

From prompts to pedagogy: Mapping prompt engineering in English language education

Pelin Derinalp * 

Gaziantep University, Gaziantep, Türkiye

Suggested citation: Derinalp, P. (2026). From prompts to pedagogy: Mapping prompt engineering in English language education. *Journal of Educational Technology and Online Learning*, 9(3), 552-567.

Highlights	Abstract
• Maps EFL-focused prompt engineering research (2022-2025) • Identifies benefits, limitations, and ethical tensions in practice • Reveals fragmented methods and lack of systematic frameworks • Provides implications for teacher education and future research	With the widespread use of generative AI tools in educational settings, prompt engineering has become increasingly important for obtaining effective and pedagogically meaningful outputs from these tools. However, research addressing prompt engineering in English as a Foreign Language (EFL) pedagogy is scattered across different contexts, skill domains, and methodological approaches, highlighting the need for a synthesis of emerging evidence. Therefore, this study aims to map research conducted between 2022 and 2025 through a scoping review. Studies focusing on prompt design in the EFL context were analysed following PRISMA-ScR guidelines. The findings show that most research was conducted in Asia, with university students as participants, and qualitative/mixed-method designs were prevalent. Prompt engineering was primarily used for writing skills and research competence, promoting learner autonomy but raising concerns about cognitive load, limited teacher guidance, and academic integrity. Three key pedagogical tensions emerged: the balance between theory and practice, learner autonomy vs. cognitive load, and technical output vs. ethical filtering. The study highlights the field's unestablished nature, emphasising the need for structured frameworks, ethics-focused research, and reliable measurement tools for assessing prompt literacy and ethical competence. It offers a roadmap for researchers, teacher educators, and institutions to address these gaps.
Article Info: Review Article	
Keywords: <i>EFL, GenAI, prompt engineering, scoping review</i>	

1. Introduction

Over recent years, breakthroughs in artificial intelligence (AI) have profoundly reshaped educational research. In particular, the rise of large language models (LLMs) like OpenAI's ChatGPT has introduced novel opportunities in teaching and learning (Derinalp, 2024; Knoth et al., 2024). As generative AI tools become increasingly accessible, they are being applied across diverse domains (Alam, 2021; Dwivedi et al., 2023), including English language education, thanks to their ability to generate human-like responses. These tools enable adaptive learning environments by tailoring instruction to individual learners' needs, delivering timely and personalised feedback, enhancing student motivation, and transforming learning experiences into more flexible and interactive ones compared to traditional teaching methods (Derinalp & Halife, 2025; Gligorea et al., 2023; Owoc et al., 2021). However, the impact of generative AI (GenAI) tools in education largely depends on how users interact with them. LLM-based systems generate output based on user input, in other words, prompts. Therefore, the accuracy, contextual appropriateness, and pedagogical value of the resulting product are directly related to the quality of the prompt used (Knoth et

* Corresponding author. ELT Department, Gaziantep University, Türkiye.
e-mail address: pelinderinalp@gmail.com / pelinderinalp@gaziantep.edu.tr

Doi: <http://doi.org/10.31681/jetol.1985220>

Received 2 Jul 2026; Accepted 27 Sep 2026

ISSN: 2618-6586. This is an open Access article under the CC BY license.



al., 2024; White et al., 2023). This has led to the emergence of a new area of expertise called prompt engineering. Challenges encountered in the use of LLMs, particularly erroneous outputs, context-disconnected responses, and limitations in perception of reality, have elevated prompt engineering to a critical skill that guides human-AI interaction and aims to achieve high-quality outputs (Bommasani et al., 2021; Floridi & Chiriatti, 2020; Liu et al., 2023). However, empirical research on how this process works, especially for non-expert users, appears to be limited.

A review of the current literature reveals that while research on the use of GenAI tools in the English as a Foreign Language (EFL) context is rapidly increasing, the majority of these studies focus on the general effects of the tools such as contributions to writing performance (Nawi et al., 2024; Tiandem-Adamou, 2024) student motivation (Renfeng et al., 2025), and learner-related factors such as AI self-efficacy and anxiety (Taşçı, 2026), or attitudes (Derinalp, 2026; Derinalp & Halife, 2025). Valuable as such findings are in the early stages of the field, they tend to limit the discussion to the level of the tools' effectiveness. However, an important factor influencing learning with these models is how the prompts at the centre of the interaction are designed. The formal structure of the prompt, its content strategies, and its pedagogical/ethical orientation can influence the accuracy, contextual appropriateness, and educational value of the output obtained. As such, recent studies have begun to highlight this aspect. For example, Noviyanti et al. (2026) examined how different prompt structures affect the quality of AI-generated feedback in EFL discussion writing. Their findings indicate that prompt design components such as the role assigned to the AI and the order of instructions can shape the pedagogical and linguistic quality of the resulting output. Similarly, Andewi et al. (2025) examined prompting strategies in the context of English language learning and found that students' use of prompts evolved from simple questions toward more complex, multi-turn interactions, emphasising that prompting is a learned competency requiring pedagogical guidance. When evaluated collectively, these studies demonstrate that prompt engineering and prompting practices have significant pedagogical implications for EFL. Despite this pedagogical significance, existing research on prompt engineering in the field of EFL appears fragmented across different contexts, objectives, and methodological approaches. This scattered nature of the field poses a challenge in obtaining an overview of its current body of knowledge. Therefore, a scoping review is warranted to systematically map how prompt engineering has been integrated into EFL pedagogy, the methodological approaches adopted, and the pedagogical and ethical tensions emerging in this area. Given that this is an emerging, rapidly evolving, and heterogeneous body of literature, a scoping review can provide an overview of the current state of the field and identify areas for future research.

2. Literature

The ability of LLMs to generate high-quality, persuasive text in natural language has fundamentally transformed the educational research agenda, providing learners with low-threshold access through interactive, multi-turn conversational interfaces (Bommasani et al., 2021). Thanks to accessible chat interfaces, LLM-based tools have rapidly proliferated in the educational field, and this spread has quickly permeated many contexts (Dwivedi et al., 2023), including EFL field (Dai & Panghe, 2025; Liu et al., 2024; Wang, 2025). Recent studies have shown that LLMs enable learning designs that enhance personalised support, timely feedback, and self-regulation through customisable flows and multi-loop interactions (Knoth et al., 2024). Nevertheless, LLM output quality exhibits high sensitivity to user input, namely the prompts. Role assignment, clear definition of objectives/context/constraints, and clarification of the output format are directly related to the accuracy, contextual appropriateness, and pedagogical value of the generated text (Knoth et al., 2024). Therefore, prompt engineering goes beyond a purely technical fine-tuning task and evolves into a pedagogical form of interaction that involves processes such as creating guidelines tailored to learning objectives, iterative validation, evidence checking, and ethical filtering. Applied studies conducted in the context of writing instruction also show that structured prompt applications can render learners' interactions more systematic and purposeful (Woo et al., 2024).

The integration of technology into language teaching has long been addressed through various theoretical frameworks. The Technological Pedagogical Content Knowledge (TPACK) framework (Mishra & Koehler, 2006) focuses on teachers integrating their technological, pedagogical, and subject knowledge,

while Computer-Assisted Language Learning (CALL) research conceptualises the integration of technology into language learning processes (Chapelle, 2005; Levy, 1997). Similarly, Technology Acceptance Models (TAM) are frequently used frameworks for understanding users' attitudes and acceptance of technology-based tools (Venkatesh et al., 2003). However, despite this comprehensive variety of models, systematic frameworks explaining the EFL context specifically for prompt engineering are extremely limited. Only a few notable initiatives stand out in the literature. The CLEAR (Concise, Logical, Explicit, Adaptive, and Reflective) Framework developed by Lo (2023a, 2023b) proposes designing prompts in a more structured manner in terms of context, learner needs, and outputs. Additionally, Zheng & Stewart (2024) conceptualised prompt design through two dimensions, navigation and generation, that both guide output and pedagogically functionalise production. Beyond these initiatives, however, it is difficult to identify established theoretical approaches that holistically address the pedagogical or ethical dimensions of prompt engineering in the context of language teaching.

On the other hand, while highlighting the opportunities offered by technology, such as personalisation, adaptive feedback, and increased participation, it is also important to emphasise the ethical and socio-technical challenges. The outputs of LLMs, when used without contextualisation and co-supervision mechanisms pose risks in terms of academic integrity (Nelson et al., 2025), learner autonomy (Arif & Naeem, 2025), and critical thinking (Andewi et al., 2025). In human-AI collaboration, responsibility becomes a shared quality alongside the effectiveness of oversight processes (Cañas, 2022). This picture necessitates approaching prompt engineering in EFL not only as a technical skill, but as a multi-layered competency area intertwined with pedagogical design and ethical filtering competencies. In this regard, taking a step back and looking at the bigger picture, it becomes clear that the determinant of success with LLMs in EFL is not access to the tool, but rather the didactic design that governs the tool, that is, what and how prompting is conducted, and within what ethical framework. The current literature suggests that publications on GenAI use in EFL are rapidly increasing; however, most studies focus on the general effects of LLMs such as research competency and writing performance (Abdelhalim, 2024; Alkhateeb et al., 2025), while the prompt itself is mostly treated as a secondary element. This leads to fragmented and incomparable knowledge about how prompt engineering is applied in EFL pedagogy, which strategies are used, and what pedagogical/ethical consequences result.

This study aims to present an overview of the field by compiling research conducted on prompt engineering in the EFL context using the scoping review method. In this regard, the countries where the studies were conducted, participant profiles, and methodological trends; how prompt engineering is applied and pedagogically positioned; which skill areas it is integrated into; and the reported benefits and limitations of this integration are systematically examined. Furthermore, prominent themes in the literature are synthesised, and the research gaps indicated by these themes are discussed. The findings of this study are expected to provide conceptual clarity for researchers and serve as a guiding reference for teacher educators and practitioners in developing future teaching frameworks.

3. Methodology

This research was conducted within the framework of a scoping review approach. A scoping review is a more exploratory and broad-scale type of review compared to systematic reviews, aiming to reveal the scope, directions, and research gaps in the existing literature in a specific field (Arksey & O'Malley, 2005; Tricco et al., 2018). Although prompt engineering has rapidly gained prominence in education in general, especially after November 2022, the field has not yet matured to a level that would allow for a systematic review in English as a Foreign Language Teaching. Therefore, a scoping review was adopted as a more appropriate method in the present study.

In line with this, the PRISMA-ScR guidelines (Tricco et al., 2018) were followed in the research, and the five-step process proposed by Arksey and O'Malley (2005) was applied: (1) defining the research questions, (2) searching for relevant studies, (3) selecting studies based on inclusion and exclusion criteria, (4) extracting and organising data, (5) synthesising and reporting findings. Accordingly, in the first stage, the research questions of the current study are articulated as follows:

RQ1) What is the general overview of research on prompt engineering in the EFL context in terms of country, participant profiles, and methodological trends?

RQ2) In these studies, how is prompt engineering applied and pedagogically positioned?

RQ3) In which areas is prompt engineering integrated into EFL pedagogy, and what are the reported benefits and limitations of this integration?

RQ4) What are the overarching themes emerging in the literature, and what research gaps do these themes indicate?

In the second stage, a comprehensive search was conducted in the Web of Science (WOS), Scopus, ERIC, and EBSCO databases to identify relevant studies. Since the use of GenAI tools, particularly ChatGPT, in education has been on the agenda since late 2022, the literature search covered studies published from late 2022 through August 2025. The keywords used in the literature review were structured to reflect both conceptual diversity and contextual scope:

("prompt engineering" OR "prompt design" OR "prompt construction" OR "prompting strategies") AND ("EFL" OR "English language learning" OR "English language teaching" OR "ESL" OR "English as a foreign language" OR "TESOL")

This combination aimed to directly relate prompt engineering and derivative concepts to the EFL/ESL context, thus ensuring a comprehensive search of the research field. The search yielded 45 publications in total. Of those, 22 were duplicates, two conference proceedings, and one book chapter, which were all excluded (Figure. 1).

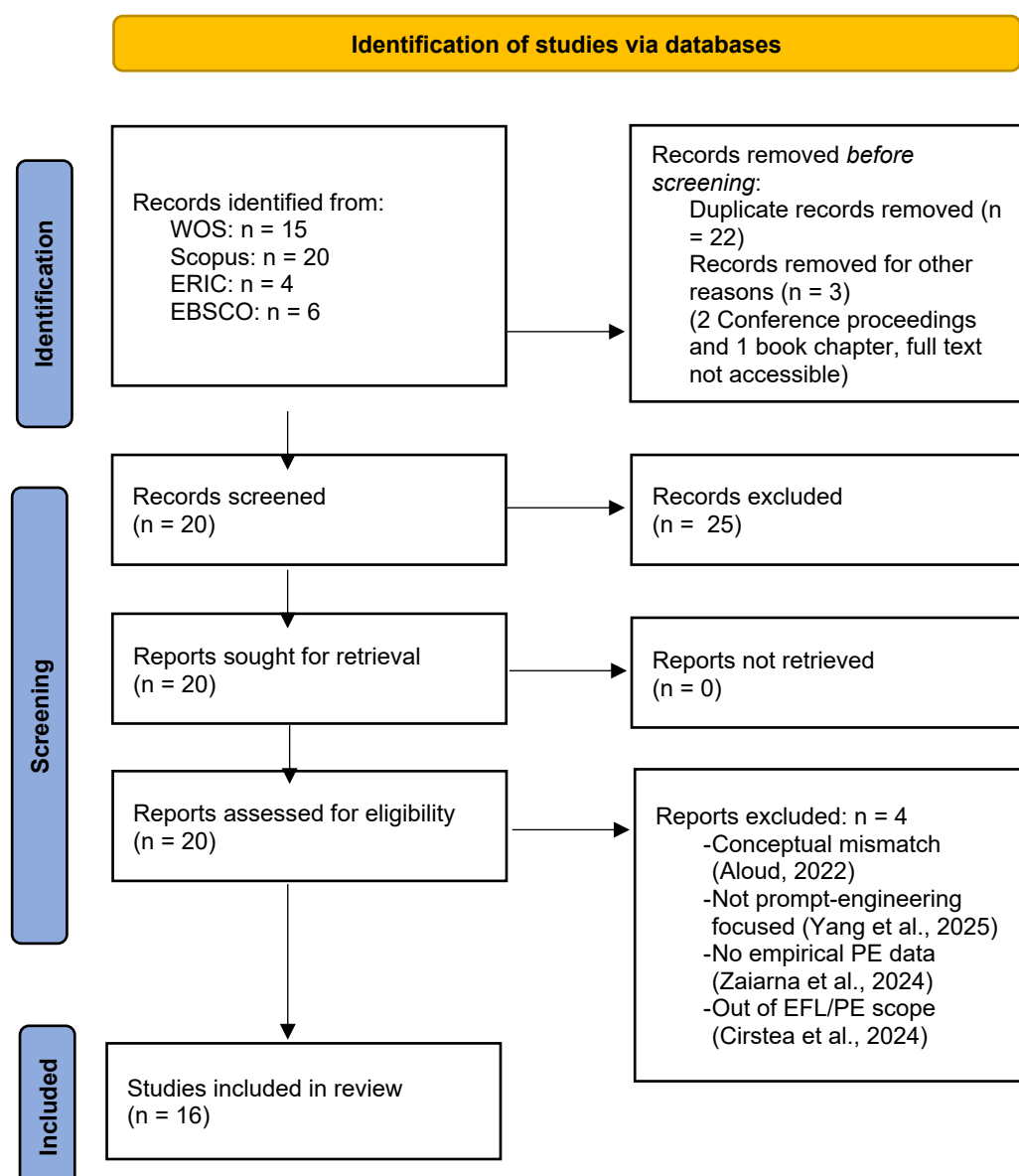


Figure 1. PRISMA flowchart of the study (Page et al., 2021)

The third stage included selecting studies based on inclusion and exclusion criteria. This study adopted three fundamental inclusion criteria: research conducted directly within an EFL context, empirical data production focused on prompt engineering, and reliance on data obtained from human participants or systematic analysis of AI outputs. Studies that did not meet these criteria were excluded (n = 4). For example, the study by Aloud (2022) was excluded due to conceptual mismatch, as it addressed the notion of prompting solely within the corrective feedback literature, specifically referring to output-prompting strategies in teacher-student interactions. Similarly, although the study by Yang et al. (2025) presents a systematic review on GenAI-supported L2 writing, it only addresses prompt engineering as a secondary dimension and does not provide primary application-oriented data; therefore, it was not included in the dataset. In the study by Zaiarna et al. (2024), although the abstract section mentions the goal of developing guidelines for prompt design in the future, no discussion or empirical findings on this topic are found in the body of the article. Therefore, the statement in question was limited to the abstract and was not systematically addressed throughout the study; consequently, it was excluded from the scope of this research. Finally, Cirstea et al.'s (2024) research focuses on digital electronic system-on-chip design and was excluded on the grounds of contextual incompatibility, as it is not related to the EFL or prompt engineering context. To assess the consistency of the screening process, an independent expert reviewed

20 studies (10 included and 10 excluded) using the predefined eligibility criteria. Discrepancies between the two assessments were resolved through discussion.

The next step involves extracting and organising the data. At this stage, the articles selected through the screening process ($n = 16$) were systematically reviewed using a standardised data extraction scheme developed in line with the research questions. Within this scheme, categories such as author(s) and publication year, country/context, sample, research method, and data collection tool were recorded for each study (See Appendix A). To ensure the comparability of findings produced by studies in different contexts, all information was tabulated and placed within a common framework. Thus, the selected articles were not only summarised at a descriptive level but also systematically organised in terms of methodological trends, participant distributions, and application forms. Data analysis was performed using a descriptive content analysis approach. The studies reviewed were systematically classified according to the categories defined in the research questions (country distribution, participant profiles, methodological preferences, etc.). The findings have been reported in a transparent and reproducible manner in accordance with the PRISMA-ScR guidelines.

4. Findings and Discussions

4.1 General Profile of Studies

The studies examined in this compilation exhibit a wide variety in terms of geographical context, participant profile, and methodological orientations. Table 1 below illustrates the characteristics of the studies included in the review.

Table 1.

Characteristics of the Studies Included in the Review

Authors & Year	Country/Context	Sample	Methods
Abdelhalim, 2024	Saudi Arabia	27 undergraduate students	Mixed methods
Abduh et al., 2025	Indonesia, high schools & university, students from diverse ethnolinguistic backgrounds	30 students & 8 teachers	Qualitative; multiple case study
Alkhateeb et al., 2025	Saudi Arabia, 2 universities	61 instructors	Mixed methods
Alsagoafi & Alomran, 2025	Saudi Arabia, public schools	169 teachers	Mixed methods
Arslan, 2025	Türkiye	230 instructors (scale); 15 instructors (interviews).	Mixed methods
Hwang et al., 2025	South Korea	10 graduate students	Mixed methods
Kerr & Kim, 2025	South Korea	8 pre-service teachers	Qualitative case study
Kohnke et al., 2025	Hong Kong - English-medium instruction university	19 pre-service students	Qualitative case study
Mizumoto & Teng, 2025	Japan	143 university students	Comparative study
Nguyen, 2024	Methodological study; author based in Vietnam	No human participants	Design-based study
Urazbayeva et al., 2024	Kazakhstan, in-service university EFL teachers from three universities	24 teachers	Mixed-methods
Woo et al. 2024	Hong Kong, secondary school EFL students in a classroom-based workshop context	21 students	Mixed-methods

Authors & Year	Country/Context	Sample	Methods
Woo et al., 2025a	Hong Kong, secondary school	67 students	Qualitative exploratory
Woo et al., 2025b	Hong Kong, secondary school	4 high-achieving students	Qualitative multiple case study
Woo et al., 2025c	Hong Kong, secondary school	107 students	Quantitative method
Zheng & Stewart, 2024	China, EFL curriculum unit on "Morals and Virtues"	No human participants	Qualitative experimental

4.2 Country and Context Distribution

The majority of the studies are concentrated in the Asian context. Specifically, Hong Kong (Kohnke et al., 2025; Woo et al., 2024a; Woo et al., 2025a; Woo et al., 2025b; Woo et al., 2025c), South Korea (Hwang et al., 2025; Kerr & Kim, 2025), Saudi Arabia (Abdelhalim, 2024; Alkhateeb et al., 2025; Alsagoafi & Alomran, 2025), Japan (Mizumoto & Teng, 2025), Türkiye (Arslan, 2025), Indonesia (Abduh et al., 2025), Kazakhstan (Urazbayeva et al., 2024), and China (Zheng & Stewart, 2024) are prominent research areas. This distribution indicates that AI-supported prompt engineering research in the field of EFL still has an Asian-centric focus.

When it comes to participant types, the studies display significant diversity. The vast majority of the examined studies were conducted directly with EFL learners. For example, studies conducted with secondary school students in Hong Kong (Woo et al., 2024a; Woo et al., 2025a; Woo et al., 2025b; Woo et al., 2025c), research conducted on university students in Japan (Mizumoto & Teng, 2025), and a study focusing on graduate students in South Korea (Hwang et al., 2025) are examples of this group. In addition, studies that directly address teachers are also noteworthy. For example, Kerr and Kim (2025) worked with pre-service teachers, Urazbayeva et al. (2024) with in-service university teachers, and Arslan (2025) and Alkhateeb et al. (2025) with EFL teachers at various levels. Furthermore, mixed samples examining both teachers and students are also available. The study with students and teachers from different ethnolinguistic groups by Abduh et al. (2025) is an example of this. Non-participant designs are also present in the literature, with Nguyen's (2024) study conducted solely on AI-generated data and Zheng and Stewart's (2024) research conducted on texts generated by ChatGPT.

Another finding of this study is that most studies are based on mixed-method designs (e.g., Abdelhalim, 2024; Alkhateeb et al., 2025; Arslan, 2025; Hwang et al., 2025; Urazbayeva et al., 2024), indicating that multidimensional approaches combining surveys/data collection tools with in-depth interviews are prominent. Qualitative designs, particularly case studies (e.g., Kerr & Kim, 2025; Kohnke et al., 2025; Woo et al., 2025a, 2025b), are also noteworthy. Quantitative methods are rather limited. Woo et al. (2025c) examined variables such as motivation and cognitive load using a comparative experimental design. Furthermore, Nguyen's (2024) study is methodologically focused, examining data generated by AI rather than human participants through content analysis. Such designs indicate that prompt engineering is beginning to take its place on the research agenda not only in terms of its pedagogical but also its methodological dimensions.

4.3 Application Forms and Pedagogical Positioning of Prompt Engineering

The studies examined classify prompt engineering applications in the EFL context under three main approaches: guidance through direct teaching frameworks, development through indirect/trial-and-error methods, and intuitive use. Approaches based on direct teaching frameworks are only seen in the study conducted by Zheng and Stewart (2024). In this study, the researchers both referenced the CLEAR framework (Lo, 2023a; 2023b) found in the literature and developed a new framework they proposed called Navigation and Generation. Thus, they suggested steps that would give students more control over outputs by adding a systematic dimension to prompt design.

Approaches that developed indirectly or through trial and error have been highlighted in studies where students developed their own strategies through experiential processes. For example, Woo et al. (2025b) report that students developed individual prompting pathways, while Kerr and Kim (2025) show that pre-service teachers discovered strategies such as iterative prompting, personalisation, and peer collaboration through their own experiences. On the other hand, intuitive use is more commonly observed in studies where teachers and students use AI bots based on their existing knowledge and experience without systematic guidance. Studies in this context (Abdelhalim, 2024; Abduh et al., 2025; Alkhateeb et al., 2025; Alsagoafi & Alomran, 2025; Arslan, 2025; Hwang et al., 2025; Kohnke et al., 2025; Mizumoto & Teng, 2025; Urazbayeva et al., 2024; Woo et al., 2024a; Woo et al., 2025a; Woo et al., 2025c) exemplify situations that are not directly taught to participants but are discovered by users through trial and error and adaptation.

Pedagogically, prompt engineering serves three main functions. First, it is used as a strategic tool in processes such as lesson planning, writing instruction, and activity development (Kerr & Kim, 2025; Woo et al., 2024a; Woo et al., 2025a; Woo et al., 2025c). Second, it is a tool for developing metacognitive and critical awareness through the analysis of biases and errors in AI outputs (Mizumoto & Teng, 2025; Zheng & Stewart, 2024). Third, it stands out as an area that emphasises the need for ethical and contextual filtering. Particularly in teacher-focused studies, it is emphasised that the outputs of prompt engineering need to be filtered in terms of cultural appropriateness, ethical responsibility, and reliability to be used in a pedagogical context (Alkhateeb et al., 2025; Arslan, 2025; Woo et al., 2025a).

4.4 The Pedagogical Integration of Prompt Engineering

The studies examined revealed that the integration of prompt engineering into EFL pedagogy occurred across a wide range of contexts. Table 2 below demonstrates this wide range of contexts systematically.

Table 2

Pedagogical Integration Areas, Benefits, and Limitations of Prompt Engineering

Pedagogical Area	Studies	Reported Benefits	Reported Limitations
Writing Instruction	Woo et al. (2024a, 2025a, 2025b, 2025c); Kohnke et al. (2025); Mizumoto & Teng (2025)	Writing quality, motivation, critical awareness, creative production	Cognitive load, misclassifications, risk of superficial output
Lesson Planning & Material Development	Arslan (2025); Alkhateeb et al. (2025); Urazbayeva et al. (2024)	Efficiency, variety of materials, instructional support	Ethical concerns, issues with contextual appropriateness
Assessment & Evaluation	Alsagoafi & Alomran (2025); Abdelhalim (2024)	Variety in question generation, connection to research competencies, metacognitive awareness	Limited scope, dependence on traditional methods

One of the most prominent areas is writing instruction. In numerous studies conducted in Hong Kong (Kohnke et al., 2025; Woo et al., 2024a, 2025a, 2025b, 2025c), prompt engineering has been integrated into lesson-based applications to support students' writing processes. The study by Mizumoto and Teng (2025) in Japan, however, revealed the limitations of prompt engineering in classifying self-regulation strategies in the writing process. Another area of integration is lesson planning and instructional design. Studies conducted specifically with instructors (Alkhateeb et al., 2025; Arslan, 2025; Urazbayeva et al., 2024) suggest that they often use ChatGPT in material development, lesson planning, and assessment processes. Such studies indicate that prompt engineering is positioned as a tool for educators. In the context of assessment and measurement, Alsagoafi and Alomran (2025) found that teachers experimentally used prompt design when generating ChatGPT-based questions. Abdelhalim (2024), in turn, revealed the relationship between prompts and research competence and metacognitive awareness. These findings suggest that prompt engineering functions as a tool for providing effectiveness and diversity in assessment tasks.

When it comes to the reported benefits, it is found that they include increased efficiency, improved writing quality, pedagogical diversity, and the development of students' critical awareness. However, limitations also exist. Cognitive load and the risk of misdirection for students (Mizumoto & Teng, 2025), and ethical concerns and contextual appropriateness issues for educators (Alkhateeb et al., 2025; Arslan, 2025) are frequently reported. These findings demonstrate that the integration of prompt engineering into EFL pedagogy is multidimensional. It functions at different levels, from writing-focused student applications to teachers' planning and assessment processes; both pedagogical opportunities and methodological/ethical limitations emerge.

4.5 Overarching Themes and Research Gaps in Prompt Engineering within EFL Studies

The studies reviewed reveal three overarching themes regarding the use of prompt engineering in the EFL context: theoretical diversity and lack of unified framework, the need for ethical and pedagogical filtering, and the lack of long-term effects on student outcomes (Table 3).

Table 3

Overarching Themes and Research Gaps in Prompt Engineering within EFL Studies

Theme	Studies	Description
Theoretical diversity & lack of unified framework	Zheng & Stewart (2024); Woo et al. (2025b); Kerr & Kim (2025)	Only a few studies propose explicit frameworks. Most provide limited or no theory, showing fragmented approaches.
Ethical and pedagogical filtering	Arslan (2025); Alkhateeb et al. (2025); Urazbayeva et al. (2024); Zheng & Stewart (2024)	Multiple studies stress filtering AI outputs for reliability, contextual fit, and academic integrity, highlighting pedagogical and ethical oversight.
Short-term gains, long-term uncertainty	Woo et al. (2025c); Mizumoto & Teng (2025)	Benefits include motivation, creativity, and better writing processes; challenges include cognitive load, overreliance, and limited critical engagement. Long-term effects remain unclear.

The studies analysed reveal that a comprehensive theoretical framework for prompt engineering applications has yet to be established. While some studies referred to broader technology-integration frameworks such as TPACK, these do not specifically address prompt engineering. Among the reviewed works, Zheng and Stewart (2024) stand out as the only study to explicitly ground their analysis in a prompt-related framework, namely CLEAR (Lo, 2023a; 2023b), while also introducing their own Navigation and Generation model. Beyond this, most studies either provide limited theoretical explanations or none at all, suggesting that a fragmented approach prevails in the field. Another overarching theme concerns the ethical and pedagogical filtering of AI outputs. Studies such as Arslan (2025), Alkhateeb et al. (2025), and Urazbayeva et al. (2024) emphasise the need for teachers to evaluate the contextual appropriateness, reliability, and academic integrity of GenAI outputs. This indicates that prompt engineering is not only a technical skill but also a pedagogical responsibility that requires critical oversight.

A further recurring theme relates to the short-term benefits versus the uncertain long-term effects of prompt engineering in EFL pedagogy. For instance, Woo et al.'s (2025c) quantitative study showed that students participating in ChatGPT-assisted writing workshops experienced higher motivation and produced more coherent texts. Similarly, Mizumoto and Teng (2025) demonstrated short-term gains in planning, revision, and evaluation strategies. Yet, both studies also reported limitations such as increased cognitive load, outputs lacking contextual depth, and insufficient development of critical evaluation skills. Together, these findings suggest that while prompt engineering holds clear short-term pedagogical promise, its longer-term impact on learning outcomes remains unclear.

5. Discussion

The findings of this review indicate that prompt engineering research exhibits a profile characterised by a regional focus on Asia, a participant-based emphasis on students, and a methodological dominance of qualitative/mixed designs. The prominence of Hong Kong, South Korea, Japan, Indonesia, and China points

to a significant imbalance in global representation. This is consistent with comprehensive reviews reporting that ChatGPT studies in EFL are particularly concentrated in Asia (Lo et al., 2024) and aligns with bibliometric findings depicting China's leadership in terms of publication volume/citation impact in the field of GenAI-education (Derinalp, 2024). In terms of participant profiles, the literature is predominantly student-focused. University and secondary school students constitute the majority of the studies, while teacher-focused or teacher-student mixed samples are relatively limited. Yang et al.'s (2025) extensive review also supports this finding and highlights the predominance of higher education samples. On the other hand, Kundu and Bej's (2025) school-based systematic review revealed that the development of language skills in student outcomes stands out, but teacher readiness and institutional infrastructure deficiencies limit research diversity. Therefore, the existing literature addresses teacher experiences and pedagogical integration dimensions only to a limited extent.

From a methodological perspective, most studies rely on qualitative or mixed designs. Case studies and small-scale mixed designs are particularly prominent (e.g., Kerr & Kim, 2025; Woo et al., 2025a, 2025b). Quantitative studies, on the other hand, are quite limited and are generally confined to short-term experimental designs (e.g., Woo et al., 2025c). Furthermore, notably absent are scale development or large-scale quantitative validation studies aimed at measuring prompt engineering as a pedagogical competency. This methodological trend indicates that while findings in the field provide deep contextual insights, they remain limited in terms of generalisability and comparability. Consequently, the current research profile exhibits an Asia-centric, student-focused, and qualitative/mixed-methods-dominant appearance. While consistent with the literature in the field, this trend clearly indicates a need for prompt engineering to be investigated in broader and more balanced pedagogical contexts, including teacher-focused and quantitative validation studies such as scale development.

The findings of this review reveal that prompt engineering applications in the EFL context are still unsystematic and fragmented. Only one study (Zheng & Stewart, 2024) referenced Lo's (2023a, 2023b) CLEAR framework and developed a new one. This demonstrates that basing prompt engineering on pedagogically structured frameworks is an exceptional example, with the general trend being based on framework-free, application-focused experiments. This is in line with the evaluations of AI and language education in the literature. Likewise, Danesi (2024) observes that while AI offers valuable opportunities for blended and interactive pedagogy, little attention has been given to systematically theorising how learners and teachers should engage with GenAI through prompts. This reinforces the view that prompt engineering in EFL remains under-theorised and requires more explicit conceptual models.

Findings indicate that students and pre-service teachers mostly develop their own prompting pathways through trial and error or intuitive strategies (Kerr & Kim, 2025; Woo et al., 2025b). Given that learner autonomy is essential to the learning process (Derinalp et al., 2025), this offers positive potential in terms of learner autonomy and strategy diversity, but it also carries risks of cognitive load and a lack of direction. As Ausubel (1964) noted in his critique of discovery learning, learning processes through trial and error do not automatically guarantee meaningful learning and are often limited to mechanical repetition. Similarly, the largely intuitive or experiential nature of prompt engineering applications in the EFL context suggests that learning outcomes may lack depth. This finding supports the view that the absence of structured pedagogical frameworks creates a critical gap in this area.

Additionally, some studies clearly demonstrate that prompt engineering is not only a technical skill but also requires pedagogical and ethical filtering. Teacher-focused research, in particular, emphasises the need to filter AI outputs for reliability, contextual appropriateness, and academic integrity (Alkhateeb et al., 2025; Arslan, 2025; Woo et al., 2025a). This finding aligns with broader ethical and pedagogical discussions in the AI-education literature. For instance, a study conducted in the Algerian context revealed that while students found ChatGPT useful for writing support, a significant portion remained uncertain about the ethical boundaries of AI usage. Approximately 17% of participants did not express a clear opinion on whether unattributed AI outputs constitute plagiarism (Cheriguene & Dahbi, 2025). This situation demonstrates that ethical boundaries are not always clear to students and highlights the need for explicit guidance on AI literacy and academic integrity in teaching programs. Furthermore, ethical responsibility is

not limited to students. When human-AI interaction is considered a collaborative process, mutual oversight and shared responsibility should be conceptualised (Cañas, 2022). In other words, teachers' or institutions' oversight of the accuracy and pedagogical appropriateness of AI outputs is not merely a technical task but also part of ethical responsibility sharing. This perspective requires rethinking the pedagogical integration of prompt engineering from an ethical collaboration perspective.

When the findings are evaluated collectively, it becomes apparent that prompt engineering practices in the EFL context are shaped around three fundamental pedagogical tensions, namely, theory versus practice, learner autonomy versus cognitive load, and technical output versus ethical judgment. First, since only two of the studies used a clear prompt engineering framework, this indicates that the field has largely developed through trial-and-error-based practices. While this approach supports students in independently developing their interactions with AI, the ongoing process of trial and error and the evaluation of outputs can increase the cognitive load. Second, well-structured prompts can provide useful and goal-oriented outputs in the short term; however, the effects of this practice on long-term learning and thinking processes are not yet fully understood. Finally, even technically successful outputs must be evaluated in terms of ethical criteria such as accuracy, academic integrity, and contextual appropriateness. These complexities demonstrate that prompt engineering is a process that transcends simply technical skill and requires maintaining a balance between learner autonomy, cognitive load, long-term learning, and ethical judgment.

5.1 Implications and future directions

Current findings indicate that three fundamental pedagogical tensions emerge in prompt engineering research: the imbalance between structured theoretical frameworks and intuitive trial-and-error practices, the dilemma between learner autonomy and cognitive load, and the tension between technical output production and the need for ethical/pedagogical filtering. This picture reveals that prompt engineering is not yet an institutionalised field of competence and highlights the urgent need for systematic frameworks, teacher training programs, and ethics-centred research. However, the lack of valid and reliable measurement tools in the field points to a significant gap. Developing scales that can assess dimensions such as students' prompt literacy, teachers' pedagogical and ethical filtering competency, or institutions' AI integration capacity will increase comparability in the literature and enable evidence-based decision-making in practice. For educators, the findings demonstrate that prompt engineering is not only a technical skill but also a process requiring pedagogical design, critical thinking, and ethical awareness. This necessitates a restructuring of teacher education at both the content and application levels. For researchers, the study highlights the lack of a theoretical framework, the narrowness of participant profiles, and the uncertainty of long-term effects, pointing to the need for more interdisciplinary and holistic approaches. At the institutional level, universities and professional associations can play a critical role in filling these gaps through ethical guidelines, teacher training programs, and implementation guides. Therefore, it can be said that the current findings not only map the existing boundaries of the field, but also shape the research, education, and institutional application agenda for the future systematic, measurable, and sustainable positioning of prompt engineering in EFL pedagogy.

5.2 Limitations

This study aims to map the current state of prompt engineering research in the EFL context and has several limitations. First, the scope of the review is limited to post-2022. This historical choice is related to the fact that GenAI tools only became a subject of research in the field of education from this period onwards. However, given the extremely fast-paced nature of the field, there is a high probability that the current findings will become outdated in a short period of time; therefore, regularly updated reviews are needed. Secondly, this review is limited to peer-reviewed journal articles indexed in WoS, Scopus, ERIC, and EBSCO, in accordance with predefined inclusion criteria. Consequently, grey literature such as thesis, dissertations, and conference papers was not included in the study. While this approach has provided greater consistency in terms of the quality and traceability of the evidence base, it may have excluded potentially relevant findings that were not published in indexed peer-reviewed journals. Future reviews could expand the evidence base by including grey literature in addition to peer-reviewed publications. Thirdly, the scoping review methodology aims to reveal common trends, differences, and gaps between studies;

however, it does not allow for inferences about the magnitude of effects or causal relationships reported in the findings. Therefore, the interpretation of findings remains largely at the level of painting a general picture of the field. In this regard, the development and systematic application of valid and reliable measurement tools in the future will increase the comparability of findings in the field, paving the way for more in-depth syntheses.

6. Conclusion

The purpose of this study was to map existing research on the use of prompt engineering in the context of EFL and to reveal trends, gaps, and future directions in the field. To this end, a search was conducted in four leading databases (WOS, Scopus, ERIC, and EBSCO) based on the PRISMA-ScR guidelines, and empirical studies published between 2022 and 2025 that met the inclusion criteria were examined. The findings point to three key points. First, the studies were largely Asia-centric in terms of region; participant profiles were mostly limited to students, and qualitative and mixed designs were dominant over quantitative studies in terms of methodology. Second, most prompt engineering applications are developed intuitively or through trial and error, with only a small number of studies directly grounded in a theoretical framework. Third, while prompt engineering is viewed as a strategic tool pedagogically, the need for ethical and contextual filtering is emphasised. It is highlighted that outputs must be filtered in terms of reliability, academic integrity, and contextual appropriateness.

These findings indicate that prompt engineering is not yet an institutionalised area of competence in EFL pedagogy. The lack of valid and reliable scales in the field limits the in-depth and measurable investigation of the pedagogical effects of prompt engineering. Therefore, it appears critical that future research focus on scale development, strengthening theoretical foundations, and systematically addressing the ethical dimensions of prompt engineering in EFL teacher education. In terms of application, prompt engineering should be positioned in teacher education programs not only as a technical skill but as a process requiring pedagogical design and ethical awareness. In conclusion, this study provides an overview of the literature by mapping current trends in prompt engineering research.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, generative AI (ChatGPT-3.5) was used solely to improve the readability, clarity, and language of the text. The author carefully reviewed and edited all AI-assisted suggestions and take full responsibility for the final content of the manuscript.

References

- Abdelhalim, S. M. (2024). Using ChatGPT to promote research competency: English as a Foreign Language undergraduates' perceptions and practices across varied metacognitive awareness levels. *Journal of Computer Assisted Learning*, 40(3), 1261-1275. <https://doi.org/10.1111/jcal.12948>
- Abduh, A., Rosmaladewi, R., Arham, M., & Asnur, M. N. A. (2025). Empowering EFL learners through generative AI: A qualitative exploration of inclusive and autonomous learning in multicultural contexts. *International Journal of Language Education*, 9(1), 168-179. <https://doi.org/10.26858/ijole.v1i1.74760>
- Alam, A. (2021, November). Possibilities and apprehensions in the landscape of artificial intelligence in education. *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, 1-8. IEEE. <https://doi.org/10.1109/ICCICA52458.2021.9697272>
- Alkhateeb, A., Hezam, A. M. M., & Almuraikhi, A. A. (2025). Assessing the use of AI tools for EFL exam preparation at Saudi universities: efficiency, benefits, and challenges. *Cogent Education*, 12(1), 2507553. <https://doi.org/10.1080/2331186X.2025.2507553>
- Aloud, R. E. (2022). Saudi Female EFL Teachers' Cognition and Practices Regarding Online Corrective Feedback in Speaking Class. *Arab World English Journal*. <https://dx.doi.org/10.24093/awej/call8.4>

- Alsagoafi, A. A., & Alomran, H. S. (2025). Revolutionizing assessment: Leveraging ChatGPT for automated item generation: An AI-driven exploratory study with EFL teachers. *World Journal of English Language*, 15(6), 385. <https://doi.org/10.5430/wjel.v15n6p385>
- Andewi, W., Waziana, W., Wibisono, D., Putra, K. A., Hastomo, T., & Oktarin, I. B. (2025). From prompting to proficiency: a mixed-methods analysis of prompting with ChatGPT versus lecturer interaction in an EFL classroom. *Journal of Studies in the English Language*, 20(2), 210-238. <https://doi.org/10.64731/jssel.v20i2.282318>
- Arif, H., & Naeem, J. (2025). The impact of generative AI on learner autonomy and critical thinking in English as a Foreign Language (EFL) writing classrooms. *Journal of Applied Linguistics and TESOL (JALT)*, 8(3), 2264-2275. <https://jalt.com.pk/index.php/jalt/article/view/1249>
- Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32. <https://doi.org/10.1080/1364557032000119616>
- Arslan, S. (2025). English-as-a-foreign language university instructors' perceptions of integrating artificial intelligence: A Turkish perspective. *System*, 131, 103680. <https://doi.org/10.1016/j.system.2025.103680>
- Ausubel, D. P. (1964). Some psychological and educational limitations of learning by discovery. *The Arithmetic Teacher*, 11(5), 290-302. <https://www.jstor.org/stable/41184964>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Cardie, C., Chang, K., Chatterji, N. S., Chen, A. S., Creel, K. A., Davis, J. Q., Demszky, D., ... Liang, P. (2021). *On the opportunities and risks of foundation models [Preprint]*. arXiv. <https://arxiv.org/abs/2108.07258>
- Cañas, J. J. (2022). AI and ethics when human beings collaborate with AI agents. *Frontiers in Psychology*, 13, 836650. <https://doi.org/10.3389/fpsyg.2022.836650>
- Chapelle, C. A. (2005). Computer-assisted language learning. In *Handbook of research in second language teaching and learning* (pp. 743-755). Routledge.
- Cheriguene, A., & Dahbi, A. (2025). Between Assistance and Abuse: Exploring the Ethical Use of ChatGPT in Algerian EFL Context. *Journal of Legal and Economic Research*, 8(2), 78-111.
- Cirstea, M., Benkrid, K., Dinu, A., Ghiriti, R., & Petreus, D. (2024). Digital electronic system-on-chip design: Methodologies, tools, evolution, and trends. *Micromachines*, 15(2), 247. <https://doi.org/10.3390/mi15020247>
- Dai, Y., & Panghe, Q. (2025). Why LLMs Driven EFL Writing: A Case Study in Senior Hight School Based on Production-Oriented Approach. *International Journal of Learning and Teaching*. <https://doi.org/10.18178/ijlt.11.5.305-313>
- Danesi, M. (2024). *AI in foreign language learning and teaching: theory and practice*. Nova Science Publishers.
- Derinalp, P. (2024). Past, Present, and Future of Artificial Intelligence in Education: A Bibliometric Study. *Sakarya University Journal of Education*, 14(Special Issue-AI in Education), 159-178. <https://doi.org/10.19126/suje.1447044>
- Derinalp, P. (2026). In-Service English Teachers' Attitudes toward Artificial Intelligence in Education. *Ege Eğitim Dergisi*, 27(2). <https://doi.org/10.12984/egedfd.1856805>
- Derinalp, P., & Halife, M. (2025). Pre-Service English as a Foreign Language Teachers' Attitudes toward Artificial Intelligence. *Journal of Theoretical Educational Sciences*, 18(3), 609-629. <https://doi.org/10.30831/akukeg.1644354>

- Derinalp, P., Karjo, C. H., Andreani, W., Ying, Y., & Herawati, A. (2025). Online learning and learner autonomy: a comparative study of Turkish and Indonesian EFL students' perspectives. *Journal of Education and Learning (EduLearn)*, 19(3), 1401-1409. DOI: 10.11591/edulearn.v19i3.21741
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). Opinion Paper: "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, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681-694. <https://doi.org/10.1007/s11023-020-09548-1>
- Gligorea, I., Cioca, M., Oancea, R., Gorski, A., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. *Education Sciences*. <https://doi.org/10.3390/educsci13121216>
- Hwang, M., Lee, K. H., & Lee, H. K. (2025). A Word to the Wise: Crafting Impactful Prompts for ChatGPT. *System*, 103756. <https://doi.org/10.1016/j.system.2025.103756>
- 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>
- Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, 6, 100225. <https://doi.org/10.1016/j.caeai.2024.100225>
- Kohnke, L., Zou, D., & Xie, H. (2025). Microlearning and generative AI for pre-service teacher education: a qualitative case study. *Education and Information Technologies*, 1-28. <https://doi.org/10.1007/s10639-025-13606-5>
- Kundu, A., Bej, T. (2025). Transforming EFL Teaching with AI: A Systematic Review of Empirical Studies. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-025-00470-0>
- Levy, M. (1997). *Computer-assisted language learning: Context and conceptualization*. Oxford University Press.
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1-35. <https://doi.org/10.1145/3560815>
- Liu, Z.-M., Hwang, G., Chen, C., Chen, X., & Ye, X. (2024). Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 39, 466 - 490. <https://doi.org/10.1080/09588221.2024.2389923>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (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), 50. <https://doi.org/10.1186/s40561-024-00342-5>
- Lo, L. S. (2023a). The art and science of prompt engineering: A new literacy in the information age. *Internet Reference Services Quarterly*, 1-8. <https://doi.org/10.1080/10875301.2023.2227621>
- Lo, L. S. (2023b). The clear path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship*, 49(4). <https://doi.org/10.1016/j.acalib.2023.102720>

- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054.
- Mizumoto, A., & Teng, M. F. (2025). Large language models fall short in classifying learners' open-ended responses. *Research Methods in Applied Linguistics*, 4(2), 100210. <https://doi.org/10.1016/j.rmal.2025.100210>
- Nawi, S. M., Sain, N., Yusof, S. M., & Kamaludin, P. N. H. (2025). Generative AI in Academic Writing: Exploring ESL students' strategies and performance. *Environment-Behaviour Proceedings Journal*, 10(SI30), 55-61. <https://doi.org/10.21834/e-bpj.v10iSI30.6879>
- Nelson, A., Santamaría, P., Javens, J. S., & Ricaurte, M. (2025). Students' Perceptions of Generative Artificial Intelligence (GenAI) Use in Academic Writing in English as a Foreign Language. *Education Sciences*. <https://doi.org/10.3390/educsci15050611>
- Nguyen, P. A. (2024). Evaluating AI-Generated Language as Models for Strategic Competence in English Language Teaching. *IAFOR Journal of Education*, 12(3), 325-349.
- Noviyanti, S. D., Hartono, R., Pratama, H., & Bahri, S. (2026). Pedagogical Mediation Through Prompt Engineering: An Expert Evaluation of AI-Generated Feedback on Islamic-Integrated EFL Argumentative Writing. *International Journal of Advanced Computer Science and Applications*, 17(3). <https://doi.org/10.14569/IJACSA.2026.0170340>
- Owoc, M. L., Sawicka, A., & Weichbroth, P. (2021). Artificial intelligence technologies in education: Benefits, challenges and strategies of implementation. In M. L. Owoc & M. Pondel (Eds.), *Artificial intelligence for knowledge management: AI4KM 2019* (pp. 37–58). Springer. https://doi.org/10.1007/978-3-030-85001-2_4
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ: British Medical Journal*, 372. <https://www.jstor.org/stable/10.2307/27352404>
- Renfeng, J., Gang, Y., & Qi, S. (2025). The motivational impact of GenAI tools in language learning: a quasi-experiment study. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12701>
- Taşçı, S. (2026). Investigating the relationship between AI self-efficacy, AI anxiety, and gender among Turkish ELT students. *Journal of Educational Technology and Online Learning*, 9(2), 273-285. <https://doi.org/10.31681/jetol.1823794>
- Tiandem-Adamou, Y. (2024). Using Generative Artificial Intelligence to Support EFL Students' Writing Proficiency in University in China. *Journal of Educational Technology and Innovation*, 6(4). <https://doi.org/10.61414/jeti.v6i4.213>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473. doi: 10.7326/M18-0850
- Urazbayeva, G., Kussainova, R., Aibergen, A., Kaliyeva, A., & Kantayeva, G. (2024). Innovation Off the Bat: Bridging the ChatGPT Gap in Digital Competence among English as a Foreign Language Teachers. *Education Sciences*, 14(9), 946. <https://doi.org/10.3390/educsci14090946>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Wang, Q. (2025). EFL learners' motivation and acceptance of using large language models in English academic writing: an extension of the UTAUT model. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1514545>

- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Albarghouthi, A., Bakshi, S., Datta, A., & Gehrke, J. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT [Preprint]. *arXiv*. <https://arxiv.org/abs/2302.11382>
- Woo, D. J., Guo, K., & Susanto, H. (2025a). Exploring EFL students' prompt engineering in human-AI story writing: an activity theory perspective. *Interactive Learning Environments*, 33(1), 863-882. <https://doi.org/10.1080/10494820.2024.2361381>
- Woo, D. J., Guo, K., & Susanto, H. (2025b). EFL secondary students' use of ChatGPT for writing task completion pathways. *The Journal of Educational Research*, 1-14. <https://doi.org/10.1080/00220671.2025.2510382>
- Woo, D. J., Wang, D., & Guo, K. (2025c). Comparing Hong Kong secondary school students' perceptions of ChatGPT-assisted EFL writing. *The Journal of Educational Research*, 1-21. <https://doi.org/10.1080/00220671.2025.2525243>
- Woo, D. J., Wang, D., Guo, K., & Susanto, H. (2024). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process. *Education and Information Technologies*, 29(18), 24963-24990.
- Yang, C., Li, R., & Yang, L. (2025). Revisiting Trends in GenAI-Assisted Second Language Writing: Retrospect and Prospect. *Journal of Educational Computing Research* 1-45. DOI: 10.1177/07356331251367309
- Zaiarna, I., Zhyhadlo, O., & Dunaievska, O. (2024). ChatGPT in foreign language teaching and assessment: exploring EFL instructors' experience. *Information Technologies and Learning Tools*, 4(102), 176-191. <https://doi.org/10.33407/itlt.v102i4.5716>
- Zheng, Y. (D.), & Stewart, N. (2024). Improving EFL students' cultural awareness: Reframing moral dilemmatic stories with ChatGPT. *Computers and Education: Artificial Intelligence*, 6, 100223.