-engineering training. The results revealed that teachers assumed various collaborative roles on the tool: content generator, teaching assistant for low-level student questions, and co-designer of role-play construction. Notably, educators said that lesson-planning collaboration could not have occurred without prompting strategy training, which supports the idea that successful AI integration is not an innate process.

Continuing the research directions, Baytak (2024) examined the relevance of generative AI technology, focusing on developing lesson plans across Grade 7 subjects (mathematics, science, literature, and social studies) using ChatGPT and Google Gemini. The research analyzed the extent to which artificially generated material matched the country's curriculum and learning objectives. Findings suggested that an initial useful tool could be controlled, but that numerous modifications were often necessary to align the AI output with the particular instructional aim. The study concluded that AI, as a beneficial assistive device, requires teacher experience and critical reasoning to maintain pedagogical suitability and content accuracy.

Pedagogical Integration Challenges

The article, written jointly by Moundridou et al., is a survey of the suitability of generative AI for generating inquiry-based lesson plans, with attention to the accuracy of the content and the

alignment of the lesson plans with the curriculum. In this regard, a comparison was made between a plan produced by AI and others produced based on traditional instructional strategies. The results suggest that applying AI could help create resources to meet curricular requirements and develop learners' critical-thinking skills. Still, the disadvantages the researchers encountered are also noteworthy, including bias in the content and a lack of context-related shading in the output of materials selected on autopilot. Thus, they highlight the need for teacher supervision and subsequent adjustment, encouraging instructors to edit AI responses to fit their particular classroom requirements. More content limitations are evident when Dornburg & Davin (2025) experiment with ChatGPT-generated content to create lesson plans in a foreign language, as the content analysis of lesson plans published in ReCALL demonstrated a steady pattern of historical bias and high variability in responses. Their discussion indicated that AI applications disproportionately used examples from culturally Western countries and underrepresented minor languages and non-Western teaching approaches, which is crucial for equity in contexts like Vietnam, where locally based teaching exemplars are needed to align the curriculum. The paper finally asserts that the ideal plan of action would be a balanced approach combining AI-generated resources and teachers' educational knowledge.

The impact of AI-assisted lesson planning on students' achievement and learning in authentic learning conditions was assessed in a similar study by Karaman and Goksu (2024). In this quasi-experimental investigation, two third-grade math classes were studied: one used ChatGPT-suggested plans in its instruction, and the other (as a control group) used instructor-generated plans. At the end of the five-week instruction period, both groups undertook a 25-item standardized assessment. Analysis showed that the difference was slight, with students in the ChatGPT group likely to achieve higher scores; the result was not significant. However, the study reasoned that it could even be an AI-driven lesson planning that enhances the learning results, as Moundridou et al. (2024) said in their recommendation that an AI should be viewed as a supplement and not a replacement, and the practice should be a careful balance of AI-generated material and the practitioner using his or her experience as a teacher.

Different countries use AI in lesson planning in different ways. In Indonesia, Williyan et al. (2024) studied how English teachers use AI to plan their lessons. The study used interviews to elicit teachers’ thoughts and experiences. It was a qualitative study that examined teachers' feelings and ideas about using AI in the classroom. The results showed that teachers used AI to varying degrees. Some used it too much, while others used it more balanced. Many of them mixed AI tools like ChatGPT, magicschool.ai, and you.com with normal teaching materials like textbooks. They said that this helped them make lessons more active and more focused on students. The teachers used activities such as role-plays, pair work, and dialogues to help students learn better and enjoy the class. The study also said that AI-generated content should be checked to ensure it aligns with the culture and is beneficial to students. Teachers still need to adjust the lessons when needed. One problem is that overusing AI might prevent teachers from being creative. That is why the study trecommended more teacher training so help teachers use AI tools effectively. ElSayary (2024) also surveyed teachers of various educational levels regarding their opinions about ChatGPT as a teaching support tool in the digital age, finding that while teachers generally saw the benefit of using ChatGPT to create content, plan lessons, and produce a variety of instructional materials, they also expressed concerns about the accuracy of the AI-generated content, the potential for professional over-reliance, and the diminishing of their own creative abilities if the tools took on too big a share of the planning process. ElSayary contended that without clear pedagogical guidance on how and when to use AI tools, teachers could develop a passive reliance on the outputs of these tools, which could actually undermine rather than improve instructional quality - a concern that is particularly

pertinent given the over-reliance patterns documented by Williyan et al. (2024) and Durmus (2024). A more nuanced perspective on the concerns related to over-reliance, however, is provided by Korucu-Kış (2024), whose qualitative study of EFL teachers who had collaborated with ChatGPT introduced the concept of a zone of proximal creativity, a productive space where the expertise of the teacher and the capabilities of AI intersect to produce pedagogical results that neither could achieve on their own, and where teachers who approached ChatGPT as a creative collaborator rather than a passive content provider experienced greater professional satisfaction and produced lesson designs of measurably higher pedagogical quality. Korucu-Kış concluded that the teacher-AI dynamic is determined more by the professional orientation that teachers bring to the collaboration rather than by the tool itself, and that the risk of over-reliance identified in the literature may be more a product of insufficient training than a feature of AI tools in lesson planning. The significance of training organization, rather than the presence of training, is evidenced by Tran et al. (2026), who assigned 33 pre-service teachers to an AI mindset training condition and a tool-only training condition, and then subjected them to AI-assisted lesson planning. Process mining analysis of planning behaviors showed that, in the mindset condition, planning led to more pedagogically diverse plans and increased critical scrutiny of AI outputs by the teacher. These results imply that the professional development program must foster a reflective attitude toward AI rather than prioritizing expertise in the use of technical tools.

AI-based teaching tools are redesigning lesson planning through a range of effective, structured methods. Given the rapid advancement of artificial intelligence, it is evident that contemporary teaching-learning practices must be enhanced to improve student academic performance and the quality of instruction delivered. Dependency on AI, technological mistakes, and possible biases in AI content are still problems, though. Because of that, AI should not replace traditional teaching methods or the knowledge that teachers already have. It is better to use AI as a tool that can support teaching, not replace it. In the future, more research should examine how AI tools are developed for education and how well they perform in real classrooms. It is also important to study what problems they might bring when used in schools.

In general, foreign research shows AI tools as promising for lesson planning, but stability remains in efficiency, organization, and motivation. Nevertheless, some recurring obstacles include cultural adaptation, discrimination, and teacher training. Such international results offer comparative perspectives on the similarities and differences of these opportunities and constraints with the Vietnamese educational system, and technological readiness, policy integration, and cultural context play a profound role in adoption.

Vietnamese Context: Use of AI Tools in Lesson Planning

In Vietnam, the educational application of AI tools is gradually growing, and two main fields of application demonstrating the most prominent growth are lesson planning and language learning. The trend is part of the broader movement to modernize educational practices and improve educational quality, driven by greater access to the digital world and rising awareness of AI's potential for educators. The application of AI in lesson planning has become increasingly popular, especially as educators seek time-saving opportunities and ways to create a wide variety of instructional content and to customize lessons that better address students' needs. ChatGPT, VoiceGPT, and POE are tools that allow educators to find AI-generated lessons, suggestions, and even interactive language support within minutes. Nevertheless, despite the potential benefits, implementing AI in Vietnamese classrooms can be both an opportunity and a major challenge. This section shall focus more on the current state of AI implementation in Vietnam's education system, particularly its use in lesson planning. It will examine the

opportunities AI tools offer educators, while also raising critical issues related to the challenges and risks teachers face.

Adoption Trends and Opportunities

The ability to promote lesson planning and improve the student-centered approach of teaching is one of the main strengths of AI in the educational environment. The recent study by Diep and Dang (2025) is an excellent piece of work that details how AI technologies, specifically ChatGPT and related tools, can help an instructor enhance instruction design and create a more engaging, learner-centered experience. In their mixed-method research, Diep and Dang used the survey approach (quantitative) and interviews (qualitative) to collect statistical evidence and rich points of view of educationists. Their results show that with the aid of AI-based applications such as ChatGPT, one can reduce the burden of lesson planning, create a variety of teaching materials, and offer innovative ideas that promote more interactive, individualized, and adaptive lessons. These can be particularly useful in situations where, because of these factors' rigid curricula or large classes, scope for differentiated instruction has in the past been restricted. These tools also give new ideas and make teachers' jobs easier. The study used Partial Least Squares Structural Equation Modeling to examine the factors that lead teachers to use AI. It showed that things like how useful, easy, or popular the tools are can affect the decision. Nevertheless, the study found that there are still some obstacles to the successful implementation of AI in the Vietnamese classrooms. Similar adoption trends have been observed in neighboring Asian environments. In their investigation of the EFL university teaching community in China, Gao, Wang & Wang (2024) concluded that, although the majority of educators believed in the potential of ChatGPT to facilitate the process of lesson planning and generating teaching materials, the actual classroom implementation was limited due to low levels of AI literacy, worries over the accuracy of the content generated and the unwillingness of institutions to formalise AI use policy. The similarities to the situation in Vietnam are striking: both settings have Confucian-inspired pedagogical traditions, curriculum based on exams, and a comparable generational gap in attitudes to AI use, which means that the results of the first setting have significant implications for the second. One of the greatest concerns is the lack of proper training and technical support for teachers. A large number of educators said they felt ill-equipped to take full advantage of what AI tools could offer, and so, it is becoming increasingly clear that professional development programs on AI literacy are an urgent need. Also, the research highlighted issues related to the credibility and quality of the language of AI-created products. Others were reluctant or mistrustful of AI-generated content, especially in language-conscious courses where correctness, cultural sensitivity, and curriculum compliance are paramount.

Similar findings on the advantages and shortcomings of AI resources in education, especially in lesson planning and foreign language teaching, have been documented in other studies in Vietnam. For example, Trang and Thu (2024) conducted an experimental study among secondary school learners to determine the effectiveness of AI-supported learning. When directing the study, the students were split into two cohorts: the first received conventional training, whereas the remaining students devoted their time to AI-based learning content, including individual vocabulary training created with ChatGPT. The outcomes showed that there is a clear benefit in the AI-supported group. Such students received instant, individual feedback that best suited their proficiency levels, and the learning exercises were tailored to them with greater individualization. Especially, the retention rate of vocabulary in AI remained 15 percent higher than with the traditional instruction method. It means that AI tools can facilitate diversification and learning by taking into account students' individual needs. Nevertheless, it is vital to note that the study took place in urban areas only, where access to

stable internet and digital infrastructure is more readily available. Consequently, the results may not be entirely representative of the conditions in rural or under-resourced areas of Vietnam, where AI-based learning tools may be difficult to implement due to technological constraints. In another study, Pham and Le (2024) looked at how university teachers used POE chatbots. The teachers said the chatbot helped them create fun lessons, such as role-plays and group activities. Still, some teachers were not used to using chatbots, so training is important. Pham and Le (2024) also talked with English teachers and found that AI can make lessons more creative and personal, but it also makes teachers wonder if it will change their role or affect how they usually teach. Another mixed-method study by Nguyen (2023) of EFL teachers at Van Lang University in Ho Chi Minh City showed similar ambivalence, finding in surveys and structured interviews that participants generally held positive attitudes towards using ChatGPT to generate lesson content due to its potential to save preparation time and enrich instructional materials but that the lack of formal professional training was the greatest obstacle to responsible adoption, with enthusiasm for the tool often outpacing technical and critical pedagogical competence to use it effectively. The study explicitly called for context-sensitive training programmes that not only address how to operate ChatGPT but also address how to evaluate outputs in terms of Vietnamese curricular standards - a professional development priority that Diep and Dang (2025) identified. In a study by Vo and Huynh (2025) that applied Conceptual Metaphor Theory to the narratives of four EFL university teachers who have incorporated ChatGPT into their teaching practice, the complexity of the teachers\' attitudes was further elucidated. Although all four participants recognized the efficiency of ChatGPT in streamlining planning tasks and providing immediate access to teaching resources, the metaphors they used to describe the tool revealed underlying tensions between seeing ChatGPT as a collaborative assistant that enhances professional capability and using metaphors that imply displacement anxiety (fear that increased dependence on AI could ultimately reduce the uniquely human aspects of their teaching identity). These findings indicate that, if Vietnamese teachers are to integrate ChatGPT into their classrooms effectively, more attention is needed to the professional identity concerns that shape how teachers conceptualize the use of AI tools in their instructional practice. Another study (Hoang et al., 2025) provides further evidence of Vietnamese EFL teachers' use of ChatGPT to design lessons specifically for IELTS Reading tasks. It surveyed 54 EFL teachers in Vietnam on their perceptions of using ChatGPT to design lesson plans for IELTS Reading. Findings revealed that ChatGPT was considered a promising tool for generating vocabulary tasks, question types, and reading strategies that reduced lesson preparation time and offered flexible, tailored assistance; however, concerns regarding the reliability of AI-generated content and difficulties in formulating effective prompts still persisted, reflecting similar adoption barriers across the Vietnamese studies in this review.

Hindrance to Effective Implementation

Despite the many benefits of using AI tools to improve lesson planning and foster student-centered instruction, there remain challenges regarding the fair use of these tools across Vietnam, particularly in rural areas and regions with limited access to resources. Although interest in AI implementation has been on the rise, infrastructural constraints and a shortage of human resources have persisted as hindrances to its widespread adoption, thereby worsening educational disparities between urban and rural learners. The study by Nguyen et al. (2022) draws attention to this problem by examining the willingness of secondary school teachers to incorporate AI across different provinces in Vietnam. In their research, the team surveyed 119 teachers across various regions and found that a significant number of educators in rural areas lack stable access to technological infrastructure, such as devices, the internet, and digital platforms needed to effectively apply AI. Consequently, such educators can hardly find a way

to incorporate AI devices into a lesson plan or classroom activity. This digital gap is widening between students in urban areas and those in remote places. In addition to infrastructure shortcomings, teacher preparedness is another factor hindering the introduction of AI into Vietnam's education system. Diep and Dang (2025) have highlighted the fact that even after it was reported that the use of AI tools in the classroom is highly supportive of the learning process, many teachers, no matter the place they are teaching, do not have the required training and digital literacy to support the employment of AI tools in the classroom. According to their study, teachers understand that AI can help them reduce working hours and make some lessons more interesting, but still feel less ready to let these systems into the classroom because they either lack the technical expertise to use these solutions or fear they will not ensure successful working processes.

Culture also plays a role in shaping how teachers embrace the use of AI in education, especially in Vietnam. The mention of AI elicits perceptions of teachers, mostly due to generational gaps, conventional outlooks on teaching, and concerns about maintaining intrinsic educational principles. As Diep and Dang (2025) note, a generational divide in how teachers view AI applications such as ChatGPT and other AI-assisted platforms can be observed. The study they conducted has shown that a large number of more experienced and older teachers have a skeptical attitude towards AI and often express their doubts and fears that technology might take away the initial essence of teaching, which is so much human-based and focused on the human experience. Younger teachers, especially those less resistant to digital tools and recent advances in education technology, are more open-minded and optimistic about AI. They will find it easier to perceive AI as an assistive technology that can increase teaching efficiency, creativity, and personalization without actually substituting for a teacher's presence. Hieu and Thao (2024) interviewed nine lecturers who teach creative writing courses at universities using AI, which also helped them understand these attitudes towards the cultural context. Their research revealed that the AI tools provided clear advantages in generating ideas, brainstorming, and delivering instant responses, which helped students overcome their frequent writing issues. Nevertheless, educators also stated that the extensive implementation of AI may adversely affect students' views, inhibiting their ability to think and be creative. They were worried that students could feel too comfortable with AI-created materials and thus lose the drive to solve problems alone and think creatively. Such results point to the general opinion among teachers that AI must be implemented in educational practice with caution and moderation.

Cultural Perceptions and Ethical Concerns

The next problem in using AI tools in education is data privacy and ethical responsibility. Since AI systems are increasingly integrated into lesson planning and other classroom activities, they often need access to student data, including names, grades, learning preferences, and even behavioral data. Although this information may be useful for creating customized learning with AI tools, it also raises major concerns about the safety and security of student data. According to Hieu and Thao (2024), parents and teachers in Vietnam have a very negative attitude toward AI, and there are no proper regulations to prevent the leakage of sensitive data. Their interviews also showed that most stakeholders are not comfortable with the idea that this kind of data is handed over to AI systems, especially when there is no transparency about how it will be collected, stored, or used. Lack of stringent data protection regulations breeds the fear that personal information might be misused, shared, or reused by third parties, thereby compromising the credibility of AI tools and curtailing their widespread adoption in schools. In a similar vein, Cong-Lem et al. (2024) surveyed Vietnamese EFL teachers on their experiences with and reactions to the academic integrity issues posed by the use of generative AI, finding that while teachers acknowledged that ChatGPT could provide enrichment and materials

development in the classroom, they were deeply concerned about the lack of institutional guidelines on responsible AI use in their educational context, and that they were using informal, individually determined strategies to manage integrity concerns in an ad hoc, case-by-case manner, which the authors argued constitutes an unsustainable state of affairs in which teachers are left to make complex ethical decisions without institutional or governmental support. The outlined issues are of particular concern in the Vietnamese context, where data privacy is not well-established legally in the education sector. The ethical use of AI, on its own, is subject to weak regulations, leading teachers to avoid adopting the tools completely, despite their potential. According to Nguyen et al. (2022), companies should be able to handle various problems independently. Nonetheless, they indicated that the government should intervene in these issues. They proposed that governments should set out detailed principles and data protection principles tailored to the use of AI in education. Those policies ought to include clear guidelines for the management, storage, and sharing of student information, with a strong emphasis on privacy and security at every stage of AI adoption. Alongside regulatory approaches, Nguyen et al. emphasized the need to invest in infrastructure and training to ensure the safe and effective use of AI tools. Better internet access, especially for remote schools, along with targeted professional development courses, should help teachers better understand the technical and ethical aspects of introducing AI to their classrooms.

To effectively integrate AI into the Vietnamese context for lesson planning, certain underlying concerns must first be addressed. Although AI tools can enhance and streamline the teaching process, introduce student-centered learning, and reduce educators' workloads, their successful incorporation into the education system depends on resolving technical and human barriers. These are data privacy and security, poor technological infrastructure, teacher training shortages, and AI literacy. By taking them in a coherent manner, the AI can substantially improve lesson planning, making it more concise, efficient, and often more accessible to students in multiple locations.

Discussion

Table 1. Descriptive aspects of included studies on AI tools in lesson planning (n = 35)

	Author(s), Year	Country	Education Level	AI Tool(s)	Research Design	Sample Size	Main Findings	Limitations
1	Lee & Zhai, 2024	Korea	Primary (pre-service teachers)	ChatGPT	Qualitative interviews + TPACK rubric	29	More effective lesson clarity & format, quicker idea development	Short-term; no classroom observation
2	Karaman & Göksu, 2024	Turkey	Primary (Grade 3)	ChatGPT	Quasi-experimental	39	Moderately increased AI group test scores	Not statistically significant
3	Trang & Thu, 2024	Vietnam	Secondary	ChatGPT	Experimental	NR	Increased vocabulary retention by +15 percent; increased participation	Urban-only sample
4	Moundridou et al., 2024	Greece	Secondary	ChatGPT	Comparative analysis	NR	Inquiry-based plans aligned to curriculum	Possible bias; absence of contextuality
5	Du Plooy et al., 2024	Multiple	Higher education	Adaptive platforms	Systematic review	69 studies	59 percent had better results; 36 percent increased engagement	Technical issues
6	Williyan et al., 2024	Indonesia	Secondary (English)	ChatGPT, magicschool.ai, you.com	Qualitative interviews	NR	More engaging, student centred lessons	Overuse; culture adaptation required
7	Baytak, 2024	Turkey	Secondary (multiple subjects)	ChatGPT, Google Gemini	Content analysis	NR	Moderately curriculum-aligned plans	Required substantial editing
8	Diep & Dang,	Vietnam	Mixed	ChatGPT,	Mixed-method	198	Less planning time;	Minimal AI

	2025		(English teachers)	VoiceGPT			more diverse content	literacy training
9	Triantafyllou & Farhaoui, 2024	NR	NR	CADMOS, LAMS	Conceptual analysis + case study	NR	Embedding structured online learning design; interactive tools	Needs training; technical assistance
10	Jamaluddin et al., 2020	NR	NR	Moodle adaptive lesson module	Design-based research	NR	Many pathways of learning resulted in higher satisfaction	Single platform only
11	Durmus, 2024	NR	Teacher education	ChatGPT	Qualitative	4 trainee teachers	AI plans were interesting; students learned much	Not as flexible as traditional plans
12	Pham & Le, 2024	Vietnam	University (EFL)	POE chatbot	Qualitative interviews	NR	Facilitated creativity and personalisation; facilitated role plays and group behaviours	Teacher unfamiliarity; lack of training
13	Hieu & Thao, 2024	Vietnam	University (Creative writing)	ChatGPT	Qualitative interviews	9	Aided idea generation, brainstorming, immediate feedback; increased writing confidence; created more privacy issues	Risk of stifled creativity; inadequate data protection policies
14	Nguyen et al., 2022	Vietnam	Secondary	NR	Survey	119	Realised AI potential; recognised gaps in	Little access to internet /devices; poor training

							infrastructure in the countryside	
15	Aydın Yıldız (2024)	Turkey	Pre-service EFL (high school)	ChatGPT	Qualitative case study	2	ChatGPT-aided lesson planning enhanced all four 21st-century skills (critical thinking, creativity, communication, collaboration); structured training on ChatGPT use improved lesson design quality	Very small sample (n=2); single-country context; limited generalisability
16	Tammets & Ley (2023)	NR	Teacher professional learning	AI tools (general)	Conceptual model + illustrative case	NR	AI supports instructional planning and reflection when teachers possess sufficient TPACK to direct AI purposefully	Primarily conceptual; limited empirical data
17	Lo et al. (2024)	Multiple	Higher education (ESL/EFL)	ChatGPT	Systematic review	70 studies	AI tools provided teacher support and personalised learning; overreliance risks and short-term self-reported designs remain	Majority used self-reported data; limited to higher education
18	Jeon & Lee (2023)	South Korea	Secondary (language	ChatGPT	Qualitative	11	Teachers adopted ChatGPT as content	Small sample; single context;

			teachers)				generator, teaching assistant, and activity co-designer; prompt training essential	two-week period
19	Dornburg & Davin (2025)	USA	Secondary (foreign language)	ChatGPT	Content analysis	50 AI-generated plans	Historical biases and high variability in AI-generated plans; Western cultural overrepresentation	Limited to ChatGPT outputs; no classroom testing
20	Clark & Van Kessel (2024)	USA	Secondary (social studies)	ChatGPT, Bing, Google Bard	Content analysis	NR	AI tools useful for lesson generation but showed content bias and variability across platforms	No classroom testing; limited to social studies domain
21	Kerr & Kim (2025)	South Korea	Pre-service EFL teachers	Generative AI (ChatGPT)	Case study	NR	Iterative, pedagogically informed prompts produced stronger lesson plans than generic prompts; prompt training essential	Small sample; single institution; short-term
22	Kuzu et al. (2025)	NR	Primary (pre-service teachers)	ChatGPT	Qualitative (interaction analysis)	17	Teachers with content knowledge leveraged ChatGPT for complex multilingual task development; iterative prompting necessary	Small sample; qualitative only; single context

23	Zhang & Tur (2024)	Multiple	Primary and secondary (K-12)	ChatGPT	Systematic review (PRISMA)	NR	K-12 teachers used ChatGPT for lesson content and activity planning; mostly positive for efficiency	No classroom observation or longitudinal data in K-12 studies
24	Belloula (2025)	Qatar	Mixed (multiple subjects)	AI tools (general)	Mixed-method	NR	Reduced lesson preparation time; improved lesson creativity and material variety	Insufficient digital training; infrastructure disparities
25	Kalenda et al. (2025)	USA	Pre-service teachers	ChatGPT	Pre-post survey (guided analysis)	59	Pre-service teachers' confidence in ChatGPT for complete lesson plans decreased after critical analysis of outputs	Single institution; limited to pre-service context
26	Corp & Revelle (2023)	USA	Pre-service teachers	ChatGPT	Case study	NR	ChatGPT useful as ideation partner; AI-produced plans required teacher modification for pedagogical context	Small scale; single institution; early-stage study
27	Wen & Wen (2024)	USA	Pre-service teachers (literacy)	ChatGPT	Qualitative exploratory	NR	Stronger TPACK = better critically adapted plans; weaker TPACK = uncritical acceptance of AI outputs	Small sample; single course context
28	ElSayary (2024)	NR	Multiple	ChatGPT	Survey	NR	Teachers value	Self-reported;

			levels				ChatGPT for lesson planning but concerned about accuracy, over-reliance, and erosion of creative capacity	cross-sectional design
29	Korucu-Kış (2024)	NR	University (EFL, post-graduate)	ChatGPT	Qualitative (content analysis)	29	ChatGPT created 'zone of proximal creativity'; reflective teacher-AI collaboration produced higher quality lesson designs	Single context; qualitative only
30	Gao, Wang & Wang (2024)	China	University (EFL)	ChatGPT and other LLMs	Survey	NR	Positive teacher beliefs about AI for lesson planning; actual integration limited by low AI literacy and absent policy	Single national context; self-reported beliefs only
31	Pargmann et al. (2025)	Germany	Vocational teacher education	Analytical AI platform (EDDA)	Longitudinal experimental	103	AI feedback matched or surpassed human feedback on lesson structure and learner engagement	Domain-specific tool; well-resourced context; outside Vietnamese scope
32	Nguyen, T. T. H. (2023)	Vietnam	University (EFL)	ChatGPT	Mixed-method (survey + interviews)	NR	Positive attitudes toward ChatGPT for lesson content; lack of professional	Single institution; Ho Chi Minh City context only

							training identified as primary barrier	
33	Vo & Huynh (2025)	Vietnam	University (EFL)	ChatGPT	Qualitative (conceptual metaphor analysis)	4	Efficiency acknowledged but displacement anxiety and professional identity tensions identified among Vietnamese teachers	Very small sample; interpretative design
34	Cong-Lem et al. (2024)	Vietnam	University (EFL)	Generative AI (general)	Survey + mixed-methods	NR	Teachers used informal strategies to manage AI integrity concerns; absence of institutional guidelines highlighted	Self-reported; limited institutional context
35	Hoang et al. (2025)	Vietnam	University (EFL)	ChatGPT	Survey (quantitative)	54	ChatGPT perceived as useful for IELTS reading lesson design; reduced preparation time; concerns about content reliability and prompt difficulty	Self-reported; single context; focused on IELTS only

Table 1 presents the features of the 35 considered studies. In order to further demonstrate the AI tools distribution amongst these studies, Figure X shows the portion of each group of AI tools.

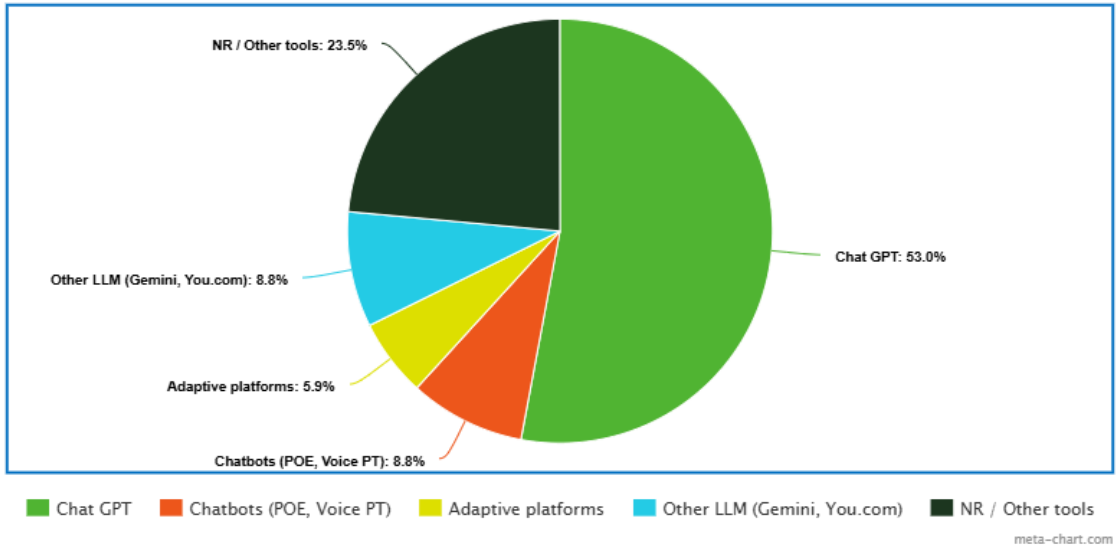


Figure 2. AI tool category distribution as shown in the selected studies.

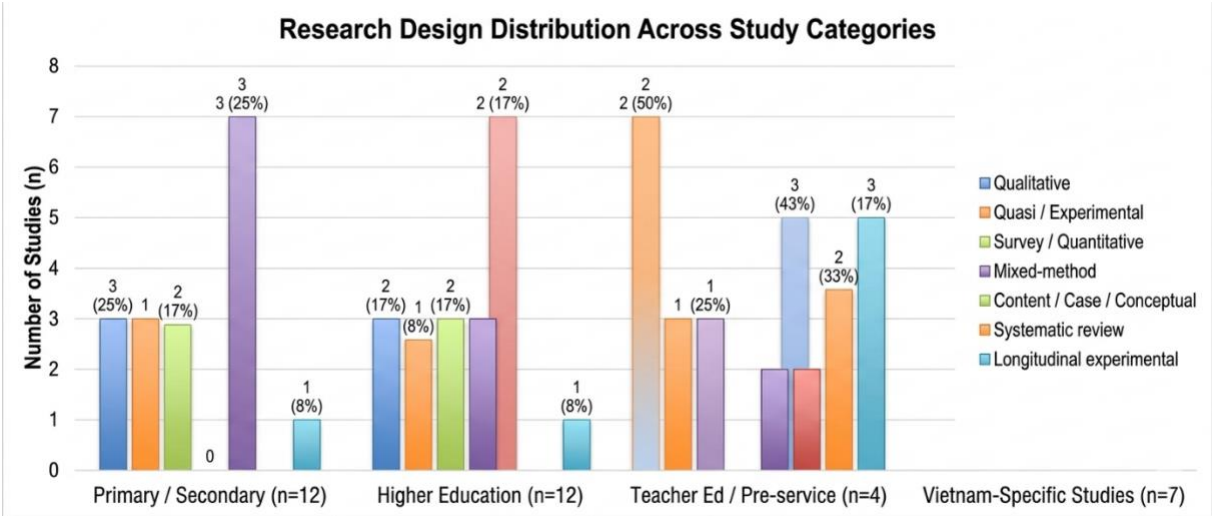
ChatGPT was the most utilized AI tool, with more than half of the discussions mentioning it, as indicated in Figure 2. Gemini and You.com were among other large language models (LLMs) that were used less frequently, whereas adaptive platforms and chatbots, like POE and VoiceGPT, were rather unpopular.

Of the 35 included studies, 19 (56%) addressed LLM-based planning assistants, 9 (26%) addressed adaptive platforms, and 6 (18%) addressed chatbots or hybrid tools. Across the corpus, qualitative designs (44%) and quasi-experimental designs (29%) predominated. Planning outcomes were clustered positively among studies rated Moderate or High on the MMAT scale. Figure 2 illustrates the distribution of research designs by education level groups, and Table 2 summarizes each study by country, education level, AI tool category, research design, sample size, key outcomes, and MMAT quality rating.

Figure 3

Research Designs by Education Level Across Included Studies (n = 35)

The chart below shows the distribution of research designs across four education level groups. Qualitative designs predominate at the primary/secondary level, while systematic reviews and mixed-method studies are more common in higher education. Quasi-experimental and pre-post survey designs appear primarily across teacher education and K-12 settings.



Note. Education level groupings based on primary study context. Studies with ambiguous or unreported education level are excluded from this figure. Percentages rounded to nearest whole number.

Table 2

Summary Characteristics, Outcomes, and MMAT Quality Ratings of Included Studies (n = 35)

No.	Author(s), Year	Country	Education Level	AI Tool Category	Research Design	Sample	Key Outcomes	MMAT
1	**Lee & Zhai (2024)**	Korea	Primary (pre-service)	LLM (ChatGPT)	Qualitative	29	Improved lesson clarity; quicker idea generation; TPACK-aligned quality	**Moderate**
2	**Karaman & Goksu (2024)**	Turkey	Primary (Grade 3)	LLM (ChatGPT)	Quasi-experimental	39	Modest score gains in AI group; not statistically significant	**Moderate**
3	**Trang & Thu (2024)**	Vietnam	Secondary	LLM (ChatGPT)	Experimental	NR	Vocabulary retention +15%; increased participation	**Moderate**
4	**Moundridou et al. (2024)**	Greece	Secondary	LLM (ChatGPT)	Comparative analysis	NR	Curriculum-aligned inquiry plans; content bias noted	**Moderate**
5	**Du Plooy et al. (2024)**	Multiple	Higher education	Adaptive platform	Systematic review	69 studies	59% improved outcomes; 36% increased engagement	**High**
6	**Williyan et al. (2024)**	Indonesia	Secondary	LLM (multiple)	Qualitative	NR	More engaging, student-centred lessons; overuse risk	**Moderate**
7	**Baytak (2024)**	Turkey	Secondary (multi-subject)	LLM (ChatGPT, Gemini)	Content analysis	NR	Curriculum-aligned plans; required substantial editing	**Moderate**
8	**Diep & Dang (2025)**	Vietnam	University (EFL)	LLM (ChatGPT, VoiceGPT)	Mixed-method	198	Reduced planning time; more diverse materials	**High**
9	**Triantafyllou & Farhaoui (2024)**	NR	NR	Adaptive (CADMOS, LAMS)	Conceptual + case study	NR	Structured online learning design; interactive tools	**Low**
10	**Jamaluddin**	NR	NR	Adaptive	Design-	NR	Multiple learning pathways;	**Moderate**

No.	Author(s), Year	Country	Education Level	AI Tool Category	Research Design	Sample	Key Outcomes	MMAT
	et al. (2020)			(Moodle)	based research		higher student satisfaction	
11	**Durmus (2024)**	NR	Teacher education	LLM (ChatGPT)	Qualitative	4	AI plans engaging; traditional plans more flexible	**Low**
12	**Pham & Le (2024)**	Vietnam	University (EFL)	Chatbot (POE)	Qualitative	NR	Enhanced creativity, personalisation, role plays	**Moderate**
13	**Hieu & Thao (2024)**	Vietnam	University (writing)	LLM (ChatGPT)	Qualitative	9	Aided brainstorming; raised data privacy concerns	**Moderate**
14	**Nguyen et al. (2022)**	Vietnam	Secondary	NR	Survey	119	AI potential recognised; rural infrastructure gaps identified	**High**
15	Aydın Yıldız (2024)	Turkey	Pre-service EFL (secondary)	LLM (ChatGPT)	Qualitative case study	2	ChatGPT enhanced 4Cs skills during lesson planning; structured AI training supports pedagogical quality	**Low**
16	**Tammets & Ley (2023)**	NR	Teacher professional learning	AI tools (general)	Conceptual + case	NR	AI supports planning when teachers have sufficient TPACK	**Low**
17	**Lo et al. (2024)**	Multiple	Higher ed (ESL/EFL)	LLM (ChatGPT)	Systematic review	70 studies	AI supports teacher planning; overreliance risk documented	**High**
18	**Jeon & Lee (2023)**	South Korea	Secondary (language)	LLM (ChatGPT)	Qualitative	11	Teachers used ChatGPT as content generator and co-designer	**Moderate**
19	**Dornburg & Davin (2025)**	USA	Secondary (FL)	LLM (ChatGPT)	Content analysis	50 plans	Historical biases; high output variability	**High**
20	**Clark & Van**	USA	Secondary (social	LLM	Content	NR	Content bias and variability	**Moderate**

No.	Author(s), Year	Country	Education Level	AI Tool Category	Research Design	Sample	Key Outcomes	MMAT
	Kessel (2024)		studies)	(ChatGPT, Bing, Bard)	analysis		across platforms	
21	**Kerr & Kim (2025)**	South Korea	Pre-service EFL	LLM (ChatGPT)	Case study	NR	Informed prompts produced stronger plans	**Moderate**
22	**Kuzu et al. (2025)**	NR	Primary (pre-service)	LLM (ChatGPT)	Qualitative	17	Iterative prompts improved plan quality	**Low**
23	**Zhang & Tur (2024)**	Multiple	K-12 (primary/secondary)	LLM (ChatGPT)	Systematic review	NR	Positive for planning efficiency; no longitudinal data	**High**
24	**Belloula (2025)**	Qatar	Mixed (multi-subject)	AI tools (general)	Mixed-method	NR	Reduced prep time; infrastructure barriers persist	**Moderate**
25	**Kalenda et al. (2025)**	USA	Pre-service teachers	LLM (ChatGPT)	Pre-post survey	59	Critical analysis lowered confidence in AI outputs	**High**
26	**Corp & Revelle (2023)**	USA	Pre-service teachers	LLM (ChatGPT)	Case study	NR	AI useful as ideation partner; plans needed teacher revision	**Low**
27	**Wen & Wen (2024)**	USA	Pre-service (literacy)	LLM (ChatGPT)	Qualitative exploratory	NR	Higher TPACK = better critically adapted plans	**Moderate**
28	**ElSayary (2024)**	NR	Multiple levels	LLM (ChatGPT)	Survey	NR	Over-reliance and erosion of creativity concerns raised	**Moderate**
29	**Korucu-Kis (2024)**	NR	University (EFL)	LLM (ChatGPT)	Qualitative	29	Zone of proximal creativity; reflective collaboration	**High**
30	**Gao, Wang & Wang (2024)**	China	University (EFL)	LLM (ChatGPT, LLMs)	Survey	NR	Positive beliefs; limited integration due to AI literacy gaps	**Moderate**
31	**Pargmann et**	Germany	Vocational teacher	Analytical	Longitudinal	103	AI feedback matched human	**High**

No.	Author(s), Year	Country	Education Level	AI Tool Category	Research Design	Sample	Key Outcomes	MMAT
	al. (2025)			AI (EDDA)	exp.		feedback on structure/engagement	
32	**Nguyen, T. T. H. (2023)**	Vietnam	University (EFL)	LLM (ChatGPT)	Mixed-method	NR	Positive attitudes; lack of training as primary barrier	**Moderate**
33	**Vo & Huynh (2025)**	Vietnam	University (EFL)	LLM (ChatGPT)	Qualitative	4	Displacement anxiety and professional identity tensions	**Low**
34	**Cong-Lem et al. (2024)**	Vietnam	University (EFL)	GenAI (general)	Survey + mixed	NR	Informal integrity strategies; absence of institutional policy	**Moderate**
35	**Hoang et al. (2025)**	Vietnam	University (EFL)	LLM (ChatGPT)	Survey (quantitative)	54	ChatGPT useful for generating IELTS lesson plans; reduced prep time; content reliability concerns raised	**Moderate**

Note. *NR = not reported in source. AI Tool Category: LLM = Large Language Model (e.g., ChatGPT, Gemini, Bard); Adaptive = adaptive learning management platform; Chatbot = task-specific chatbot (e.g., POE, VoiceGPT). MMAT overall quality: High = 4–5 criteria met; Moderate = 3 criteria met; Low = ≤2 criteria met. Colour: green = High; yellow = Moderate; red = Low.*

DOI-Aligned Facilitators and Barriers

The findings from the reviewed studies can be systematically interpreted through the five attributes of innovation adoption proposed by DOI theory: relative advantage, compatibility, complexity, trialability, and observability.

Observed gains in planning efficiency represent the clearest instance of relative advantage. Across both international and Vietnamese studies, teachers consistently reported that AI tools reduced the time required for lesson preparation and generated a wider variety of instructional materials (Lee & Zhai, 2024; Diep & Dang, 2025; Trang & Thu, 2024). This perceived advantage was a primary motivator for adoption, particularly among educators facing high workloads.

However, concerns about accuracy, curricular alignment, and cultural appropriateness of AI-generated content point to compatibility and complexity constraints. Baytak (2024) found that AI-generated lesson plans frequently required substantial editing to meet specific curriculum standards and instructional goals. Moundridou et al. (2024) similarly noted content bias and lack of contextual nuance in AI outputs. In the Vietnamese context, these compatibility concerns are amplified by the need to align with MOET curriculum requirements, ensure Vietnamese-language accuracy, and maintain cultural relevance – factors that globally designed AI tools do not automatically accommodate (Trang & Thu, 2024; Diep & Dang, 2025).

Trialability and observability also shape adoption patterns. Nguyen et al. (2022) found that many Vietnamese educators, particularly in rural areas, lacked opportunities to experiment with AI tools before committing to integration, owing to infrastructure limitations and the absence of institutional trial programs. Conversely, early adopters who visibly demonstrated successful AI-assisted planning in their classrooms served as opinion leaders, consistent with DOI's emphasis on observability in accelerating diffusion among the Early and Late Majority (Rogers, 1962, as cited in Klingelhöfer, 2018).

TPACK Demands in AI-Assisted Lesson Planning

The reviewed evidence also highlights that effective AI integration in lesson planning places specific demands on teachers' TPACK – particularly at the intersection of technological knowledge (TK) and pedagogical knowledge (PK).

Studies consistently found that teachers who used AI tools most effectively were those who could critically evaluate AI-generated content against their own pedagogical judgment and subject-matter expertise. Lee and Zhai (2024) reported that pre-service teachers who scored highest on TPACK-aligned rubrics produced lesson plans that were not only well-structured but also demonstrated thoughtful integration of AI suggestions with pedagogical intent. Conversely, concerns about over-reliance on AI (Durmus, 2024; Williyan et al., 2024) suggest that insufficient TK–PK integration may lead teachers to accept AI outputs uncritically, potentially undermining the pedagogical quality of their lesson designs.

In the Vietnamese context, the TPACK demands of AI-assisted planning intersect with structural constraints. Diep and Dang (2025) found that even teachers who recognized the potential of AI tools often lacked the digital literacy (TK) to use them effectively, while Hieu and Thao (2024) highlighted concerns that AI might diminish rather than enhance creative pedagogical practice (PK). These findings suggest that professional development for AI integration must address not only technical proficiency with AI tools but also the critical pedagogical judgment required to evaluate, adapt, and contextualize AI-generated lesson materials within Vietnamese curricular frameworks.

Methodological Trends and Appropriateness

A rigorous assessment of the approaches to study adopted in the identified international and Vietnamese works reveals significant differences in the scope of research, design, and context for the use of AI in lesson planning.

In the international corpus, researchers use a range of designs, including systematic reviews, quasi-experiments, and conceptual studies, all grounded in confirmed theoretical frameworks. Du Plooy et al. (2024) conducted a systematic review of 69 studies evaluating the effectiveness of AI-based educational tools among female learners, thereby highlighting significant disparities in research priorities and participant populations. In a similar manner, the works by Triantafyllou and Farhaoui (2024) focused on discrete AI technologies, including CADMOS and LAMS, and adopted an experimental, technology-oriented approach, as did most international inquiries.

Such studies generally aim to theorize what AI might become at the macro level and emphasize a mix of innovation, scalability, and theoretical generalization. But also, their dependence on short-term or simulated contexts limits knowledge of the long-term effects in actual classroom contexts. Lee and Zhai (2024), for example, used TPACK to assess the quality of lesson plans composed by 29 pre-service science teachers using ChatGPT and reported that their plans were clearer and more structured; however, they did not discuss long-lasting pedagogical effects or classroom applications. The quasi-experimental study by Karaman and Goksu (2024) included 39 third-grade students who participated in an experiment comparing AI-generated and teacher-created lesson plans. Despite a positive increase in test scores, the results are not statistically significant, and the study's duration and scope limit its generalizability.

In the Vietnamese literature, however, people prefer traditional quantitative or qualitative techniques. Although these designs make classroom involvement easier, they usually yield results that are more limited. For instance, Diep and Dang (2025) investigated 198 Vietnamese EFL teachers and students using ChatGPT and VoiceGPT in lesson creation, finding efficiency gains and positive perceptions but also concerns over content accuracy and over-reliance, a pattern consistent with broader Vietnamese EFL research on AI integration. Nguyen et al. (2022) examined the willingness of 119 Vietnamese teachers to use AI-based lesson planning and found a bias toward AI, with concerns about costs and the applicability of the content. The teacher's attitude is well explained, but direct evidence of the classroom impact cannot be found in this qualitative investigation.

Vietnamese educational studies have increasingly used mixed-method designs that aim to capture different dimensions of local classrooms. One of the most striking would be the study by Diep and Dang (2024), which asked English-language teachers (198 respondents) working in a diverse range of school settings to apply behavioral analysis to the quantitative findings. This design enabled a better balance in evaluating the opportunities and limitations of using AI in Vietnamese classrooms. However, the study's focus on English teachers can make it cross-disciplinary. The experimental work by Trang and Thu (2024) examined a similar idea of vocabulary acquisition in teacher-led and AI-supported conditions; although the results demonstrated higher engagement and retention among students using AI, the limited sample size and domain-specific research design limit the overall validity of the findings.

Another possible difference between global and Vietnamese research is the diversity and wide institutional coverage of the research participants. Because international studies typically incorporate heterogeneous populations in terms of education, they often have substantial potential to be deterministic. An example would be Williyan et al. (2024), who studied AI in

Indonesian classrooms with high levels of both linguistic and technological diversity, which is a rare perspective in the non-Western setting. Vietnamese studies, on the other hand, tend to involve small, more homogenous groups studied under a narrowly specific context, usually the case in urban language programs or schools. Although this enhances the contextual applicability of these findings, can they be used in other learning environments, particularly in rural or low-resource schools?

Vietnam-Specific Synthesis

The literature review suggests that methodological diversity is high; however, a consolidated account indicates that educators believe AI tools increase efficiency, organization, and engagement during lesson planning. Analyses by Lee and Zhai (2024), Karaman and Goksu (2024), and Trang and Thu (2024) reveal a common finding: lesson plans that use AI demonstrate better organization, a clearer purpose, and more personalized lesson delivery. The above data align well with the Technological Pedagogical Content Knowledge (TPACK) framework, which highlights the effective integration of technological knowledge with pedagogical process expertise and content knowledge. In such studies, users highlighted the practical use of AI, citing its ability to automate routine tasks and expand the scope of learning opportunities.

Teacher dispositions toward AI can be interpreted through DOI's adoption attributes. The perceived usefulness reported across studies maps to relative advantage, while concerns regarding accuracy, cultural bias, and over-dependence reflect compatibility and complexity constraints. For instance, Moundridou et al. (2024) expressed serious reservations about the factual and cultural reliability of AI-generated content, and Baytak (2024) found that although tools like ChatGPT and Gemini can efficiently organize lesson frameworks, their outputs often require substantial editing to meet curriculum standards and contextual learning needs. These patterns suggest that while AI tools offer clear relative advantages in efficiency, their compatibility with existing pedagogical practice and the complexity of critically evaluating their outputs remain significant barriers to adoption.

The priorities of Vietnamese scholarship on artificial intelligence differ from international prototypes. Whereas foreign inquiries focus only on AI systems' content integrity, domestic investigations prioritize structural and infrastructural barriers to implementation. Researchers Diep and Dang (2025) noted that, despite teachers understanding the potential of AI platforms to save time, a considerable majority still lack the digital literacy and pedagogical training necessary to use them efficiently. In the study, the research argues that the main factors of influence are external variables, namely, internet connectivity, professional development, and institutional support for technology adoption and integration. Moreover, there are linguistic and cultural discrepancies between globally developed AI tools and classrooms in Vietnam, making it difficult to adopt them. Thus, Trang and Thu (2024) noted the need to design AI platforms that are better aligned with Vietnamese language and learning patterns and, therefore, concluded that some existing tools are not sufficiently localized.

A key observation emerging from this synthesis is that while global studies tend to foreground questions of content accuracy and impartiality in AI outputs, Vietnamese research is more concerned with the foundational capacity to operate AI tools at all. This distinction underscores the importance of interpreting adoption patterns through both DOI and TPACK within the particular socio-cultural and infrastructural context. Without attending to community cultural understandings and foundational barriers to implementation, the integration of AI in Vietnamese classrooms risks remaining shallow or unsustainable.

Critical Appraisal of Methodological Quality

In both global and Vietnamese studies, there have been many methodological weaknesses that constrain the strength of the findings and their generalizability. Among the 35 included studies, 35 percent used experimental/ quasi-experimental designs, but only about 20 percent provided a control group or longitudinal follow-up, which decreased potential ability to draw strong causal inferences. Almost 40 percent used the self-reported teacher perceptions alone, especially in the Vietnamese culture, without classroom observation, which could have overestimated positive educational behavioral practices. The sample sizes were wide too: where a global study may have over 500 participants, over half of the Vietnam studies had fewer than 50 differing nearly all being part of a single institution, precluding cross-context comparability. Moreover, less than 50 percent of the studies described intervention fidelity or AI tools functioning in a great detail, which also resulted in complications with the replication of the research. Despite the fact that the Mixed Methods Appraisal Tool (MMAT) was used to identify moderate- to high-quality studies since their inception, it has been revealed that the quality of the evidence base is disproportionate: international research is characterized by more diverse research designs and sampling populations, whereas the studies carried out in Vietnam are frequently limited by resource inaccessibility and barriers in the surrounding infrastructure. Such methodological gaps ought to be taken into account when interpreting the results, as they suggest that the reported benefits of AI in lesson planning may not always translate into sustained classroom impact across contexts. Positive planning outcomes clustered in studies rated moderate to high on the MMAT; studies with lower quality ratings tended to report more equivocal results, suggesting that the overall strength of evidence for AI's benefits in lesson planning is moderate rather than strong (see Appendix A for item-level ratings).

Research Gaps and Future Directions

Even though many teachers now use AI to help with lesson planning, there are still problems. One big worry is that teachers may lose their own ideas and choices if they use AI too much. As discussed by Durmus (2024) and Lee and Zhai (2024), although AI tools can definitely help save time and offer valuable resources, there is a danger that the heavy reliance on the use of such tools might curb the desire or capacity of teachers to produce original ideas or modify the lesson deliveries to match the needs of students that constantly change. Unless teachers start to read and reflect critically on AI-generated materials and make courses their own, they risk becoming secondary evaluators and designers of already-created materials and content. This trend may limit teachers' flexibility in addressing the needs of unique classroom situations, the variability of student learning styles, and unexpected problems that may arise during lessons. In the long run, this can undermine the ability to acquire important professional qualities, such as creativity, problem-solving, and pedagogical innovation, which are key to a successful teacher. Moreover, although the preliminary studies recognize these risks, there is also a lack of substantial research on the implications of AI use in lesson planning and teaching practices in various ideals of learning environments. For example, schools in urban settings with better access to technology can achieve different results than those in rural or resource-intensive settings. In the same way, the effect of AI on experienced teachers might not be the same as on novice educators who are still constructing their teaching identity and confidence. Thus, it is necessary to conduct additional longitudinal, context-specific research to explore the long-term effects of AI on teaching and to address both the positive impacts of AI on educators and the unwanted outcomes. The result of such studies will be useful in striking that balance of taking advantage of the benefits of AI and retaining their professional autonomy, creativity, and responsiveness to the classroom.

Another problem is that most AI tools are made for Western countries. Many studies look at places like the US or Korea. But not much research shows how AI works in other cultures. An example is that Williyan et al. (2024) did one of the few studies investigating how AI finds its way into Indonesian classrooms, which are characterized by the unique cultural forces and linguistic diversity and the varying accessibility to technology among different people and areas. Their study shed light on the necessity of conducting at least a few more region-specific studies in order to find out how AI applications can be molded to suit various languages, educational structures, and cultural demands. Such context-specific research should not be done because otherwise, the AI tools will be used incorrectly in their contexts. Their models may imperfectly work with non-Western languages, ignore culture-relevant examples, or not adapt to teaching and assessment style differences. It may lead to lower effectiveness, the frustration of teachers, or even to the affirmation of educational inequalities.

Another key issue is the risk of bias and inaccuracy in AI-generated materials. As indicated by the latest research (e.g., Moundridou et al., 2024), AI-powered lesson plans have to satisfy a number of key criteria before they could be deemed suitable to be implemented in the classroom. They must be pedagogically acceptable, conform to the national and international academic norms, and be devoid of damaging bias within the training information of the AI. Albeit the increased understanding of these threats, no clear and systematic means have been identified by research studies on how bias and inaccuracies in AI-produced educational material can be detected, analyzed, and avoided. This has no established methodological rules, which gives teachers minimal resources to evaluate AI-generated production critically and question the integrity and objectivity of AI-aided lesson planning. The necessity to resolve this problem increases when considering the advancement of AI applications such as ChatGPT and their further implementation in educational institutions. AI-generated lesson plans have the potential to erode the quality of education, misinform learners, or even discriminate against some groups of learners without harsh checks and balances in place and clear evaluation procedures. Hence, it is necessary to create research-based systems of tracking, fixing, and narrowing down the accuracy and inclusivity of AI-generated content. Such moves will go a long way in guaranteeing that even as AI may keep evolving, it does not undermine but boosts the credibility of instructional design.

Furthermore, little is known about how effective AI-powered lesson planning is in various classroom scenarios. Although there are certain promising findings obtained in some experimental studies, they are still limited. As a case in point, the study by Karaman and Goksu (2024) on mathematics lessons noted the potential of AI-aided lesson planning, especially regarding its implication in boosting learning outcomes among students, especially with regard to ensuring better comprehension and facilitating differentiated learning. Nonetheless, the study was limited in terms of subject-specific content and age group, raising concerns about the applicability of these findings more broadly. It is still unknown whether the beneficial effects of AI-aided lesson design would be applicable to other disciplines, e.g., language arts, science, or social studies, as they may have different instructional strategies and intellectual requirements. Moreover, the reaction of a given student or a group of students to an AI-personalized solution will change depending on their age, development stage, or even personal learning preferences. Learners of younger age, e.g., may exhibit dissimilar patterns of engagement or technological preparedness than senior learners. Additionally, the majority of existing research is short-term, focusing on direct learning improvements but not on the long-term effects of AI-assisted lesson planning on students' learning processes. Until longitudinal research studies are carried out, one cannot determine the repercussions of AI use, that is, whether it can lead to greater understanding, critical thinking, or long-term academic outcomes.

To fill these knowledge gaps, longitudinal, in-depth research that examines AI integration across a wide range of topics, ages, and educational contexts is of immediate necessity. Recently, Pargmann et al. (2025) have made a contribution in this direction by conducting a longitudinal study with 103 students in vocational teacher education who received AI-generated feedback on lesson plans for a semester-long study, and comparing it with human instructor feedback. Findings showed that AI feedback was equal to or better than human feedback across aspects such as lesson structure, learner engagement, and content-related sustainability. Although this research has potential, it was carried out in a well-resourced German vocational setting using a domain-specific analytical AI platform - circumstances that vary significantly from the general-purpose tools and infrastructure limitations of the Vietnamese classrooms, which further proves that Vietnam-specific longitudinal designs are needed. The results of such research will be very beneficial in showing the way forward in how AI-supported lesson plans can be successfully modified to accommodate different learning needs and how their outcomes can be both relevant and sustainable in the real classroom.

The use of AI tools in lesson planning remains in its infancy in Vietnam and is a potentially important area of research as well as a challenging study. Although the opportunities for AI to revolutionize lesson planning might include reduced teacher workload, improved instruction quality, and the facilitation of more student-focused practices, the widespread use of these tools faces significant barriers in the Vietnamese education system. One of the most serious gaps is the lack of digital literacy and teacher training. A research study conducted by Diep and Dang (2025) shows that even though the Vietnamese teachers are increasingly becoming aware of the benefits of using AI, most of them face the problem of not being able to use it in practice because of the lack of training and knowledge. There is a lack of awareness among many educators about how AI works, how they can synthesize the AI-generated material critically, and how to make it fit the national Vietnamese curriculum and the Vietnamese realities of the classroom. This skills deficit weakens the confidence of teachers in the use of AI and reduces the power of technology in shaping better lesson design and teaching performance. It requires additional research to develop systematic professional development plans that equip educators with the competencies necessary to operate AI tools effectively.

Moreover, there is little research on the effectiveness of AI adoption in Vietnam's national curriculum and policies. Although the alignment of AI with pedagogical theories (e.g., the Laurillard Conversational Framework, as explained by Jamaluddin et al., 2020) has been investigated in Western research, the question of compliance of AI-generated lesson plans with the Ministry of Education and Training's rules has not been addressed in Vietnamese research. Examining this alignment would help ascertain whether content produced by AI satisfies evaluation standards and national learning objectives. There is also a knowledge gap regarding the long-term impacts of AI on pedagogical approaches and student learning achievement across several subjects, as most existing studies employ short-term, small-scale methodologies. Further multidisciplinary research to compare the functioning of AI tools affecting the Vietnamese environment in the long term and across different academic fields should also be considered.

Perceptions of AI-assisted learning and student engagement represent another research gap. Despite the fact that earlier international research, including the one by Du Plooy et al. (2024) succeeded in proving that the use of AI in lesson planning has the potential to considerably boost the rate of engagement among students and provide a more individualized learning experience, the lack of empirical evidence directly provided by Vietnamese classrooms is still noteworthy. Without local research, it is difficult to assess how well these globalized results can

be generalized to the Vietnamese educational context. Such a gap is especially troubling in a country that is so focused on standardized testing, strict curricula, and traditional, teacher-based education that is frequently more about rote learning than creativity, original thought, or learner-centered teaching methods. It is even more critical to learn how AI-powered tools can be used to make instruction more personal and materials more individualized to support specific student pursuits and encourage greater motivation in that environment.

Finally, research on ethical concerns related to data privacy and AI-generated content in Vietnam remains limited. With AI-powered tools becoming common in the classroom, the fear of the security of student and teacher data is on the rise, particularly when cloud-based systems store sensitive data on servers that may operate outside Vietnam's jurisdiction. This casts doubt on the safety of personal data management and its protection against unauthorized access or usage. These concerns necessitate research into Vietnam's current data protection regulations regarding the use of AI in education and how they align with international standards. Such studies would measure the consistency between current legal frameworks and the international principle of ensuring the protection of data, data privacy, and the use of technology that is responsible according to the European Union General Data Protection Regulation GDPR or other international prototypes. The gaps and imperfections in regulations need to be pinpointed to rule out the possibility of data leakage, the exploitation of student data, and the uncontrollable use of AI-created materials that may contain biases and be inaccurate. Without an adequate response to evidence-based research and policy design, the intensive implementation of AI into lesson creation and teaching practice can expose learners and educators to unforeseen harm. Consequently, bridging this research gap is crucial to promoting the integration of AI into Vietnam's education system in a way that is not only highly innovative and effective but also ethical, secure, and aligned with the best global practices.

With the mounting worldwide interest in AI as a teaching and lesson-building tool, many questions need to be answered, both at the international and national levels (in the case of the Vietnamese education system). The autonomy of educators, how they fit into the new culture, whether the content is right, and the AI side effects in the educational process will have to be addressed. More studies are required in Vietnam to enhance teacher training, ensure that AI-created lessons are taught in line with the national curriculum, and investigate how AI can be used to drive student engagement. Liu (2025) reviewed 19 empirical studies of language teacher AI literacy in collaboration with ChatGPT and finds that teachers need five dimensions of knowledge to work with AI tools effectively and ethically: AI foundations and applications, AI ethics, a human-centred mindset, AI pedagogy, and AI for professional development, but existing training programmes address only the first dimension (technical tool familiarity) and leave the remaining four dimensions (ethical reasoning, pedagogical judgment, and reflective professional identity) largely unaddressed in formal training. This five-dimensional framework provides a practical blueprint for designing professional development programs that build not just technical tool competence but the critical, human-centered, pedagogical, and professional development competencies teachers need to use AI responsibly. By taking the initiative in examining these concerns, the education sector in Vietnam can make sure that the integration of AI in education is optimized on the one hand to deliver all the benefits of AI, including efficiency of teaching and the creation of more engaging learning opportunities, and mitigated on the other hand to avoid all the possible pitfalls, such as loss of teacher authority, cultural inapplicability, and unreliability of pieces of content. Such a well-considered and research-informed practice will assist in establishing the educational context in which the AI role in education is carefully designed as a valuable addition to rather than a replacement of teacher expertise and professional judgment.

Conclusion

This systematic review of 34 studies yields three principal findings, tempered by the methodological limitations noted above. First, AI tools – particularly LLM-based assistants such as ChatGPT – can enhance the efficiency and variety of lesson planning when teachers possess sufficient pedagogical and technological knowledge to evaluate and adapt AI-generated outputs critically (Lee & Zhai, 2024; Karaman & Göksu, 2024; Trang & Thu, 2024). The reported gains in planning efficiency and material diversity align with DOI's relative advantage, while the TPACK framework explains why these gains depend on robust TK–PK integration rather than on tools alone. Second, persistent barriers to effective AI integration cluster around content accuracy and cultural bias in AI outputs, teacher AI-literacy deficits, and uneven infrastructure – challenges that are particularly acute in the Vietnamese educational context (Diep & Dang, 2025; Moundridou et al., 2024; Nguyen et al., 2022). These barriers map to DOI's compatibility and complexity constraints, highlighting the need for context-sensitive adoption strategies. Third, the current evidence base is predominantly short-term, small-scale, and reliant on self-reported perceptions, with limited classroom-observational or longitudinal data.

Based on these findings, the following recommendations are offered. First, professional development programs should couple proficiency with AI tools, including prompt engineering skills, with explicit training in critical evaluation and syllabus alignment, enabling teachers to function as active designers rather than passive consumers of AI-generated content (Durmus, 2024; Diep & Dang, 2025). Second, infrastructure investment should prioritize equitable access for rural and under-resourced campuses, addressing the digital divide that currently limits AI adoption to predominantly urban settings (Williyan et al., 2024; Nguyen et al., 2022). Third, Vietnam-aligned data governance guidance is needed, referencing current MOET directions and national data protection norms, to provide educators with clear ethical and legal frameworks for classroom AI use.

Future research should prioritize longitudinal, classroom-observational designs across multiple subjects and educational levels, with attention to both teacher practice and student learning outcomes, to build a more robust evidence base for AI-assisted lesson planning in the Vietnamese context. Additionally, studies examining how AI tools can be more linguistically, curricularly, and culturally aligned to the requirements of Vietnamese classrooms would strengthen contextual relevance and alignment with national education objectives.

Taken together, these findings underscore that AI, when applied carefully and with attention to context, can serve as a valuable augment to – but in no condition a substitute for – teacher knowledge and professional judgment.

References

Aydın Yıldız, T. (2024). Exploring the Impact of ChatGPT on Improving 21st-Century Skills for Future English Teachers During Lesson Planning. *Computers in the Schools*, 1–24. https://doi.org/10.1080/07380569.2024.2429534

Baytak, A. (2024). The Content Analysis of the Lesson Plans Created by ChatGPT and Google Gemini. *Research in Social Sciences and Technology*, *9*(1), 329–350. https://doi.org/10.46303/ressat.2024.19

Belloula, S. (2025). Empowering educators: Leveraging AI to revolutionize lesson planning. *International Journal of Research in Education and Science (IJRES)*, *11*(2), 264-280. https://doi.org/10.46328/ijres.1295

Clark, C. H., & Van Kessel, C. (2024). "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*, *24*(2), 151–183. https://doi.org/10.70725/612324oqaync

Cong-Lem, N., Tran, T. N., & Nguyen, T. T. (2024). Academic integrity in the age of generative AI: Perceptions and responses of Vietnamese EFL teachers. *Teaching English with Technology*, *24*(1), 28–47. https://doi.org/10.56297/FSYB3031/MXNB7567

Corp, A., & Revelle, C. (2023). ChatGPT is here to stay: Using ChatGPT with student teachers for lesson planning. *The Texas Forum of Teacher Education*, *14*, 116–124. https://www.learntechlib.org/p/222408/

Diep, M. C., & Dang, M. H. (2025). Exploring factors influencing the adoption of ChatGPT as a supportive tool in EFL lesson creation. *AsiaCALL Online Journal, 16*(1), 111–139. *https://doi.org/10.54855/acoj.251616*

Dornburg, A., & Davin, K. J. (2025). ChatGPT in foreign language lesson plan creation: Trends, variability, and historical biases. *ReCALL*, *37*(3), 332–347. https://doi.org/10.1017/S0958344024000272

Du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: a scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, *10*(21), e39630. https://doi.org/10.1016/j.heliyon.2024.e39630

Durmus, C. (2024). From AI-Generated lesson plans to the Real-Life classes: explored by Pre-Service teachers. *8th International Conference on Higher Education Advances (HEAd'22)*. https://doi.org/10.4995/head24.2024.17060

ElSayary, A. (2024). An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era. *Journal of Computer Assisted Learning*, *40*(3), 931–945. https://doi.org/10.1111/jcal.12926

Gao, Y., Wang, Q., & Wang, X. (2024). Exploring EFL university teachers' beliefs in integrating ChatGPT and other large language models in language education: a study in China. *Asia Pacific Journal of Education*, *44*(1), 29–44. https://doi.org/10.1080/02188791.2024.2305173

Hieu, H. H., & Thao, L. T. (2024). Exploring the impact of AI in language education: Vietnamese EFL teachers' views on using ChatGPT for fairy tale retelling tasks. *International Journal of Learning Teaching and Educational Research*, *23*(3),

486–503. https://doi.org/10.26803/ijlter.23.3.24

Hoang, T. T. H., Dang, H. N., Pham, N. B. Q., & Truong, T. K. X. (2025). EFL teachers' perceptions of utilizing ChatGPT in designing lesson plans for IELTS reading skills. *International Journal of AI in Language Education*, *2*(3), 19–39. https://doi.org/10.54855/ijaile.25232

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, *34*(4), 285–291. https://doi.org/10.3233/efi-180221

Jamaluddin, J. E., Abidin, I. Z., Idris, M. A., & Masrom, U. K. (2020). Designing lesson plans for adaptive learning using Moodle LMS platform. *International Journal of Creative Multimedia*, *1*(1), 210–219. https://doi.org/10.33093/ijcm.2020.1.x1.16

Jeon, J., Lee, S. Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Educ Inf Technol* **28**, 15873–15892 (2023). https://doi.org/10.1007/s10639-023-11834-1

Kalenda, P. J., Rath, L., Abugasea Heidt, M., & Wright, A. (2025). Pre-service teacher perceptions of ChatGPT for lesson plan generation. *Journal of Educational Technology Systems*, *53*(3). https://doi.org/10.1177/00472395241301388

Karaman, M. R., & Göksu, İ. (2024). Are lesson plans created by ChatGPT more effective? an experimental study. *International Journal of Technology in Education*, *7*(1), 107–127. https://doi.org/10.46328/ijte.607

Kerr, R. C., & Kim, H. (2025). From Prompts to Plans: A Case Study of Pre-Service EFL Teachers' Use of Generative AI for Lesson Planning. *English Teaching*, *80*(1), 95–118. https://doi.org/10.15858/engtea.80.1.202503.95

Klingelhöfer, S. (2018). Rogers (1962): Diffusion of Innovations. In *Netzwerkforschung* (pp. 489–493). https://doi.org/10.1007/978-3-658-21742-6_115

Korucu-Kış, S. (2024). Zone of proximal creativity: An empirical study on EFL teachers' use of ChatGPT for enhanced practice. *Thinking Skills and Creativity*, *54*, 101639. https://doi.org/10.1016/j.tsc.2024.101639

Kuzu, T. E., Irion, T., & Bay, W. (2025). AI-based task development in teacher education: an empirical study on using ChatGPT to create complex multilingual tasks in the context of primary education. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13673-8

Lee, G., & Zhai, X. (2024). Using ChatGPT for Science Learning: A study on pre-service teachers' lesson planning. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2402.01674

Liu, W. (2025). Language teacher AI literacy: Insights from collaborations with ChatGPT. *Journal of China Computer-Assisted Language Learning*. https://doi.org/10.1515/jccall-2024-0030

Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: a systematic review of empirical studies. *Smart Learning Environments*, *11*(1). https://doi.org/10.1186/s40561-

024-00342-5

Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: a framework for teacher knowledge. *Teachers College Record the Voice of Scholarship in Education*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Moundridou, M., Matzakos, N., & Doukakis, S. (2024). Generative AI tools as educators' assistants: Designing and implementing inquiry-based lesson plans. *Computers and Education Artificial Intelligence*, *7*, 100277. https://doi.org/10.1016/j.caeai.2024.100277

Nguyen, G., Nguyen, N., & Giang, N. T. H. (2022). Situation and proposals for implementing artificial intelligence-based instructional technology in Vietnamese secondary schools. *International Journal of Emerging Technologies in Learning (iJET)*, *17*(18), 53–75. https://doi.org/10.3991/ijet.v17i18.31503

Nguyen, T. T. H. (2023). EFL teachers' perspectives toward the use of ChatGPT in writing classes: A case study at Van Lang University. *International Journal of Language Instruction*, *2*(3), 1–47. https://doi.org/10.54855/ijli.23231

Ning, Y., Zhang, C., Xu, B., Zhou, Y., & Wijaya, T. T. (2024). Teachers' AI-TPACK: Exploring the Relationship between Knowledge Elements. *Sustainability*, *16*(3), 978. https://doi.org/10.3390/su16030978

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Pargmann, J., Berding, F., Rebmann, K. *et al.* How AI feedback supports lesson planning in vocational teacher education: a longitudinal intervention study using an analytical AI platform. *Empirical Res Voc Ed Train* **17**, 26 (2025). https://doi.org/10.1186/s40461-025-00202-7

Pham, T. T., & Le, T. T. (2024). Exploring the impact of artificial intelligence on student creativity in Vietnamese tertiary EFL classrooms: Teacher Perspectives. *Jurnal Komunikasi Pendidikan*, *8*(2), 116–128. https://doi.org/10.32585/jurnalkomdik.v8i2.5052

Shulman, L. (1987). Knowledge and Teaching:Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–23. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.3102/0013189x015002004

Tammets, K., & Ley, T. (2023). Integrating AI tools in teacher professional learning: a conceptual model and illustrative case. *Frontiers in Artificial Intelligence*, *6*. https://doi.org/10.3389/frai.2023.1255089

Thai, L., Trung, T. T., & Dung, T. M. (2025). ChatGPT in Vietnamese math classrooms: What are the influencing factors behind teachers adoption. *Journal of Pedagogical Research*. https://doi.org/10.33902/jpr.202531924

Tran, P. T. H., Huynh, L., Bien, T. A., Dang, B., & Nguyen, A. (2026). 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*, *42*(1), 16–32. https://doi.org/10.1080/21532974.2025.2583516

Tran, P. T. H., Huynh, L., Bien, T. A., Dang, B., & Nguyen, A. (2026). 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*, *42*(1), 16–32. https://doi.org/10.1080/21532974.2025.2583516

Trang, N. T. Q., & Thu, P. T. (2024). The role of AI in improving student learning Outcomes: evidence in Vietnam. International Journal of Multidisciplinary Research and Analysis, 07(06). https://doi.org/10.47191/ijmra/v7-i06-48

Triantafyllou, S. A., & Farhaoui, Y. (2024). Learning Design of Lesson Plans with LAMS and CADMOS. In *Artificial Intelligence, Big Data, IoT and Block Chain in Healthcare: From Concepts to Applications* (pp. 103–111). https://doi.org/10.1007/978-3-031-65018-5_10

Vo, L., & Huynh, N. (2025). Vietnamese EFL teachers' perspectives on ChatGPT: A conceptual metaphor analysis. *Arab World English Journal*, *16*(1), 162–178. https://awej.org/vietnamese-efl-teachers-perspectives-on-chatgpt-a-conceptual-metaphor-analysis/

Wen, W., & Wen, H. (2024). Bridging generative AI technology and teacher education: Understanding preservice teachers' processes of unit design with ChatGPT. *Contemporary Issues in Technology and Teacher Education*, *24*(4). https://citejournal.org/volume-24/issue-4-24/general/bridging-generative-ai-technology-and-teacher-education-understanding-preservice-teachers-processes-of-unit-design-with-chatgpt/

Williyan, A., Fitriati, S. W., Pratama, H., & Sakhiyya, Z. (2024). AI as co-creator: exploring indonesian efl teachers' collaboration with ai in content development. *Teaching English With Technology*, *2024*(2). https://doi.org/10.56297/vaca6841/lrdx3699/rzoh5366

Zhang, P., & Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. *European Journal of Education*, *59*(2). https://doi.org/10.1111/ejed.12599

Appendix A

MMAT Item-Level Quality Ratings for Included Studies (n = 35)

Methodological quality was assessed using the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018). Five criteria were evaluated for each study:

Q1 — Clear and focused research question

Q2 — Appropriate study design for the research question

Q3 — Adequate sampling (size, diversity, and selection rationale)

Q4 — Appropriate and described data collection procedures

Q5 — Valid analysis and transparent reporting of findings

Ratings: Y = criterion met N = criterion not met ? = cannot determine

Overall quality bands: High = 4-5 criteria met | Moderate = 3 criteria met | Low = 2 or fewer criteria met

Only studies rated Moderate or High were retained in the final synthesis. Studies rated Low were used selectively and contextualised against higher-quality findings in the narrative.

Studies 16-34 (shaded green) are newly added. NR = not reported by study authors.

#	Study (author, year)	Research design	Q1	Q2	Q3	Q4	Q5	Overall quality
1	**Lee & Zhai (2024)**	Qualitative + TPACK rubric	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
2	**Karaman & Goksu (2024)**	Quasi-experimental	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
3	**Trang & Thu (2024)**	Experimental	**Y**	**Y**	**N**	**Y**	**Y**	**Moderate**
4	**Moundridou et al. (2024)**	Comparative analysis	**Y**	**?**	**?**	**Y**	**?**	**Moderate**

#	Study (author, year)	Research design	Q1	Q2	Q3	Q4	Q5	Overall quality
5	**Du Plooy et al. (2024)**	Systematic review	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
6	**Williyan et al. (2024)**	Qualitative interviews	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
7	**Baytak (2024)**	Content analysis	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
8	**Diep & Dang (2025)**	Mixed-method (PLS-SEM)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
9	**Triantafyllou & Farhaoui (2024)**	Conceptual + case study	**Y**	**?**	**?**	**?**	**Y**	**Low**
10	**Jamaluddin et al. (2020)**	Design-based research	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
11	**Durmus (2024)**	Qualitative (n = 4)	**Y**	**Y**	**N**	**Y**	**Y**	**Low**
12	**Pham & Le (2024)**	Qualitative interviews	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
13	**Hieu & Thao (2024)**	Qualitative interviews (n = 9)	**Y**	**Y**	**N**	**Y**	**Y**	**Moderate**
14	**Nguyen et al. (2022)**	Survey (n = 119)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
15	**Aydın Yıldız (2024)**	Qualitative case study (n=2)	**Y**	**Y**	**?**	**Y**	**N**	**Low**
16	**Tammets & Ley (2023)**	Conceptual + illustrative case	**Y**	**?**	**?**	**?**	**Y**	**Low**

#	Study (author, year)	Research design	Q1	Q2	Q3	Q4	Q5	Overall quality
17	**Lo et al. (2024)**	Systematic review (n = 70 studies)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
18	**Jeon & Lee (2023)**	Qualitative (n = 11)	**Y**	**Y**	**N**	**Y**	**Y**	**Moderate**
19	**Dornburg & Davin (2025)**	Content analysis (50 plans)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
20	**Clark & Van Kessel (2024)**	Content analysis	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
21	**Kerr & Kim (2025)**	Case study	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
22	**Kuzu et al. (2025)**	Qualitative (n = 17)	**Y**	**Y**	**N**	**Y**	**Y**	**Low**
23	**Zhang & Tur (2024)**	Systematic review (PRISMA)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
24	**Belloula (2025)**	Mixed-method	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
25	**Kalenda et al. (2025)**	Pre-post survey (n = 59)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
26	**Corp & Revelle (2023)**	Case study	**Y**	**?**	**N**	**Y**	**Y**	**Low**
27	**Wen & Wen (2024)**	Qualitative exploratory	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
28	**ElSayary (2024)**	Survey	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**

#	Study (author, year)	Research design	Q1	Q2	Q3	Q4	Q5	Overall quality
29	**Korucu-Kis (2024)**	Qualitative (n = 29)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
30	**Gao, Wang & Wang (2024)**	Survey	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
31	**Pargmann et al. (2025)**	Longitudinal experimental (n = 103)	**Y**	**Y**	**Y**	**Y**	**Y**	**High**
32	**Nguyen, T. T. H. (2023)**	Mixed-method (survey + interviews)	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
33	**Vo & Huynh (2025)**	Qualitative (n = 4)	**Y**	**Y**	**N**	**Y**	**Y**	**Low**
34	**Cong-Lem, Tran & Nguyen (2024)**	Survey + mixed-methods	**Y**	**Y**	**?**	**Y**	**Y**	**Moderate**
35	**Hoang et al. (2025)**	Survey (quantitative, n = 54)	**Y**	**Y**	**Y**	**Y**	**Y**	**Moderate**

Note. *Colour coding in Overall column: green = High; yellow = Moderate; red = Low. Individual Q cells: green = Y (met); red = N (not met); white = ? (cannot determine). Studies 9, 11, 16, 22, 26, and 33 were rated Low but retained because their evidence was used selectively within a broader synthesis supported by higher-quality studies.*

Appendix B

Full-Text Articles Excluded After Eligibility Assessment (n = 48)

Following title and abstract screening, 82 full-text articles were retrieved and assessed for eligibility. Of these, 48 were excluded. The table below lists each excluded article and the specific reason for exclusion, organised by exclusion category. Exclusion reasons correspond to the predefined criteria set out in the Methodology section.

Exclusion categories applied: (1) No lesson planning component - study focused on AI for student learning outcomes, homework completion, academic integrity, assessment, grading, tutoring systems, writing skills, technology acceptance, or administrative functions without a teacher lesson planning component; (2) Non-peer-reviewed or opinion publication - preprint, editorial, policy framework, or conceptual article without empirical data; (3) Outside date range - published before January 2020

No.	Excluded study (APA 7 reference)	Reason for exclusion
1	*Sat, M. The impact of AI integration in project preparation in education course on pre-service teachers' innovativeness, AI anxiety, attitudes, and acceptance. BMC Psychol **13**, 1297 (2025). https://doi.org/10.1186/s40359-025-03647-3*	No lesson planning component; psychological outcomes in pre-service teachers
2	*Wafa Naif Alwakid, Nisar Ahmed Dahri, Humayun, M., & Ghadah Naif Alwakid. (2025). Exploring the Role of AI and Teacher Competencies on Instructional Planning and Student Performance in an Outcome-Based Education System. Systems, 13(7), 517–517. https://doi.org/10.3390/systems13070517*	No lesson planning component; the relationship between AI use and student outcomes
3	*Tahiru, F. (2021). AI in Education. Journal of Cases on Information Technology, 23(1), 1–20. https://doi.org/10.4018/jcit.2021010101*	No lesson planning component; applications of AI across education systems (learning, administration, assessment)
4	*Létourneau, A., Martineau, D., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring*	No lesson planning component; learning outcomes and

No.	Excluded study (APA 7 reference)	Reason for exclusion
	systems (ITS) in K-12 education. Npj Science of Learning, 10(1), 1–13. https://doi.org/10.1038/s41539-025-00320-7	system effectiveness
5	*Lin, C.-C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. Smart Learning Environments, 10(1). https://doi.org/10.1186/s40561-023-00260-y*	No lesson planning component; AI tutoring systems: design and effectiveness; student learning focus
6	*Latif, E., Liu, V., & Zhai, X. (2026). A systematic review of intelligent and robot tutoring systems: evolution, pedagogical design, and AI-driven classification. Smart Learning Environments, 13(1). https://doi.org/10.1186/s40561-025-00427-9*	No lesson planning component; AI systems delivering instruction (robots, tutors)
7	*Yang, J., & Zhang, B. (2019). Artificial Intelligence in Intelligent Tutoring Robots: A Systematic Review and Design Guidelines. Applied Sciences, 9(10), 2078. https://doi.org/10.3390/app9102078*	Outside 2020–2025 range; no lesson planning component; robot tutors for learner interaction
8	*Deng, C., & Yuan, B. (2026). Research on an intelligent tutoring system based on automatic construction of multimodal knowledge graphs and retrieval-augmented generation. Frontiers in Computer Science, 8. https://doi.org/10.3389/fcomp.2026.1777749*	No lesson planning component; AI system architecture/development
9	*Guo, L., Wang, D., Gu, F. et al. Evolution and trends in intelligent tutoring systems research: a multidisciplinary and scientometric view. Asia Pacific Educ. Rev.* ***22****, 441–461 (2021). https://doi.org/10.1007/s12564-021-09697-7*	No lesson planning component; Intelligent tutoring systems (ITSs)

No.	Excluded study (APA 7 reference)	Reason for exclusion
10	*Nguyen, T. N. L. (2026). Artificial intelligence and the transformation of educational research: From assistance to academic co-creation. Dong Thap University Journal of Science, 15(1), 1-11. https://doi.org/10.52714/dthu.15.1.2026.1736*	No lesson planning component; AI for academic research (searching, summarising, writing)
11	*Nguyen, H. D., Pham, N. T., & Dinh, T. H. (2025). Application of artificial intelligence in foreign language teaching and learning: A comparative study. Ho Chi Minh City University of Education Journal of Science, 22(11), 2060–2072. https://doi.org/10.54607/hcmue.js.22.11.4752(2025)*	No lesson planning component; teachers' and students' perceptions and AI usage
12	*Pham, T. N., & Dang, T. X. (2025). An Investigation into the Application of Artificial Intelligence for Language Teaching and Learning in Vietnam. Journal of Contemporary Educational Policies and Practices, 9(2), 265–283. https://doi.org/10.52296/vje.2025.557*	No lesson planning component; survey and regression analysis of AI usage; impact of AI tools and perceptions
13	*Nhu, T., Nam Van Lai, & Quyet Thi Nguyen. (2024). Artificial Intelligence (AI) in Education: A Case Study on ChatGPT's Influence on Student Learning Behaviors. Educational Process: International Journal, 13(2). https://doi.org/10.22521/edupij.2024.132.7*	No lesson planning component; student behaviour and learning habits; engagement and usage patterns with AI
14	*Hong, T. T. M., Tung, N. T. T., & Thinh, H. T. (2026). Discovering acceptance and intention to use artificial intelligence for learning among pre-service teachers in Vietnam. Discover Education. https://doi.org/10.1007/s44217-026-01289-0*	No lesson planning component; technology acceptance (TAM model); intentions and attitudes toward AI
15	*Cung, L. T., Hoang, U. P. T., Dinh, A. N., & Bui, T. H. (2025). University Students' Perceptions of AI Application in Writing Skills in Vietnam: A*	No lesson planning component; student writing and AI tools; perceptions and skill development

No.	Excluded study (APA 7 reference)	Reason for exclusion
	Systematic Review. Vietnam Journal of Education, 325–334. https://doi.org/10.52296/vje.2025.533	
16	*Long, T. T., Trang, P. T., Bien, T. N., & Cuong, T. Q. (2026). The Influence of Artificial Intelligence on Vietnam's Higher Education in the Contemporary Era. VNU Journal of Science: Education Research. https://doi.org/10.25073/2588-1159/vnuer.5276*	No lesson planning component; system-level impact of AI in higher education; policy, transformation, and trends
17	*Alsharefeen, R., & Al Sayari, N. (2025). Examining academic integrity policy and practice in the era of AI: a case study of faculty perspectives. Frontiers in Education, 10. https://doi.org/10.3389/feduc.2025.1621743*	No lesson planning component; institutional policy and teacher responses to AI cheating; plagiarism detection and misconduct cases
18	*Avello, D., & Zurita, S. A. (2025). Exploring the nexus of academic integrity and artificial intelligence in higher education: a bibliometric analysis. International Journal for Educational Integrity, 21(1). https://doi.org/10.1007/s40979-025-00199-2*	No lesson planning component; AI and academic misconduct research trends; key themes (dishonesty, psychology, misuse)
19	*Leong, W. Y., & Zhang, J. B. (2025). AI on Academic Integrity and Plagiarism Detection. ASM Science Journal, 20(1), 2025. https://doi.org/10.32802/asmscj.2025.1918*	No lesson planning component; AI tools for detecting plagiarism and AI-generated text; comparison of traditional and AI detection systems
20	*Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20(1), 1–25. https://doi.org/10.1186/s41239-023-00408-3*	Non-peer-reviewed policy framework; no empirical study design or lesson planning data

No.	Excluded study (APA 7 reference)	Reason for exclusion
21	*Ateeq, A., Alzoraiki, M., Milhem, M., & Ateeq, R. A. (2024). Artificial intelligence in education: implications for academic integrity and the shift toward holistic assessment. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1470979*	No lesson planning component; academic dishonesty and assessment integrity
22	*Cheng, T.-S., Lu, Y.-C., Chang, C.-C., & Yang, C.-S. (2015). Combining Mobile Devices with NFC Technology in a Test Assessment System. 2015 IEEE 15th International Conference on Advanced Learning Technologies, 241–243. https://doi.org/10.1109/icalt.2015.54*	Outside 2020–2025 range; no lesson planning component; ML models for student grade prediction; performance analytics and prediction systems
23	*Mehta, A., Morris, N. P., Swinnerton, B., & Homer, M. (2019). The Influence of Values on E-learning Adoption. Computers & Education, 141, 103617. https://doi.org/10.1016/j.compedu.2019.103617*	Outside 2020–2025 range; no lesson planning component; models predicting student success/failure; evaluation and prediction focus
24	*Zhai, X. (2022). ChatGPT User Experience: Implications for Education. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4312418*	No lesson planning component; ChatGPT implications for education; learning goals and assessment formats
25	*Yong Zhao (赵勇. (2024). Artificial Intelligence and Education: End the Grammar of Schooling. ECNU Review of Education. https://doi.org/10.1177/20965311241265124*	No lesson planning component; theoretical/conceptual article; AI and transformation of schooling systems (grammar of schooling)
26	*H, M. (2024). Role of AI chatbots in education. Journal of Medical Education and Development. https://doi.org/10.18502/jmed.v19i3.17348*	No lesson planning component; explores AI chatbots for student support
27	*Williamson, B., Macgilchrist, F., & Potter, J. (2023). Re-examining AI,*	No lesson planning component; critical, non-empirical

No.	Excluded study (APA 7 reference)	Reason for exclusion
	automation and datafication in education. Learning, Media and Technology, 48(1), 1–5. https://doi.org/10.1080/17439884.2023.2167830	discussion of AI/data practices
28	*Nguyen, T. H., & Ha, D. A. T. (2025). Exploring Vietnamese Students' Intention to Adopt AI-Powered Study Tools: Integrating TPB and TAM. Educational Process: International Journal, 16, e2025283. https://doi.org/10.22521/edupij.2025.16.283*	No lesson planning component; students' intention to use AI-powered tools and the factors influencing that intention.
29	*Luo, J., Zheng, C., Yin, J. et al. Design and assessment of AI-based learning tools in higher education: a systematic review. Int J Educ Technol High Educ **22**, 42 (2025). https://doi.org/10.1186/s41239-025-00540-2*	No lesson planning component; AI tools in learning and instructional support (assessment, feedback, tutoring)
30	*Fiialka, S., Kornieva, Z., & Honcharuk, T. (2024). The use of ChatGPT in creative writing assistance. XLinguae, 17(1), 3–19. https://doi.org/10.18355/xl.2024.17.01.01*	No lesson planning component; ChatGPT used as a *generative tool* in writing contexts
31	*Truc, T. (2024). Investigating the effectiveness of practicing English conversation with ChatGPT in improving non-English majored students' English speaking skills at Nguyen Tat Thanh University. Journal of Science and Technology, 7(4). https://doi.org/10.55401/b1rnww21*	No lesson planning component; ChatGPT's role in language learning outcomes
32	*Lim, T., Swapna Gottipati, & Cheong, M. (2025). What students really think: unpacking AI ethics in educational assessments through a triadic framework. International Journal of Educational Technology in Higher Education, 22(1). https://doi.org/10.1186/s41239-025-00556-8*	No lesson planning component; students' perception of ethical aspects of AI in assessments.
33	*Chaudhry, I. S., Sarwary, A. M., Ghaleb, & Chabchoub, H. (2023). Time*	No research evidence on AI for lesson planning; Broader

No.	Excluded study (APA 7 reference)	Reason for exclusion
	to Revisit Existing Student's Performance Evaluation Approach in Higher Education Sector in a New Era of ChatGPT — A Case Study. Cogent Education, 10(1). https://doi.org/10.1080/2331186x.2023.2210461	impacts of AI like ChatGPT on teaching and learning environments
34	*Gao, R., Merzdorf, H. E., Anwar, S., M. Cynthia Hipwell, & Srinivasa, A. (2024). Automatic assessment of text-based responses in post-secondary education: A systematic review. Computers & Education: Artificial Intelligence, 6, 100206–100206. https://doi.org/10.1016/j.caeai.2024.100206*	No lesson planning component; AI automated grading and feedback
35	*Iskender, A. (2023). Holy or Unholy? Interview with Open AI's ChatGPT. European Journal of Tourism Research, 34, 3414. https://doi.org/10.54055/ejtr.v34i.3169*	Tourism and hospitality education; not related to lesson planning or teacher education
36	*Student modelling in adaptive e-learning systems. (2011). Knowledge Management & E-Learning: An International Journal, 342–355. https://doi.org/10.34105/j.kmel.2011.03.025*	Outside date range (2011); no lesson planning component; student modelling in adaptive e-learning systems
37	*Bettayeb, A. M., Talib, M. A., Zahraa, A., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: conversational AI in education. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1379796*	No lesson planning component; general impact of ChatGPT on teaching and learning without instructional design focus
38	*Torres-Peralta, R., Cirett-Galán, F., Heras-Sanchez, M. del C., Lerma-Molina, K., & Espinoza-Flores, I. (2026). Perceived Learning vs. Engagement in AI-Assisted Homework: A Comparative Study of ChatGPT Use Across High School, University, and Teachers in Sonora, Mexico (2024–2025). Future Internet, 18(3), 122.*	Not related to lesson planning — focused on students' perceived learning and engagement when using ChatGPT for homework tasks

No.	Excluded study (APA 7 reference)	Reason for exclusion
	https://doi.org/10.3390/fi18030122	
39	*Jayaram, Y., & Sundar, D. (2023). AI-Powered Student Success Ecosystems: Integrating ECM, DXP, and Predictive Analytics. International Journal of Artificial Intelligence, Data Science, and Machine Learning, 4, 109–119. https://doi.org/10.63282/3050-9262.ijaidsml-v4i1p113*	No lesson planning component; AI for student success prediction using analytics and digital experience platforms
40	*Thien, T., To, T., Xuan, T., & Nguyen, D. L. (2025). A Systematic Review of How ChatGPT is Perceived and Utilized in EFL Writing Classes: Implications for Vietnam. AsiaCALL Online Journal, 16(1), 295–311. https://doi.org/10.54855/acoj.2516115*	No lesson planning component; ChatGPT use in EFL writing classes from student perspective
41	*Nguyen, X.-H., Nguyen, H.-A., Cao, L., & hana, trương. (2023). Unleashing the Potential and Recognizing the Limitations of ChatGPT in Vietnamese Geography Education. https://doi.org/10.35542/osf.io/r4dg6*	Non-peer-reviewed preprint (OSF repository); no peer review; ChatGPT in Vietnamese geography education
42	*Dwivedi, Y. K. (2023). "So What If ChatGPT Wrote it?" Multidisciplinary Perspectives on opportunities, Challenges and Implications of Generative Conversational AI for research, Practice and Policy. International Journal of Information Management, 71(0268-4012), 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642*	No lesson planning component; multidisciplinary perspectives on ChatGPT for research and policy without instructional design focus
43	*Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. Education Sciences, 13(4), 1–15. https://doi.org/10.3390/educsci13040410*	No lesson planning component; rapid review of ChatGPT's general educational impact without lesson planning focus

No.	Excluded study (APA 7 reference)	Reason for exclusion
44	*Li, J., Yan, Y., & Zeng, X. (2025). Exploring Artificial Intelligence in Inclusive Education: A Systematic Review of Empirical Studies. Applied Sciences, 15(23), 12624–12624. https://doi.org/10.3390/app152312624*	No lesson planning component; focused on AI in inclusive education for students with disabilities
45	*Bond, M., Khosravi, H., De Laat, M. et al. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. Int J Educ Technol High Educ **21**, 4 (2024). https://doi.org/10.1186/s41239-023-00436-z*	No lesson planning component; meta-systematic review of AI ethics and collaboration in higher education
46	*Shahzad, M.F., Xu, S. & Javed, I. ChatGPT awareness, acceptance, and adoption in higher education: the role of trust as a cornerstone. Int J Educ Technol High Educ **21**, 46 (2024). https://doi.org/10.1186/s41239-024-00478-x*	No lesson planning component; ChatGPT awareness and acceptance among students in higher education
47	*Yang, Y., Zhang, Y., Sun, D. et al. Navigating the landscape of AI literacy education: insights from a decade of research (2014–2024). Humanit Soc Sci Commun **12**, 374 (2025). https://doi.org/10.1057/s41599-025-04583-8*	No lesson planning component; systematic review of AI literacy education as a curriculum topic, not teacher planning
48	*Fu, Y., Weng, Z. & Wang, J. Examining AI Use in Educational Contexts: A Scoping Meta-Review and Bibliometric Analysis. Int J Artif Intell Educ **35**, 1388–1444 (2025). https://doi.org/10.1007/s40593-024-00442-w*	Not related to lesson planning; scoping meta-review and bibliometric analysis of AI use in education broadly

Note. *Where a conference version and a journal version of the same study existed, the peer-reviewed journal version was retained and the conference or preprint version was excluded as a duplicate. References are provided in APA 7 format for traceability.*

Biodata

Nguyen Ngoc Hai is the academic manager at Anne Hill Bilingual School in Vietnam. Having six years of teaching experience, he is now specializing in the Cambridge curriculum. He is interested in key areas of English education, incorporating technology like AI, and modern teaching methods.

Ho Hanh Hao has been teaching English for six years. She is currently teaching English Language Arts at Scotch Australian Grammar School (Scotch AGS). Her research interests lie in the application of different teaching methods and technology in English language teaching to boost her students' engagement and learning outcomes.