Journal of Education, Society & Sustainable Practice, 2026, 2, 30–42
https://journals.enviromindsolutions.com/jessp
eISSN: 3070-6246

Check for updates

Prompt engineering in education: A systematic review of theory, applications, and future directions

Saeed Anowar[a*] and Reshma Khatun[a]

[a] Department of Education, Aliah University, Park Circus Campus, Kolkata-14, West Bengal, India

Original Article

Article info

Received: 11 February 2026
Revised: 20 April 2026
Accepted: 22 April 2026
Published: 24 April 2026

Keywords

Prompt engineering
Generative AI
ChatGPT
Educational technology
Teacher readiness

How to cite

Anowar, S., Khatun, R. (2026). Prompt engineering in education: A systematic review of theory, applications, and future directions. Journal of Education, Society & Sustainable Practice, 2, 30–42. https://doi.org/10.63697/jessp.2026.10093

Corresponding author

saeed.edu.rs@aliah.ac.in

Abstract

The integration of Generative Artificial Intelligence (GenAI), particularly large language models (LLMs) such as ChatGPT, has significantly influenced current educational practices. Among the competencies necessary for effective interaction with LLMs, prompt engineering has emerged as a focal point. Prompt engineering encompasses the design, organization, and optimization of inputs to GenAI systems to produce accurate, relevant, and pedagogically meaningful outputs. Although empirical, conceptual, and policy-oriented studies on this topic are increasing, research on prompt engineering in education remains fragmented across disciplines and educational contexts. This paper systematically reviews the application of prompt engineering in education, synthesizing theoretical foundations, developmental trajectories, models, frameworks, and application domains. Drawing on recent peer-reviewed literature, with an emphasis on studies published in the past three years, the review consolidates findings on motivation, attitudes, knowledge, and skills related to prompt engineering among both learners and educators. The expanding role of AI in education prompts a critical inquiry: how can AI be integrated into teaching in ways that preserve and enhance the human dimensions of knowledge sharing? This paper introduces a framework for the responsible use of AI that reinforces, rather than replaces, the essential human elements at the heart of education.

1 Introduction

Artificial intelligence (AI) has seen a profound epistemic and useful revolution within educational contexts, developing from an outlying technical help to an important driver of pedagogical reform (Cruz-Aguilar, 2025). Early research in artificial intelligence in education (AIEd) primarily emphasized intelligent tutoring systems, adaptive learning, and automated evaluation systems intended to optimize instructional competence and learner performance (Holmes et al., 2019; Zawacki-Richter et al., 2019). These applications primarily mirrored a computational and system-centric approach, whereby AI was started as a tool for supporting preset learning drives rather than as a dynamic participant in knowledge generation.

The rise of generative AI (GenAI), especially large language models (LLMs) such as ChatGPT, has revolutionized education. Earlier AI systems possess the capacity to produce contextually rich, human-like textual outputs, engage in dialogs, and support complex academic and pedagogical tasks, including lesson planning, assessment design, academic writing, and research (Walter, 2024; Nguyen, 2024). Experimental studies suggest that both students and educators are progressively bringing GenAI tools into teaching and learning practices, driven by perceived gains in speed, personalization, and cognitive support (Dobrovská et al., 2024; Uwosomah and Dooly, 2025). This change presents new educational challenges related to ethics, trust, and the proof of AI-generated content (Alkaissi and McFarlane, 2023; Mijwil et al., 2023).

A growing volume of textual and empirical data shows that successful interaction with GenAI systems is neither obvious nor automatic. The educational value, accuracy, and usefulness of AI-generated outputs rely heavily on how users create prompts, define limits, and frame their questions (Spasić and Janković, 2023; Nazari and Saadi, 2024). As a result, prompt engineering has become a crucial intermediate skill that mediates communication between humans and AI tools. Recent research views prompt engineering as a multifaceted skill that integrates linguistic precision, domain knowledge, pedagogical reasoning, and ethical awareness, despite its origins in technical and computational domains (Federiakin et al., 2024; Walter, 2024). Meanwhile, meaning is co-constructed through interaction, iteration, and circumstantial conciliation; prompt engineering is hypothetically consistent with constructivist and socio-cultural perspectives on learning. It serves as cognitive pieces that influence learners' epistemic agency, support inquiry, and shape learning routes in educational settings. Emerging models conceptualize prompt engineering within broader AI literacy frameworks, coupling it to teachers' digital competencies, professional identity, and motivation to employ AI-integrated pedagogies (Ng et al., 2023; Lan, 2024; Guan et al., 2025). Empirical studies further suggest that focused prompt engineering interventions may boost learners' AI self-efficacy, conceptual comprehension, and task performance, although effects differ among disciplinary and instructional settings (Woo et al., 2024; Woo et al., 2025).

Prompt engineering has been applied across an extensive range of educational areas, demonstrating its versatility as both an instructional and investigative tool. In classroom teaching, structured prompts have been utilized to enhance lesson planning, content explanation, diversity, and the design of inquiry-based learning (Spasić and Janković, 2023; Ramnarain et al., 2024). Well-designed prompts improve instructional efficiency, allowing adaptive responses to varied learner needs (Ng et al., 2023; Ashby, 2024). In teacher education, prompt engineering training has been shown to improve pre-service teachers' readiness to integrate AI into pedagogical practice by strengthening AI self-efficacy, instructional confidence, and professional development (Lan, 2024; Guan et al., 2025; Uwosomah and Dooly, 2025). These discoveries align with broader research representation that positive attitudes and perceived competence are significant predictors of technology acceptance (Wong et al., 2008; Gbormittah et al., 2024). Assessment and feedback establish another major application area. Instructors or educators have employed prompt engineering to generate formative feedback, practice questions, rubrics, and exemplars, while highlighting verification strategies to alleviate hallucination and bias (Alkaissi and McFarlane, 2023; Tassoti, 2024). Curriculum project bids typically include syllabus development, preparation of learning outcomes, and inclusive and accessible creation, primarily for learners with diverse needs (Moni et al., 2007; Drigas et al., 2023). In education, prompt engineering has supported literature search, experimental design, and exploratory analysis, although scholars remain cautious about the increased dependence on AI-generated content without thorough human verification (Mijwil et al., 2023; Zhang et al., 2024).

Despite its growing reputation, prompt engineering remains under-theorized and unevenly integrated into teacher education programs and curricular frameworks. Prior reviews on artificial intelligence in education have mainly focused on technological applications, learning analytics, ethical concerns, or policy implications, with limited attention to prompt engineering as a diverse educational competence (Su et al., 2022; Holmes et al., 2023). Studies have often discussed prompt engineering as an area that is frequently discipline-specific, operationally fragmented, or oriented toward higher education and technical fields, leaving significant gaps in relation to pedagogical theory, assessment practices, and pre-service teacher preparation (Ramnarain et al., 2024; Lee and Palmer, 2025).

This review aims to examine the theoretical foundations of prompt engineering as an educational competence and to trace its development through existing models, frameworks, and assessment approaches. It further analyzes its applications across diverse educational contexts and learner populations, while critically assessing the methodological, pedagogical, and ethical challenges that shape future research. The unique contribution of this article is predominantly conceptual and pedagogical. Rather than offering a technical account of prompting, the review synthesizes prompt engineering as a multidimensional educational competence at the intersection of AI literacy, pedagogy, and ethical practice.

2 Prompt engineering

2.1 Theoretical foundations of prompt engineering in education

Prompt engineering in education draws on theoretical traditions that show the cognitive and social bonds between humans and GenAI. Instead of seeing it as a way to "talk to the machine," we should view it as a learning medium driven by pedagogical logic and how we regulate our own thinking. Recent research increasingly recognizes the ongoing ability as a distinct educational skill, drawing on established theories of motivation and human-technology interaction (Federiakin et al., 2024; Walter, 2024; Lee and Palmer, 2025).

From a constructivist point of view, refining a prompt is essentially a way of building knowledge through trial and error. This type of interaction proves the old socio-constructivist point: we learn by actively making sense of things, not just by sitting there and accepting it as is. When a student tests a hypothesis or develops a prompt based on a previous answer, they are basically using the same "scaffolding" and discovery-based learning methods that teachers have used for decades (Holmes et al., 2019; Su et al., 2022). In the classroom, this dialogue pushes students to analyze, synthesize, and develop their thought processes, as long as they aren't blindly trusting whatever the AI says (Zhang et al., 2024).

Bandura's social cognitive theory provides additional explanatory value, particularly through the concept of self-efficacy, which has been frequently identified as a significant factor in AI engagement. Empirical studies suggest that students and teachers with stronger AI self-efficacy are more likely to experiment with advanced prompting strategies, adaptive prompt structures, and persist in refining their prompts when AI outputs are unsatisfactory (Woo et al., 2024; Woo et al., 2025). Poorly perceived competence may result in shallow prompt use or avoidance of GenAI tools (Pokrivcakova, 2023; Dobrovská et al., 2024). These findings are consistent with broader motivational meta-analyses demonstrating that perceived competence significantly predicts autonomous engagement with learning technologies (Bureau et al., 2022; Li et al., 2025).

The technical pedagogical content knowledge (TPACK) framework positions prompt engineering at the intersection of technological knowledge, goal, and disciplinary understanding. Effective educational prompts require a technical understanding of how AI function; it also requires alignment with instructional purposes, learner characteristics, and subject-specific methods of knowing (Ng et al., 2023; Guan et al., 2025). In an extended TPACK framework, knowledge of AI systems positions prompt engineering as a core component of broader AI competency, alongside data awareness, ethical reasoning, bias identification, and the critical evaluation of AI-generated outputs (Zawacki-Richter et al., 2019; Walter, 2024). Researchers emphasize that the necessity of prompt engineering be taught together with ethical protections, particularly in response to documented risks, such as hallucination, misinformation, and academic integrity (Alkaissi and McFarlane, 2023; Mijwil et al., 2023). Ethical rules for AI in education further emphasize the significance of transparency, accountability, and human AI interactions (Aiken and Epstein, 2000; Holmes et al., 2023). From a human AI interaction (HAI) perspective, prompt engineering functions as a linguistic and cognitive interface through which users exert control over AI systems. Theoretical models of HAI emphasize serviceability, feedback loops, cognitive load, and interaction efficiency that directly influence prompt effectiveness and learning outcomes (Lee and Palmer, 2025). Studies indicate that inadequately structured prompts increase extraneous cognitive burden, whereas carefully designed prompts can improve learning efficiency when combined with verification and reflection strategies (Zhang et al., 2024; Tassoti, 2024).

2.2 Development and models of prompt engineering

At the inception, prompt engineering was largely confined to the domain of computer science and technological disciplines. At that time, a prompt was just a functional tool, a way to tighten system outputs through precise syntax, with very little thought given to actual teaching or learning. But as GenAI started to move into the classroom, the view began to shift. Researchers started seeing prompt engineering not just as a technological input, but as an educational skill that could be taught, practiced, and measured (Federiakin et al., 2024;

Walter, 2024). These academic debates tended to place prompt engineering under the broad category of "AI literacy," focused on how to make the tool run rather than how to properly incorporate it into the learning process (Holmes et al., 2019; Su et al., 2022). However, as the field advanced, a distinction arose. Researchers started to recognize that developing a prompt is separate from general communication; it is a modern-day capacity that needs strategic thinking, a full awareness of context, and a sense of ethical obligation (Federiakin et al., 2024). This accomplishment indicates a huge move away from a "tool-first" mindset and toward a much more human-centered approach to how human connect with AI.

Recent research has proposed multi-dimensional frameworks that conceptualize prompt engineering as a structured cognitive process rather than an ad hoc activity. These models typically integrate dimensions such as prompt structure comprehension, contextual framing, role and audience specification, constraint setting, and iterative refinement (Spasić and Janković, 2023; Nazari and Saadi, 2024). Two-layered prompt models differentiate between core task components (e.g., objective, content scope, and output format) and adaptive elements (e.g., role assignment, tone, constraints, and evaluation criteria). This distinction emphasizes the reflective and intentional nature of effective prompting, emphasizing metacognitive regulation over trial-and-error usage (Lee and Palmer, 2025). Empirical evidence suggests that beginners trained using structured prompt frameworks demonstrate higher quality outputs, enhanced critical assessment, and better AI self-efficacy compared to untrained peers.

3 Materials and methods

3.1 Systematic review process and analysis

This study adopted a semi-systematic review method following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework, which supports transparent reporting of study identification, screening, eligibility, and inclusion procedures (Page et al., 2021). A semi-systematic review was considered because the research on prompt engineering application in education is an emerging field, interdisciplinary in nature, and methodologically diverse. Unlike a narrowly bounded systematic review designed for a mature and highly standardized body of evidence, a semi-systematic review enables the integration of conceptual, empirical, and review-based studies to map the development of the field, identify major themes, and synthesize knowledge gaps (Snyder, 2019).

The review is organized around a layered conceptual model (**Fig. 1**). The top layer emphasizes theoretical foundations such as constructivist and socio-cultural perspectives, cognitive artifacts and scaffolding, and AI literacy frameworks. The middle layer highlights the practical expansion of prompt engineering across academic use, curriculum and lesson design, and assessment and feedback. The bottom layer captures the principal concerns emerging from the literature, particularly academic integrity, teacher education, and future pedagogical directions.

3.2 Data sources and search strategy

To ensure broad and representative coverage of the literature, the search was conducted on the following scholarly databases: Scopus, Web of Science, ERIC (Educational Resources Information Center), ScienceDirect, SpringerLink, Taylor & Francis Online, and Google Scholar. These databases were selected because they collectively provide strong scholarly coverage of peer-reviewed articles in education, educational technology, teacher education, and interdisciplinary research on AI.

The search strategy combined keywords related to prompt engineering, GenAI, and education. Representative search strings included: “prompt engineering” AND education, “ChatGPT” AND teacher education, “generative AI” AND pedagogy, “prompt engineering” AND pre-service teachers, “AI literacy” AND teacher readiness, and “large language models” AND education. Boolean operators such as “AND” and “OR” were used to combine terms, and minor modifications were made across databases according to platform-specific search

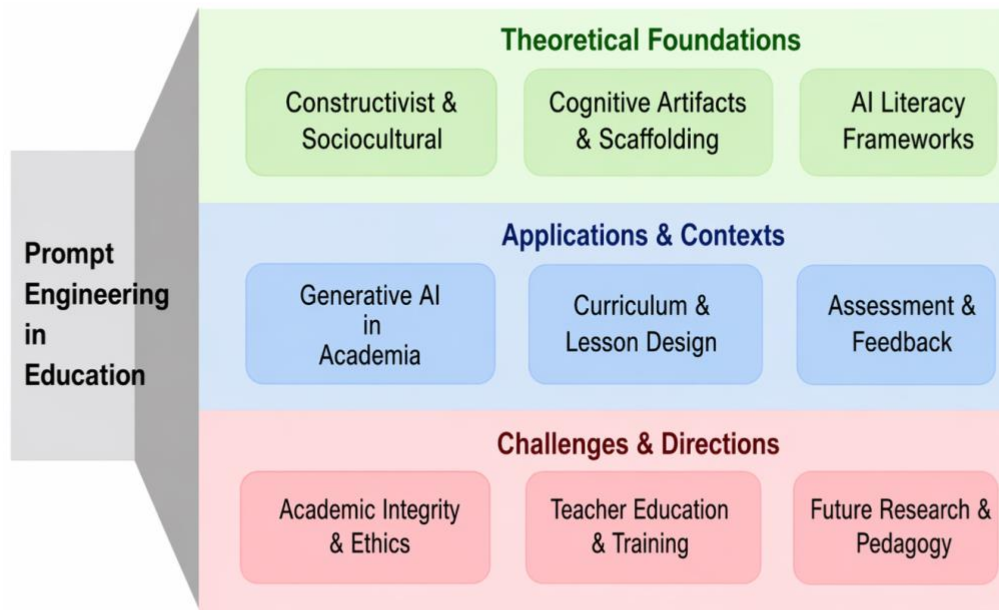

Figure 1. *Structural diagram of conceptual frameworks for prompt engineering in education.*

requirements. The search was primarily focused on studies published between 2019 and 2025, with particular emphasis on literature from the past three years.

3.3 Inclusion and exclusion criterion

Relevant research was included if it was (a) focused on prompt engineering or structured interaction with GenAI tools, (b) situated within educational settings, including teaching, learning, teacher education, assessment, curriculum, and educational research, (c) consisted of peer-reviewed journal articles, review papers, and policy-relevant research, and (d) published in English. To ensure topical relevance, approximately 30-40% of the included studies were published in the previous three years.

Studies were excluded if they were primarily focused on model optimization without clear educational relevance, were purely opinion-based or viewpoint articles without an analytical foundation or were focused solely on non-GenAI systems.

3.4 Screening and selection process

The screening and selection process followed the PRISMA 2020 sequence of identification, screening, eligibility, and inclusion, which provides an updated reporting framework for transparent documentation of study selection in the review process (Page et al., 2021). After the initial database search, duplicate records were removed. Titles and abstracts were then assessed to determine their relevance to prompt engineering and educational contexts. Full texts of potentially relevant studies were subsequently investigated against the inclusion and exclusion criteria. This process facilitated the selection of literature that was conceptually relevant, methodologically informative, and aligned with the objectives of the review.

3.5 Data analysis and synthesis

The selected studies were analyzed using thematic synthesis, a method that involves line-by-line coding of texts, the development of descriptive themes, and the generation of higher-order analytical themes to support cross-study interpretation and integration (Thomas and Harden, 2008). Principles of thematic analysis were also considered during the interpretive refinement of themes (Braun and Clarke, 2006).

4 Results and discussion

4.1 Classification of prompt engineering in education

Prompt engineering in education is broadly classified into four categories: conceptual and framework studies, empirical training studies, application studies, and ethical and policy studies (**Table 1**). This classification demonstrates that the field has expanded beyond technical experimentation and includes conceptualization, pedagogical interventions, classroom applications, and normative concerns related to ethics and governance. The conceptual and framework-oriented studies primarily define prompt engineering and identify key competencies, whereas empirical studies on training focus on skill development, AI self-efficacy, and instructional readiness. Application studies demonstrate the practical use of prompt engineering in classrooms, curriculum design, and assessment, while ethical and policy related studies address concerns related to academic integrity, privacy, and responsible AI governance.

There is a persistent disparity between positive dispositions toward GenAI use and the capacity to employ prompt engineering as a pedagogical practice. Across educational contexts, pre-service teachers report curiosity, confidence, and perceived usefulness of these tools, yet struggle to translate these perceptions into instructional strategies that align with learning objectives, learner variability, and assessment integrity (Walter, 2024; Guan et al., 2025). This disparity suggests that readiness for prompt engineering should be understood as a multi-dimensional construct that includes technical knowledge, pedagogical reasoning, ethical awareness, and professional identity. Similar trends have been observed in broader technology-oriented research, in which positive attitudes coexist with limited instructional change (Ng et al., 2023; Dobrovská et al., 2024).

4.2 Mapping of prompt engineering and the limits of technical optimization

The analysis revealed a connection between psychological readiness and professional disposition with the practical ability to use prompt engineering effectively in educational settings (**Table 2**). Motivation is commonly associated with perceived usefulness and relevance, and willingness to engage. Attitude is linked to acceptance, trust, and ethical

Table 1. *Classification of prompt engineering studies in education.*

Category	Focus	Typical methods	Key outcomes	References
Conceptual and framework studies	Definitions, models, and competencies	Conceptual analysis, and framework synthesis	Clarified constructs; proposed competency dimensions	Lee and Palmer (2025) Federiakin et al. (2024) Nazari and Saadi (2024) Ng et al. (2023)
Empirical training studies	Skill development and self-efficacy	Experimental design, and mixed methods	Improved prompt quality; AI self-efficacy	Woo et al. (2025) Tassoti (2024) Spasić and Janković (2023)
Application studies	Classroom, assessment, and curriculum	Case studies, and design-based research	Enhanced efficiency; pedagogical support	Spasić and Janković (2023) Spasić et al. (2024) Nguyen (2024) Ashby (2024) Nazari and Saadi (2024)
Ethical and policy studies	Integrity, privacy, and governance	Reviews, and policy analysis	Ethical guidelines; responsible use	Mijwil et al. (2023) Alkaissi and McFarlane (2023) Aiken and Epstein (2000) Holmes et al. (2019) Holmes et al. (2023)

Table 2. *Motivation, attitude, and competency mapping in prompt engineering research.*

Construct	Indicators	Findings	Citations
Motivation	Perceived usefulness, relevance, and willingness to engage with AI-supported learning	Structured interventions and AI-related training tend to increase motivation to use prompt engineering meaningfully in educational tasks	Woo et al. (2024) Li et al. (2025) Ramnarain et al. (2024)
Attitude	Acceptance, trust, ethical concern, and openness to adoption	Attitudes are generally positive, but often accompanied by concerns related to over-reliance, ethics, and quality of AI-generated outputs	Dobrovská et al. (2024) Pokrivcakova (2023) Uwosomah and Dooly (2025) Gbormittah et al. (2024)
Competency	Prompt structure, contextualization, refinement, and pedagogical alignment	Guided instruction and framework-based training improve prompt engineering performance and pedagogically relevant use of AI tools	Federiakin et al. (2024) Woo et al. (2025) Tassoti (2024) Ng et al. (2023) Walter (2024)

concern. While competency is essentially the ability to structure, contextualize, and refine prompts appropriately. The studies suggest that guided instruction strengthens competency, while motivation and attitude influence whether teachers and learners engage with AI-supported pedagogy in a meaningful and professional manner.

Ethical and moral considerations emerge as central, rather than incidental, challenges in understanding and practicing prompt engineering. Even when enhanced prompts improve the relevance and coherence of AI-generated outputs, they do not inherently resolve issues such as bias, hallucination, or epistemic responsibility (Alkaissi and McFarlane, 2023; Mijwil et al., 2023). In some cases, advanced prompting may even obscure fundamental limitations of AI systems, creating an illusion of reliability that undermines critical scrutiny. These tensions are particularly significant in teacher education, where future educators are expected to model accountable knowledge practices. Research on AI-supported writing and feedback highlights the delicate balance between efficiency and the preservation of authorship, originality, and reflective learning (Woo et al., 2024). From this perspective, ethical responsibility in educational AI cannot be delegated solely to technological systems; it must be actively cultivated through pedagogical guidance and institutional governance (Aiken and Epstein, 2000; Holmes et al., 2019). Therefore, prompt engineering should be framed not only as a technical skill but also as an ethically guided practice that requires critical evaluation.

Although GenAI is recommended as a tool in education, meaningful interaction with prompt-based systems requires linguistic ability, domain expertise, and metacognitive regulation, which are not equally accessible to all pre-service teachers (Woo et al., 2025). Consequently, those with stronger academic and digital understanding may experience greater advantages, while others remain marginal participants. This pattern aligns with critical perspectives in learning analytics and AI research, which caution that technological innovation can reproduce existing inequalities if contextual and cultural factors are overlooked (Zawacki-Richter et al., 2019).

4.3 Educational applications of prompt engineering

The major educational application domains and their associated outcomes are summarized, linking the conceptual and theoretical discussion to concrete pedagogical uses reported in the literature (**Table 3**). The reported outcomes include greater time efficiency, improved lesson differentiation, stronger teacher readiness, enhanced reflective practice, and research support. At the same time, several studies emphasize the continuing need for verification, ethical judgment, and human oversight. Accordingly, this review supports the argument that prompt

Table 3. *Educational applications and reported outcomes of prompt engineering.*

Domain	Application	Reported outcomes	Citations
Teaching	Lesson planning, content explanation, and classroom support	Greater time efficiency; improved differentiation; and stronger support for instructional preparation	Spasić and Janković (2023) Ashby (2024) Nguyen (2024)
Teacher education	Training workshops, professional preparation, and readiness building	Improved teacher readiness; enhanced reflective practice; and stronger AI-related pedagogical confidence	Woo et al. (2025) Guan et al. (2025) Ramnarain et al. (2024) Walter (2024)
Assessment	Feedback generation, question design, and rubric support	Faster feedback processes; improved assessment support; and continuing need for verification and oversight	Nguyen (2024) Nazari and Saadi (2024) Zhang et al. (2024) Tassoti (2024)
Curriculum design	Syllabus development, activity design, and alignment with learning outcomes	Better alignment with intended learning outcomes; and flexible integration of AI-supported instructional design	Lee and Palmer (2025) Nazari and Saadi (2024) Holmes et al. (2023) Su et al. (2022)
Educational research	Literature exploration, writing support, and research assistance	Improved efficiency in drafting and synthesis; and continuing need for validation and academic integrity safeguards	Mijwil et al. (2023) Alkaissi and McFarlane (2023) Abdullah and Elewah (2024)

engineering is emerging as an applied educational competence with implications for both instructional practice and professional development.

Despite growing evidence of the potential use of GenAI and prompt engineering in education, the literature reveals critical limitations that warrant careful consideration. Ethical issues often emerge, particularly regarding copying, lies, machine prejudice, and data privacy. Studies suggest that AI-generated content may include truthful errors or unfair information, particularly when prompts are not well-ordered or when verification mechanisms are insufficient (Alkaissi and McFarlane, 2023; Mijwil et al., 2023). In educational scenarios, such hazards call into question important principles like academic ethics and responsible knowledge production. Another concern is that overreliance on AI-generated information can inadvertently impede dangerous thinking, introspective decision-making, and creativity. Studies on AI-generated writing and feedback reveal that while productivity and output may improve, overreliance on generative technologies may diminish cognitive participation and metacognitive regulation (Zhang et al., 2024; Woo et al., 2024).

These findings are consistent with previous research outcomes in educational technology that needs automation, when adopted without careful evaluation, may replace rather than supplement higher-order learning processes (Holmes et al., 2019). Although many educators and pre-service teachers report good views toward AI, their levels of AI knowledge, teaching confidence, and prompt engineering skills remain unevenly developed (Ng et al., 2023; Guan et al., 2025). Several studies suggest that teachers often hold surface-level understanding of AI tools but lack deeper educational strategies for creating prompts matched with learning goals, student diversity, and assessment integrity (Walter, 2024; Federiakin et al., 2024). This suggests that positive views alone are insufficient for real educational change. Structural imbalances complicate the utilization of prompt engineering in education. Digital inequalities linked to access, institutional infrastructure, language proficiency, and professional development continue to define who benefits from AI-enabled learning (Zawacki-Richter et al., 2019). Previous studies suggested that students and instructors with superior academic, linguistic, and digital knowledge are more likely to utilize prompt engineering (Woo et al., 2024). Studies in higher education and teacher education highlight concerns that current assessment practices may no longer adequately measure individual knowledge or skill development in AI-supported learning environments (Nazari and Saadi, 2024; Tassoti, 2024).

While prompt engineering can enhance efficiency in lesson planning and content generation, excessive reliance may reduce opportunities for reflective practice, pedagogical reasoning, and the development of adaptive expertise (Federiakin et al., 2024). Empirical data indicate that when verification and assessment are limited, AI-supported processes may postpone learning and impair theoretical engagement (Zhang et al., 2024). GenAI must be positioned as an augmentative device rather than a replacement for professional competence. Studies on the future of education frequently indicate that contextual judgment, ethical decision-making, and interpersonal knowledge remain vital features of successful pedagogy (Drigas et al., 2023). Accordingly, prompt engineering must be integrated into educational frameworks that support teachers' capacity to make informed instructional decisions, engage in reflective inquiry, and maintain human oversight.

5 Challenges and future research directions

Despite increasing academic interest, many challenges remain in the systematic and sustainable integration of prompt engineering in education. From a methodological perspective, most available research relies on short-term interventions, cross-sectional designs, and self-reported measures of attitudes and perceptions (Dobrovská et al., 2024; Ramnarain et al., 2024). Although such studies provide valuable exploratory insights, they limit causal interference and offer only a partial understanding of how prompt engineering skills develop over time. Future research should focus on longitudinal and mixed-methods designs capable of capturing development trajectories, transferability across settings, and long-term instructional impacts.

Many studies focus primarily on the use of AI tools rather than on how prompt engineering could be embedded within subject-specific pedagogies, assessment methodologies, and learning theories such as constructivism, self-determination theory, and TPACK (Su et al., 2022; Li et al., 2025). Future research should examine how prompt engineering can enhance inquiry-based learning, differentiation, and inclusive teaching across academic disciplines. Evidence suggests that effective prompt engineering requires linguistic proficiency, domain expertise, and metacognitive skills, which are not evenly distributed among learners and instructors (Woo et al., 2025).

From a scientific perspective, the opacity of LLMs presents ongoing challenges related to explainability, bias, and reliability. As GenAI increasingly influences educational decision-making, research should investigate human-centered and explainable AI approaches that facilitate transparent collaboration between educators and AI tools (Holmes et al., 2023). At the policy level, the absence of clear institutional and regulatory guidelines continues to hinder responsible adoption. Future research should assess how national and international policy frameworks can address ethical use, assessment practices, data governance, and academic integrity in AI-supported education (Aiken and Epstein, 2000; Jamal, 2023).

6 Conclusion

Prompt engineering is not just a technical skill but an emerging educational competence that shapes how educators conceptualize, interact, and evaluate GenAI in educational settings. This review demonstrates that effective practice depends on the integration of cognitive, pedagogical, and ethical aspects, which require more than technical familiarity with AI tools. While GenAI can enhance instructional design, feedback processes, and professional learning, its educational value relies on teacher readiness, theoretical clarity, and institutional support structures. Ongoing challenges include methodological fragmentation, limited professional capacity, ethical guidelines, and inequitable access. This indicates that indiscriminate adoption increases risks of superficial engagement rather than transformative learning. Accordingly, prompt engineering should be methodically embedded in teacher education curricula, aligned with learning schemes, and guided by accountable AI policies to sustain pedagogical integrity.

7 Ethical statements

This research was constructed entirely on the synthesis and review of existing literature; it did not involve gathering primary data from human participants or animal subjects. For this reason, formal approval from an institutional review board was not necessary. We have ensured that every source and data used in this review were pulled from publicly obtainable records or accessed via academic databases, strictly following the standard ethical codes of research. During the preparation of this manuscript, the authors used Grammarly for language editing and structural refinement. All outputs were critically reviewed and edited by the authors. The authors take full responsibility for the published version of the manuscript.

8 Conflict of interest

We confirm that no financial, personal, or professional ties exist that could be understood as swaying the research, the writing process, or the eventual publication of this work.

9 Data availability statement

If readers require additional materials related to this research, such as datasets, figures, or the specific analytical materials developed during the review, these can be obtained by contacting the corresponding author.

10 Author contributions

S. Anowar: Conceptualization, formal analysis, investigation, methodology, and writing – original draft. R. Khatun: Methodology, supervision, and writing – review & editing. The final version of this paper was reviewed, refined, and formally accepted by both authors.

References

Abdullah, M. F., & Elewah, Y. A. (2024). AI in Education: Examining ChatGPT Awareness and Usage among University of Science & Technology Students, Aden. International Research Journal of Innovations in Engineering and Technology, 8(8), 26–36. https://doi.org/10.47001/irjiet/2024.808004

Aiken, R. M., & Epstein, R. G. (2000). Ethical guidelines for AI in education: Starting a conversation. International Journal of Artificial Intelligence in Education, 11, 163–176.

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. Cureus, 15(2), e35179. https://doi.org/10.7759/cureus.35179

Ashby, S. (2024). What are New Zealand primary teachers' beliefs and understandings of using ChatGPT to support their own practice? Kairaranga, 25(1), 22–36. https://doi.org/10.54322/fptqqb79

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bureau, J. S., Howard, J. L., Chong, J. X. Y., & Guay, F. (2022). Pathways to student motivation: A meta-analysis of antecedents of autonomous and controlled motivations. Review of Educational Research, 92(1), 46–72. https://doi.org/10.3102/00346543211042426

Cruz-Aguilar, M. A. (2025). The epistemic revolution of AI: Reconfiguring the foundations of scientific knowledge. AI & Society. https://doi.org/10.1007/s00146-025-02658-3

Dobrovská, D., Vaněček, D., & Yorulmaz, Y. I. (2024). Students' attitudes towards AI in teaching and learning. International Journal of Engineering Pedagogy (iJEP), 14(8), 88–106. https://doi.org/10.3991/ijep.v14i8.52731

Drigas, A., Chaidi, I., & Papoutsi, C. (2023). Teacher of the future. International Journal of Emerging Technologies in Learning (iJET), 18(16), 87–114. https://doi.org/10.3991/ijet.v18i16.36169

Federiakin, D., Molerov, D., Zlatkin-Troitschanskaia, O., & Maur, A. (2024). Prompt engineering as a new 21st century skill. Frontiers in Education, 9, 1366434. https://doi.org/10.3389/feduc.2024.1366434

Gbormittah, D., Korsah, D. P., Osiakwan, J. K., & Adom, G. (2024). Perceived attitude of preservice teachers towards the integration of technology in teaching and learning college geometry: cases of two selected colleges of education in Ghana. Discover Education, 3(1). https://doi.org/10.1007/s44217-024-00149-z

Guan, L., Zhang, Y., & Gu, M. M. (2025). Pre-service teachers preparedness for AI-integrated education: An investigation from perceptions, capabilities, and teachers' identity changes. Computers and Education: Artificial Intelligence, 8, 100341. https://doi.org/10.1016/j.caeai.2024.100341

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promise and implications for teaching and learning. The Center for Curriculum Redesign, Boston, MA, 02130, USA.

Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education. In Data ethics: building trust: how digital technologies can serve humanity (pp. 621–653). Globethics Publications. https://doi.org/10.58863/20.500.12424/4276068

Jamal, A. (2023). The role of artificial intelligence (AI) in teacher education: Opportunities and challenges. International Journal of Research and Analytical Reviews, 10(1), 139–144.

Lan, Y. (2024). Through tensions to identity-based motivations: Exploring teacher professional identity in Artificial Intelligence-enhanced teacher training. Teaching and Teacher Education, 151, 104736. https://doi.org/10.1016/j.tate.2024.104736

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: a systematic review to help inform curricula. International Journal of Educational Technology in Higher Education, 22(1). https://doi.org/10.1186/s41239-025-00503-7

Li, J., King, R. B., Chai, C. S., Zhai, X., & Lee, V. W. Y. (2025). The AI motivation scale (AIMS): A self-determination theory perspective. Journal of Research on Technology in Education, 1–22. https://doi.org/10.1080/15391523.2025.2478424

Mijwil, M., Hiran, K. K., Doshi, R., Dadhich, M., Al-Mistarehi, A.-H., & Bala, I. (2023). ChatGPT and the future of academic integrity in the artificial intelligence era: A new frontier. Al-Salam Journal for Engineering and Technology, 2(2), 116–127. https://doi.org/10.55145/ajest.2023.02.02.015

Moni, K. B., Jobling, A., van Kraayenoord, C. E., Elkins, J., Miller, R., & Koppenhaver, D. (2007). Teachers' knowledge, attitudes and the implementation of practices around the teaching of writing in inclusive middle years' classrooms: No quick fix. Educational and Child Psychology, 24(3), 18–36. https://doi.org/10.53841/bpsecp.2007.24.3.18

Nazari, M., & Saadi, G. (2024). Developing effective prompts to improve communication with ChatGPT: A formula for higher education stakeholders. Discover Education, 3(1), 45. https://doi.org/10.1007/s44217-024-00122-w

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. Educational Technology Research and Development, 71(1), 137–161. https://doi.org/10.1007/s11423-023-10203-6

Nguyen, T. C. (2024). University teachers' perceptions of using ChatGPT in language teaching and assessment. Proceedings of the AsiaCALL International Conference, 4, 116–128. https://doi.org/10.54855/paic.2349

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, 372, n71. https://doi.org/10.1136/bmj.n71

Pokrivcakova, S. (2023). Pre-service teachers' attitudes towards artificial intelligence and its integration into EFL teaching and learning. Journal of Language and Cultural Education, 11(3), 100–114. https://doi.org/10.2478/jolace-2023-0031

Ramnarain, U., Ogegbo, A. A., Penn, M., Ojetunde, S., & Mdlalose, N. (2024). Pre-service science teachers' intention to use generative artificial intelligence in inquiry-based teaching. Journal of Science Education and Technology, 34(6), 1272–1285. https://doi.org/10.1007/s10956-024-10159-z

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Spasić, A. J., & Janković, D. S. (2023). Using ChatGPT standard prompt engineering techniques in lesson preparation: Role, instructions and seed-word prompts. In 2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST) (pp. 47–50). IEEE. 2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST). https://doi.org/10.1109/icest58410.2023.10187269

Spasić, A., Mladenović, M., Nikolić, B., & Stošić, L. (2024). ChatGPT-assisted educational game development: A prompt engineering technique for teachers without coding experience. Unpublished. https://doi.org/10.13140/RG.2.2.16885.49122

Su, J., Zhong, Y., & Ng, D. T. K. (2022). A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region. Computers and Education: Artificial Intelligence, 3, 100065. https://doi.org/10.1016/j.caeai.2022.100065

Tassoti, S. (2024). Assessment of students use of generative artificial intelligence: Prompting strategies and prompt engineering in chemistry education. Journal of Chemical Education, 101(6), 2475–2482. https://doi.org/10.1021/acs.jchemed.4c00212

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8, 45. https://doi.org/10.1186/1471-2288-8-45

Uwosomah, E. E., & Dooly, M. (2025). It is not the huge enemy: Preservice teachers' evolving perspectives on AI. Education Sciences, 15(2), 152. https://doi.org/10.3390/educsci15020152

Walter, Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. International Journal of Educational Technology in Higher Education, 21, 15. https://doi.org/10.1186/s41239-024-00448-3

Wong, A. F., Chong, S., Choy, D., Wong, I., & Goh, K. (2008). A comparison of perceptions of knowledge and skills held by primary and secondary teachers: From the entry to exit of their preservice programme. Australian Journal of Teacher Education, 33(3). https://doi.org/10.14221/ajte.2008v33n3.6

Woo, D. J., Wang, D., Yung, T., & Guo, K. (2025). Effects of a prompt engineering intervention on undergraduate students' AI self-efficacy, AI knowledge and prompt engineering ability: A mixed methods study. British Educational Research Journal. https://doi.org/10.1002/berj.70087

Woo, D. J., Wang, D., Yung, T., & Guo, K. (2024). Learning prompt engineering for academic writing with ChatGPT: Postgraduate students' motivation, cognitive load, and satisfaction. https://doi.org/10.13140/RG.2.2.10781.52967

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1). https://doi.org/10.1186/s41239-019-0171-0

Zhang, R., Bi, N. C., Salerno, K., & Chen, P. (2024). ChatGPT might slow you down: Verification moderates AI uses and learning efficiency among college students. *Journal of Media Education, 15*(4), 32–37.