Journal of Linguistic and Language Education in Technology
https://ejournal.lp3msakti.com/index.php/JLLET/
Email: journaljllet@gmail.com
DOI: https://doi.org/10.36312/jllet.v1i2.74

November 2026, Vol. 1, No. 2
e-ISSN: 3164-7499
pp. 102-118

Designing Reasoning-Rich Tasks for ChatGPT-Assisted Academic Writing: A Qualitative Classroom-Based Study

[1]*Hasan Shaikh, [2]Aboe Bakar, [3]Yohanes Kabula
[1]English Discipline, Khulna University, Kobi Jibanananda Das Academic Building, Gollamari, Bangladesh
[2] Department of English and Business Communication, Lumbini Banijya Campus, Tribhuvan University, Devinagar, Butwal Sub-Metropolitan City-11/13, Rupandehi, Nepal
[3]English Language Education, Faculty of Humanities and Creative Industries, Maranatha Christian University, Jl. Prof. drg. Surya Sumantri, Bandung, West Java, Indonesia

*Corresponding Author Email: hasan.shaikh@ku.ac.bd

Received: September 2026, Revised: October 2026; Accepted: November 2026; Published: November 2026

Abstract

The increasing use of ChatGPT in academic writing raises an important pedagogical question concerning the conditions under which generative artificial intelligence promotes reasoning rather than encourages cognitive outsourcing. This qualitative classroom-based study investigated how Task-Based Language Teaching (TBLT) task characteristics shaped learners' reasoning and interaction with ChatGPT and identified task conditions and forms of teacher mediation that promoted critical engagement. Participants were 21 EFL learners from private educational institutions in Bangladesh who had experience with online learning, academic writing, and ChatGPT. Data were collected through semi-structured interviews and analyzed using thematic analysis combining deductive and inductive coding. The findings revealed a continuum between reasoning-poor and reasoning-rich task conditions. Descriptive, summarizing, grammar-focused, and reproduction-oriented tasks generally encouraged transactional interaction, including direct requests for drafts, paraphrases, and corrections. Conversely, argumentative, comparative, problem-based, source-based, and revision-oriented tasks promoted dialogic interaction characterized by follow-up questioning, counterargument generation, evidence verification, comparison, and explicit justification. Critical engagement was further strengthened by teacher-mediated verification requirements, justification of AI-related decisions, modeling of critical AI evaluation, dialogic questioning, structured comparison, peer discussion, and guided autonomy. The findings indicate that ChatGPT does not inherently strengthen or weaken reasoning; its cognitive value depends substantially on task architecture and instructional mediation. The study concludes that reasoning-rich TBLT should preserve learners' epistemic responsibility by requiring them to interrogate, verify, compare, revise, and justify AI-generated content. These findings provide practical implications for designing AI-assisted academic writing activities that promote critical reasoning rather than passive dependence.

Keywords: ChatGPT; Academic writing; Task-based language teaching; Critical reasoning; Generative artificial intelligence

Journal of Linguistic Metacognition and Language Learning is Licensed Under a CC BY-SA Creative Commons Attribution-ShareAlike 4.0 International License.

How to cite: Shaikh, H., Bakar, A., & Kabula, Y. (2026). Designing Reasoning-Rich Tasks for ChatGPT-Assisted Academic Writing: A Qualitative Classroom-Based Study, *Journal of Linguistic and Language Education in Technology*, 1(2), 102-118, Doi: https://doi.org/10.36312/jllet.v1i2.74

INTRODUCTION

The rapid expansion of generative artificial intelligence (GenAI) in higher education has transformed academic writing practices by giving students immediate access to tools capable of generating ideas, producing drafts, reformulating sentences, explaining concepts, and providing feedback. ChatGPT, in particular, has attracted substantial attention because of its potential to support personalized learning, writing development, problem solving, and

disciplinary learning across higher education contexts (Adamakis & Rachiotis, 2025; Ardyansyah et al., 2024; Naznin et al., 2025). At the same time, its integration raises pedagogical concerns regarding academic integrity, learner autonomy, over-reliance, accuracy, and the possible displacement of cognitive processes that students would otherwise perform themselves (Álvarez-Jiménez et al., 2026; Coman et al., 2026; Faisal, 2026). Consequently, the educational question is no longer simply whether ChatGPT should be used in university learning. A more consequential issue concerns the conditions under which its use strengthens rather than weakens reasoning. This concern is especially important in academic writing, where students are expected not merely to produce linguistically acceptable text but to construct arguments, evaluate evidence, distinguish credible from questionable claims, consider alternative interpretations, and make defensible rhetorical and epistemic decisions.

Research increasingly suggests that the cognitive consequences of GenAI use are not inherent properties of the technology itself. Rather, they are mediated by the way activities are designed and by what students are required to do with AI-generated information. GenAI may reduce unnecessary effort in routine processes, but the same efficiency can become problematic when learners transfer essential analytical work to the system (Adamakis & Rachiotis, 2025; Álvarez-Jiménez et al., 2026; Coman et al., 2026). A student asked simply to produce a polished essay with ChatGPT may interact with the system transactionally by requesting a complete draft, paraphrase, summary, or grammatical correction. By contrast, students required to challenge an AI-generated claim, compare competing explanations, verify evidence, identify logical weaknesses, or justify revisions must engage with the tool in more cognitively demanding ways (Ardyansyah et al., 2024; Li et al., 2026; Su et al., 2026; Wang et al., 2025). This distinction highlights a central pedagogical problem: AI-assisted writing cannot be adequately understood through a binary comparison between AI use and non-use.

Task-Based Language Teaching (TBLT) provides a useful framework for examining this problem because it places task design at the center of learning. TBLT emphasizes purposeful language use through activities that require learners to achieve communicative or problem-oriented outcomes. The cognitive demands created by a task—such as information gaps, competing perspectives, evidence constraints, decision making, and requirements for justification—shape how learners interact with language, information, and one another (Balacheff, 2024; Fu & Wang, 2026). When extended to AI-assisted academic writing, this perspective suggests that ChatGPT can function either as a cognitive scaffold or as a shortcut, depending on how the task organizes learner-AI interaction. Tasks that require argumentation, comparison, evaluation, and evidence-based decision making are more likely to position AI-generated content as material to interrogate rather than as an authoritative answer (Balacheff, 2024; Bollaert & Coppin, 2024; Decelles, 2023; Houdé, 2023). Therefore, the crucial pedagogical variable is not the presence of ChatGPT but the structure of the task within which ChatGPT is used.

Several task characteristics appear particularly relevant to reasoning-rich AI use. First, information gaps require learners to locate missing evidence, reconcile incomplete information, or construct explanations rather than merely reproduce generated content. Second, competing perspectives encourage students to evaluate alternative claims and defend their preferred interpretation. Third, evidence constraints oblige learners to verify AI-generated statements against scholarly sources or disciplinary standards. Fourth, iterative decision making requires students to revise initial positions in response to new information while articulating why particular changes are justified. Finally, explicit requirements to explain why AI suggestions are accepted, modified, or rejected create accountability for reasoning (Adamakis & Rachiotis, 2025; Álvarez-Jiménez et al., 2026; Su et al., 2026; Wang et al., 2025). Under these conditions, ChatGPT becomes less a producer of final answers and more a cognitive artifact that students

must evaluate, transform, and integrate into their own reasoning processes (Li et al., 2026; Naznin et al., 2025).

Task type may also shape students' interactional patterns with ChatGPT. Descriptive or reproduction-oriented assignments can encourage transactional interaction, in which students seek definitions, summaries, corrections, or ready-made paragraphs. Argumentative, source-based, comparative, and evaluative tasks, however, provide greater opportunities for dialogic interaction. Students may ask ChatGPT to generate counterarguments, test assumptions, identify weaknesses in a claim, compare interpretations, or critique draft reasoning. Studies across academic writing, mathematics, chemistry, medicine, and other higher education contexts emphasize the value of verification, explanation, and critical scrutiny when learners use AI systems (Ardyansyah et al., 2024; Nieto-Márquez et al., 2026; Su et al., 2026; Trčková et al., 2024; Wang et al., 2025). These findings indicate that task design can influence not only what students produce but also how they think, question, and make decisions during AI-assisted learning.

The integration of ChatGPT into TBLT simultaneously changes the instructional role of teachers. Teachers can no longer be understood solely as providers of linguistic knowledge or evaluators of final products. In AI-mediated environments, they increasingly function as designers of cognitive conditions, facilitators of AI literacy, monitors of reasoning processes, and evaluators of evidence-based decision making (Adamakis & Rachiotis, 2025; Arroyo-Bello et al., 2025; Naznin et al., 2025; Wang et al., 2025). This involves teaching students how to formulate productive prompts, verify claims, detect unsupported information, recognize bias, compare AI responses with disciplinary sources, and document their revision decisions. Assessment practices must therefore reward reasoning quality, source evaluation, justification, and transparency rather than judging only fluency or the quality of the final text (Adamakis & Rachiotis, 2025; Nieto-Márquez et al., 2026; Ostos et al., 2026). Teacher mediation becomes particularly important because students may otherwise interpret fluent AI-generated language as evidence of reliability.

Despite growing research on ChatGPT in higher education, an important gap remains. Much existing work examines students' attitudes, writing performance, ethical concerns, AI literacy, or the general benefits and risks of GenAI (Adamakis & Rachiotis, 2025; Coman et al., 2026; Harunasari, 2023; Naznin et al., 2025). Less attention has been given to how variation in TBLT task architecture systematically shapes students' reasoning behavior during ChatGPT-assisted academic writing. In particular, there is a need to distinguish task conditions that foster active interrogation, evidence checking, and reflective decision making from those that encourage passive acceptance or cognitive outsourcing. There is also limited understanding of how teachers perceive their instructional responsibilities when task design, rather than content delivery alone, becomes the principal mechanism for regulating productive AI use.

Accordingly, this study aims to investigate how the design of TBLT tasks shapes reasoning in ChatGPT-assisted academic writing, with particular attention to task characteristics, patterns of learner-AI interaction, conditions that promote critical engagement, and teachers' changing pedagogical roles. The novelty of the study lies in shifting the analytical focus from whether ChatGPT improves writing to how specific task architectures determine the quality of reasoning that occurs during AI-assisted writing. Rather than treating ChatGPT use as a uniform instructional condition, the study conceptualizes AI-supported writing along a continuum from reasoning-poor to reasoning-rich task design and examines the mechanisms through which students retain or relinquish epistemic responsibility. The study is guided by the following research questions: *How do different TBLT task characteristics and task types shape students' reasoning and patterns of interaction with ChatGPT during academic writing? And Which task conditions and forms of teacher mediation promote critical engagement with ChatGPT rather than passive acceptance of AI-generated output?*

RESEARCH METHOD

Research Design

This study employed a qualitative research design to investigate how Task-Based Language Teaching (TBLT) task characteristics shaped learners' reasoning and interaction with ChatGPT during academic writing. A qualitative approach was considered appropriate because the study sought to understand participants' experiences, interpretations, decision-making processes, and perceptions rather than to measure the magnitude of an intervention or establish causal relationships. Specifically, the study focused on how learners perceived different task conditions, how they evaluated and responded to ChatGPT-generated output, and which task characteristics encouraged critical engagement rather than passive acceptance. These questions required detailed accounts of learners' reasoning processes that could not be adequately captured through numerical measures alone.

The qualitative design was also suitable for examining variation in learner–AI interaction across academic writing tasks. Students may use ChatGPT in substantially different ways depending on whether a task requires description, argumentation, comparison, source evaluation, revision, or justification. A qualitative approach enabled the researchers to explore these differences through participants' explanations of how they generated prompts, evaluated responses, verified information, accepted or rejected suggestions, and revised their writing. In this sense, the design was closely aligned with the research objectives because it emphasized the processes through which TBLT task architecture may influence cognitive engagement with generative AI. The study adopted an interpretive orientation, assuming that learners construct meaning from their interactions with ChatGPT within particular instructional and task contexts. Rather than treating ChatGPT use as a uniform behavior, the study examined how participants interpreted task requirements and how those interpretations affected the ways in which they used AI during academic writing. This orientation allowed attention to be given to both common patterns and individual differences in reasoning, including cases in which students questioned AI responses, verified information independently, revised AI-generated content, or relied heavily on generated answers.

Nevertheless, the selected design involved several limitations. First, qualitative findings from a relatively small sample cannot be statistically generalized to all EFL learners in Bangladesh or other educational contexts. The purpose of the study was therefore analytical and interpretive rather than population-level generalization. Second, because much of the evidence depended on participants' retrospective descriptions during interviews, responses were potentially affected by recall bias or social desirability. Participants might, for example, present themselves as more critical or responsible users of ChatGPT than they actually were. To minimize this risk, interview questions focused on concrete experiences and specific examples rather than asking participants to make broad evaluations of their own ability. Probing questions were also used to clarify how particular decisions were made. A further potential source of bias concerned researcher interpretation during qualitative coding. This was addressed through systematic coding procedures, maintenance of a coding record, repeated comparison between themes and original interview data, and peer examination of a subset of coded data. Negative or contradictory cases were retained rather than excluded when they challenged emerging interpretations. These strategies strengthened the credibility and dependability of the analysis while acknowledging that qualitative interpretation inevitably involves researcher judgment.

Research Participants

The participants consisted of 21 EFL learners enrolled in private educational institutions in Bangladesh. Participants were recruited using purposive sampling, as the study required learners who had relevant experience with online learning, academic writing, and the use of digital technologies in English-language learning. Purposive selection was appropriate because

the research questions concerned specific experiences of ChatGPT-assisted academic writing rather than the characteristics of a statistically representative student population. To be eligible for participation, learners were required to meet several criteria. First, they had to be studying English as a foreign language in a private educational setting in Bangladesh. Second, they needed prior experience with online or technology-mediated learning. Third, participants were expected to have experience completing academic writing activities in English, such as paragraphs, essays, argumentative writing, summaries, reports, or related university-level writing assignments. Fourth, participants needed sufficient familiarity with ChatGPT to interact with the system during or in relation to academic writing tasks. Learners without prior experience of online academic writing or generative AI-assisted learning were not included because they would have limited experiential grounds for responding to the interview questions.

The sample size of 21 participants was considered appropriate for an in-depth qualitative inquiry because it enabled detailed examination of individual experiences while allowing recurring and contrasting patterns to emerge across participants. The study did not assume that 21 learners represented the broader population of Bangladeshi EFL learners. Instead, sample adequacy was assessed in relation to the depth and relevance of the data generated. Data collection continued until interviews produced substantial repetition in the major categories relevant to task design, interaction with ChatGPT, verification practices, reasoning, and learner decision making.

Demographic information was collected to contextualize participants' responses. This included age group, gender, educational level, length of English study, previous experience with online learning, frequency of ChatGPT use, and prior experience using ChatGPT for academic writing. These characteristics were used descriptively rather than as variables for statistical comparison. No personally identifying information was reported, and pseudonyms or participant codes were used during transcription and analysis to protect confidentiality. Participation was voluntary. Before data collection, participants were informed about the purpose of the study, the interview procedures, their right to decline to answer particular questions, and their right to withdraw without penalty. Informed consent was obtained before the interviews were conducted.

Instruments and Data Collection Technique

Data were collected primarily through semi-structured interviews. This instrument was selected because it provided sufficient consistency across participants while allowing the researcher to explore individual experiences in greater depth. The semi-structured format was particularly appropriate for investigating learner reasoning because participants could explain not only what they did with ChatGPT but also why they accepted, modified, questioned, or rejected particular AI-generated responses. The interview protocol was developed directly from the study objectives and research questions. Questions were organized around four major areas: task characteristics, learner–ChatGPT interaction, critical engagement with AI-generated output, and reasoning during academic writing. The interview also explored participants' perceptions of task conditions that encouraged or discouraged independent thinking.

Questions concerning task characteristics examined participants' experiences of descriptive, argumentative, comparative, evaluative, source-based, and problem-oriented academic writing tasks. Participants were asked whether particular task requirements changed the way they used ChatGPT. Questions concerning learner–AI interaction explored the types of prompts students used, whether interaction was limited to requesting ready-made content or developed into iterative exchanges, and how students responded when ChatGPT provided incomplete, questionable, or conflicting information. A third group of questions investigated critical engagement. Participants were asked whether and how they verified ChatGPT-generated information, compared AI responses with external sources, detected questionable

claims, considered alternative arguments, and decided whether to accept or reject suggestions. Further questions explored revision practices and whether students could explain the reasons behind their final writing decisions. These questions were particularly important for distinguishing passive acceptance from reasoning-rich engagement.

The initial interview protocol was subjected to expert review to establish content relevance and clarity. Reviewers with expertise in EFL instruction, academic writing, qualitative research, and technology-enhanced learning were asked to evaluate whether the questions adequately represented the study objectives and whether any items were ambiguous, leading, repetitive, or overly technical. The interview questions were subsequently revised based on their comments. A small-scale pilot interview was also conducted with learners who had characteristics similar to those of the intended participants but who were not included in the final dataset. The pilot was used to assess question clarity, sequencing, interview duration, and the effectiveness of probing questions. Questions that produced repetitive answers or failed to generate information directly related to the research objectives were reformulated or removed.

Because concepts such as validity and reliability operate differently in qualitative research, the study emphasized credibility, dependability, confirmability, and transparency. Credibility was supported by using follow-up questions to clarify participants' intended meanings and, where necessary, summarizing responses during interviews so participants could confirm or correct the researcher's interpretation. Dependability was strengthened by maintaining a consistent interview guide and documenting modifications to the coding framework. Confirmability was supported by retaining interview transcripts, coding decisions, analytical notes, and theme-development records as an audit trail.

Data Analysis

The interview data were analyzed using thematic analysis. This approach was selected because it enabled systematic identification and interpretation of recurring patterns concerning task characteristics, learner–ChatGPT interaction, critical evaluation, and reasoning. The analysis combined deductive and inductive coding. Deductive coding was informed by the research questions and conceptual framework, while inductive coding allowed unexpected patterns to emerge from participants' accounts. The analysis proceeded through several stages. First, recorded interviews were transcribed verbatim and checked against the audio recordings for accuracy. The transcripts were read repeatedly to achieve familiarity with the complete dataset. Initial notes were made regarding recurring expressions, critical incidents, differences between task types, and examples of learners accepting, questioning, verifying, or transforming ChatGPT output.

Second, initial codes were assigned to meaningful segments of the transcripts. Examples of possible codes included *requesting complete answers*, *grammar-focused use*, *asking for alternatives*, *checking external sources*, *questioning AI accuracy*, *comparing arguments*, *accepting without verification*, *rejecting AI suggestions*, *justifying revisions*, *prompt refinement*, and *independent position formation*. Coding was conducted at the semantic level while remaining attentive to underlying patterns in how participants described cognitive responsibility. Third, related codes were grouped into broader candidate categories. For example, codes such as requesting complete drafts, copying generated paragraphs, and accepting responses without verification could be grouped under passive or transactional AI use. Codes including requesting counterarguments, checking evidence, comparing interpretations, and challenging AI responses could form a category of critical-dialogic interaction. Similarly, codes relating to evidence requirements, competing perspectives, justification, and revision could be organized as reasoning-rich task conditions.

Fourth, candidate categories were compared across the 21 participants and developed into higher-order themes. Potential themes included task architecture as a driver of cognitive

engagement, transactional versus dialogic interaction with ChatGPT, verification as a mechanism of epistemic responsibility, justification and revision as reasoning triggers, and conditions promoting or constraining learner autonomy. Themes were reviewed against the entire dataset to determine whether they adequately represented both dominant patterns and contradictory cases.

A constant comparative process was used throughout the analysis. Statements from different participants were compared to identify similarities, differences, and variations associated with task type and learner behavior. Particular attention was given to negative cases, such as participants who continued to rely passively on ChatGPT even when tasks contained reasoning-oriented requirements, because such cases helped prevent overly deterministic conclusions about the influence of task design. To enhance the trustworthiness of the analysis, a subset of transcripts and codes was independently reviewed by another researcher or qualified peer. Differences in interpretation were discussed and coding definitions were refined where necessary. The purpose was not to achieve mechanical agreement but to identify unsupported interpretations and strengthen the transparency of theme development. The researcher also maintained analytical memos documenting why codes were merged, separated, or elevated into themes.

Finally, findings were reported using thematic descriptions supported by selected participant quotations. Quotations were chosen because they clearly illustrated analytical patterns rather than merely repeating the researcher's claims. Participant codes were used to preserve anonymity. Interpretation focused on explaining how particular task characteristics shaped different forms of ChatGPT use and how those interaction patterns either maintained learners' responsibility for academic reasoning or encouraged cognitive delegation to AI. The analysis therefore moved beyond identifying whether students used ChatGPT positively or negatively. Instead, it examined the mechanisms linking task design, learner–AI interaction, and reasoning behavior, which directly addressed the study's objective of distinguishing reasoning-rich from reasoning-poor forms of ChatGPT-assisted academic writing.

RESEARCH FINDINGS AND DISCUSSION

Research Findings

How do different TBLT task characteristics and task types shape students' reasoning and patterns of interaction with ChatGPT during academic writing?

The thematic analysis indicated that the nature of the TBLT task strongly influenced how learners interacted with ChatGPT and the extent to which they engaged in reasoning during academic writing. Participants generally described less cognitively demanding interaction when completing descriptive, summarizing, or language-correction tasks. In these situations, ChatGPT was frequently used transactionally to generate definitions, improve grammar, paraphrase sentences, or produce an initial draft. Learners tended to evaluate the output primarily in terms of linguistic fluency and relevance rather than evidential accuracy or argumentative strength. One learner explained, *"When the task is only to describe a topic, I normally ask ChatGPT to give me a paragraph and then I change some words"* (P4). Another stated, *"For grammar correction, I usually accept the suggestion because I think ChatGPT is better than me in grammar"* (P9). Similarly, P13 reported, *"If the teacher asks for a summary, I give the text to ChatGPT and ask it to make it shorter. I do not question it very much."* P17 commented, *"For simple writing, I use it mainly to save time because I already know what the answer should look like."* P21 added, *"When there is no requirement to explain my opinion, I mostly use ChatGPT for the final wording."* These responses suggest that relatively closed or reproduction-oriented tasks created fewer incentives for students to interrogate AI output and therefore more readily supported passive or efficiency-oriented use.

In contrast, argumentative, problem-based, comparative, and source-based tasks elicited more dialogic and iterative interaction with ChatGPT. Participants reported asking follow-up questions, requesting counterarguments, exploring multiple perspectives, and comparing the system's responses with their own ideas. Such tasks appeared to shift ChatGPT from a text-producing tool to a conversational resource for developing or testing reasoning. P2 explained, *"In argumentative writing, I ask ChatGPT to give the opposite side because sometimes I only see my own position."* P6 similarly remarked, *"When two viewpoints are given, I ask ChatGPT to explain both, but I decide which one is stronger after checking the evidence."* P10 stated, *"For a problem-solving task, one answer is not enough. I keep asking why and what would happen if the situation changes."* P15 noted, *"If I have to compare two ideas, I cannot just copy ChatGPT because I need to show the differences and explain why they matter."* Likewise, P18 explained, *"Argumentative tasks make me talk more with ChatGPT. I ask questions after its answer instead of using only the first response."* These excerpts indicate that tasks involving contestable claims and alternative solutions stimulated more sustained learner–AI dialogue because students needed to evaluate rather than merely retrieve information.

A third theme concerned evidence constraints and source verification. Participants indicated that tasks requiring scholarly evidence, citations, or source comparison made them more cautious toward ChatGPT-generated claims. Learners described checking references, searching for supporting sources, and rejecting information they could not verify. P1 stated, *"If the assignment asks for journal evidence, I cannot trust ChatGPT directly because sometimes the references are not real."* P7 explained, *"I search Google Scholar after ChatGPT gives me an idea, especially when I have to cite the statement."* P11 observed, *"When the teacher requires sources, I use ChatGPT only to find the direction, not as the evidence."* P14 similarly stated, *"I once found that the article suggested by ChatGPT did not exist, so now I check the title before using it."* P20 commented, *"A source-based task makes me more careful because I know I must prove where the information comes from."* These responses demonstrate that evidential requirements redistributed responsibility back to learners by requiring them to validate AI output against independently identifiable sources.

The analysis also showed that tasks requiring justification and revision produced greater metacognitive engagement. When learners had to explain why they accepted, modified, or rejected ChatGPT suggestions, they became more aware of their own decision-making processes. P3 explained, *"If I have to explain why I changed the paragraph, I cannot just say ChatGPT suggested it. I have to understand the reason."* P5 stated, *"The revision task made me compare my original paragraph and ChatGPT's version, and sometimes I kept my own sentence because it was clearer for my purpose."* P8 noted, *"When the teacher asks us to justify each major change, I think before accepting the correction."* P12 reported, *"I rejected one suggestion because ChatGPT changed the meaning of my argument. Before, I might not notice that."* P19 added, *"Writing a reflection after using ChatGPT makes me remember what I actually learned and what the AI did for me."* These accounts suggest that justification requirements functioned as accountability mechanisms, slowing down automatic acceptance and encouraging learners to articulate the basis for their choices.

The findings suggest a continuum between reasoning-poor and reasoning-rich task conditions. Descriptive, surface-editing, and production-focused tasks tended to support shorter, transactional exchanges in which students sought correct or fluent outputs. In contrast, argumentative, comparative, problem-based, source-based, and revision-oriented tasks generated more questioning, verification, iterative prompting, and explicit reasoning. Importantly, the results do not suggest that a task type automatically produces critical engagement. Some learners continued to rely heavily on ChatGPT even in complex tasks, whereas others independently verified information in relatively simple tasks. Nevertheless, the

dominant pattern indicates that task architecture strongly influenced how much cognitive responsibility students retained during ChatGPT-assisted writing.

Table 1. Summary of Different TBLT Task Characteristics and Task Types of Students' Reasoning and Patterns of Interaction with ChatGPT

Theme	Task characteristics	Typical learner–ChatGPT interaction	Reasoning outcome
Transactional AI use	Descriptive, summarizing, grammar correction, reproduction-oriented tasks	Single prompts, paraphrasing, correction, complete draft requests	Limited evaluation; emphasis on fluency and efficiency
Dialogic AI interaction	Argumentative, problem-based, comparative tasks	Follow-up questions, counterarguments, alternatives, iterative prompting	Greater analytical reasoning and perspective taking
Evidence verification	Source-based tasks, citation requirements, evidence constraints	Checking references, comparing sources, validating claims	Increased epistemic caution and source evaluation
Justification and revision	Reflective tasks, revision histories, explanation requirements	Comparing versions, accepting/rejecting suggestions, explaining changes	Greater metacognition and decision awareness
Task complexity as a continuum	Increasing openness, uncertainty, competing perspectives	Movement from retrieval toward dialogue and evaluation	Increased learner responsibility for reasoning

The findings indicate that task characteristics did more than determine the form of the final writing product; they shaped the cognitive process through which that product was created. Relatively closed tasks allowed students to treat ChatGPT primarily as a language-production service, whereas open-ended tasks containing competing claims, uncertainty, or evidence requirements encouraged learners to interact with the tool as a source to be questioned. Thus, variation in reasoning was associated not simply with ChatGPT use itself but with what students were required to accomplish after receiving an AI response.

The results also suggest that the strongest reasoning occurred when task design made students responsible for evaluating and defending decisions. Evidence checking, comparison, and revision requirements transformed AI output from a ready-made answer into an object of analysis. However, the presence of a complex task alone was not sufficient. Reasoning was strongest when complexity was combined with explicit accountability, such as requiring learners to verify claims, document revisions, or explain why an AI suggestion was accepted or rejected. This finding supports the interpretation that reasoning-rich TBLT depends on the interaction between task demand and learner accountability.

Which task conditions and forms of teacher mediation promote critical engagement with ChatGPT rather than passive acceptance of AI-generated output?

The second research question examined the task conditions and forms of teacher mediation that encouraged learners to critically evaluate ChatGPT output. The first major theme was the importance of explicit verification requirements. Participants reported being more critical when teachers required them to validate AI-generated claims against scholarly or course-based sources rather than simply use ChatGPT as a reference point. P2 explained, *"If the teacher says every important claim must be checked from another source, I automatically become more careful with ChatGPT."* P7 stated, *"My teacher told us that AI can sound confident even when it is wrong, so now I verify facts before putting them in my assignment."* P10 commented, *"When we must attach the real source, I cannot use an AI answer unless I*

find support for it." P16 similarly said, *"The rule that we need two academic sources makes me compare instead of just accepting the ChatGPT explanation."* P21 added, *"I trust ChatGPT less when I know the teacher will ask where the information came from."* These responses indicate that verification requirements created an external accountability structure that encouraged learners to evaluate credibility rather than equate linguistic confidence with factual reliability.

A second influential condition involved mandatory justification of AI-related decisions. Students became more reflective when teachers asked them to explain how ChatGPT had influenced their writing and why particular suggestions were retained, modified, or discarded. P3 reported, *"My teacher asked why I accepted one suggestion but rejected another. That question made me realize I needed my own reason."* P6 noted, *"When we have to write a short explanation of our changes, copying becomes difficult because we need to understand every important revision."* P11 explained, *"I used to accept most corrections, but now I ask whether the change fits my argument because I have to justify it."* P14 stated, *"The reflection part is useful because it shows whether I really understand the paragraph or only follow ChatGPT."* P19 added, *"If I know the teacher will ask me to defend my final version, I read the AI response more carefully."* These excerpts suggest that justification requirements supported metacognitive monitoring by making learners accountable not only for the final product but also for the reasoning process behind it.

The third theme concerned the teacher's role in modeling critical AI use. Participants valued teacher demonstrations showing how AI-generated output could be inaccurate, overly general, biased, or rhetorically inappropriate. Such demonstrations appeared to challenge students' assumptions that fluent answers were necessarily reliable. P1 explained, *"The teacher showed us one ChatGPT answer that looked excellent, but two references were false. After that I stopped thinking good English means correct information."* P5 stated, *"Our teacher compared an AI paragraph with a journal article and asked us to find the weaknesses. That taught me how to check it."* P9 noted, *"Before, I only learned how to make prompts. The teacher taught us that questioning the answer is more important than making a complicated prompt."* P13 remarked, *"When the teacher intentionally gave us a wrong AI response to analyze, I became more suspicious in a good way."* P18 similarly stated, *"Seeing the teacher reject an AI answer because the evidence was weak showed me that ChatGPT should not be treated as an expert."* Teacher modeling therefore appeared to develop critical AI literacy by making evaluation practices visible rather than assuming that students would discover them independently.

Another important theme was teacher questioning and dialogic scaffolding. Learners reported greater cognitive engagement when teachers responded to AI-assisted writing with questions such as "Why do you believe this claim?", "What evidence supports it?", "What alternative explanation exists?", or "Would you keep this sentence if ChatGPT had not generated it?" P4 stated, *"When the teacher asks me why I used a statement, I have to go back and understand it."* P8 explained, *"My teacher usually asks, 'What is your evidence?' That makes me check the response again."* P12 said, *"Instead of telling us that ChatGPT is wrong, the teacher asks us to find the problem ourselves."* P15 observed, *"Questions from the teacher make me compare my opinion with the AI opinion."* P20 added, *"If the teacher only says correct or incorrect, I do not think much, but when she asks me to defend the paragraph, I think more deeply."* These excerpts indicate that teacher questioning was particularly effective when it preserved learner responsibility rather than simply replacing AI authority with teacher authority.

The analysis further suggested that structured comparison and peer discussion promoted critical engagement. Students reported benefiting from tasks in which AI-generated responses were compared with peer-produced texts, scholarly excerpts, or alternative AI outputs. P2

stated, *"When we compared three ChatGPT answers, I saw that the answers could be different even for the same topic."* P6 explained, *"Peer discussion helped because my classmate noticed a problem in the AI answer that I did not see."* P10 commented, *"Comparing our own paragraph with the AI paragraph helped me see that the AI version was fluent but less specific."* P17 stated, *"If we discuss the output in a group, we do not accept it so quickly because everyone asks different questions."* P21 added, *"The comparison task showed me that sometimes my original argument was better than ChatGPT's."* Such activities appeared to reduce the perceived authority of AI by presenting its output as one possible contribution among several rather than the final standard against which all writing should be judged.

Participants emphasized that teacher mediation was most effective when it established clear boundaries without completely prohibiting AI. Overly restrictive approaches were perceived as less useful because they encouraged avoidance rather than responsible engagement. P3 explained, *"If the teacher only says 'do not use ChatGPT,' we do not learn how to use it correctly."* P7 stated, *"I prefer when the teacher tells us what AI can help with and what we must do ourselves."* P11 commented, *"Clear rules are useful because I know I can use ChatGPT for ideas but not submit its paragraph directly."* P16 said, *"The teacher allowed AI but required us to show our prompts and changes, so I became more responsible."* P18 added, *"I think controlled use is better than banning it because we need to learn how to judge the output."* These accounts suggest that effective teacher mediation involved guided autonomy: allowing AI use while preserving students' responsibility for evidence, reasoning, and authorship.

Table 2. Summary of Task Conditions and Forms of Teacher Mediation Promote Critical Engagement with ChatGPT

Theme	Task condition / teacher mediation	Learner response	Contribution to critical engagement
Verification requirements	Cross-checking AI claims with credible sources	Fact-checking, source searching, citation verification	Reduces uncritical trust in AI
Justification requirements	Explaining acceptance, rejection, or revision of AI output	Reflection on decisions and reasoning	Strengthens metacognition and accountability
Teacher modeling	Demonstrating inaccurate or weak AI responses	Increased skepticism and evaluative awareness	Develops critical AI literacy
Dialogic teacher questioning	Asking for evidence, rationale, and alternatives	Reconsideration and defense of claims	Sustains learner responsibility for reasoning
Structured comparison	Comparing AI, peer, and scholarly texts	Identification of differences and weaknesses	Reduces perception of AI as final authority
Peer discussion	Collaborative critique of AI output	Multiple perspectives and shared evaluation	Promotes deeper scrutiny
Guided autonomy	Clear rules allowing responsible AI use	More transparent and purposeful AI interaction	Balances assistance with independent thinking

The findings suggest that critical engagement did not emerge simply because students had access to ChatGPT. Instead, it developed when task conditions created explicit reasons to question, verify, compare, and justify AI-generated content. Verification requirements and justification tasks were particularly important because they directly interrupted passive acceptance. These conditions transformed the learner from a recipient of AI-produced text into an evaluator responsible for establishing whether an output was credible, relevant, and rhetorically appropriate.

Teacher mediation was also most productive when it supported learner judgment rather than substituting one authority for another. Modeling, probing questions, comparative activities, and transparent AI-use guidelines appeared to cultivate a form of guided autonomy in which students were permitted to use ChatGPT but remained responsible for evidential and rhetorical decisions. The findings therefore suggest that the teacher's emerging role in ChatGPT-assisted TBLT is not primarily to police AI use, but to design conditions under which learners must critically engage with it. In this sense, effective mediation combines task structure, AI literacy, reflective accountability, and dialogue to ensure that ChatGPT functions as a reasoning scaffold rather than a replacement for learner thinking.

Discussion

The present study examined how TBLT task characteristics shape students' reasoning and patterns of interaction with ChatGPT during academic writing, as well as the task conditions and teacher mediation strategies that promote critical rather than passive engagement with AI-generated output. The findings show that the educational value of ChatGPT is not determined by access to the technology alone. Instead, students' cognitive engagement depended substantially on what the task required them to do before, during, and after interacting with the AI. Descriptive, summarizing, grammar-correction, and reproduction-oriented activities tended to elicit transactional use, whereas argumentative, comparative, problem-based, source-based, and revision-oriented tasks promoted more iterative questioning, verification, comparison, and justification. This pattern supports the study's initial assumption that task architecture distributes cognitive responsibility between the learner and AI rather than ChatGPT automatically promoting or diminishing reasoning.

The tendency for simpler production-oriented tasks to encourage transactional AI use is consistent with previous concerns regarding cognitive outsourcing in GenAI-assisted learning. Adamakis and Rachiotis (2025) argue that although AI can improve efficiency and provide immediate assistance, pedagogically weak integration may reduce opportunities for independent reasoning and critical evaluation. Similarly, Álvarez-Jiménez et al. (2026) highlight the tension between performance enhancement and learner autonomy when generative AI is used in higher education. The current findings extend these arguments by identifying a more specific mechanism through which cognitive outsourcing may occur: task requirements that reward completion, linguistic fluency, or rapid production without demanding explanation or verification. In such contexts, students may have little incentive to challenge fluent AI-generated text. Thus, the problem does not appear to be efficiency itself, but efficiency without epistemic accountability.

Conversely, the stronger reasoning observed in argumentative, comparative, and problem-oriented tasks corresponds with research emphasizing the value of cognitively demanding interactions with generative AI. Ardyansyah et al. (2024) reported that students can use GPT productively when they actively engage with generated explanations rather than treat them as unquestionable answers. Wang et al. (2025) similarly found that restructuring tutor and peer roles around ChatGPT-supported problem-based learning could foster critical thinking when learners remained responsible for evaluating and negotiating ideas. The present study reinforces these findings in EFL academic writing by showing that argumentative tasks prompted learners to request counterarguments, consider alternative viewpoints, ask follow-up questions, and evaluate the comparative strength of different positions. This suggests that productive AI use is more likely when task completion cannot be achieved through a single generated response.

A particularly important finding concerns evidence constraints and source verification. Participants reported greater caution when assignments required academic sources, citation checking, or comparison between ChatGPT responses and independently identifiable evidence. This result aligns with Adamakis and Rachiotis (2025), who emphasize AI literacy, verification,

and pedagogical integrity as central conditions for responsible GenAI integration. It also supports Naznin et al. (2025), whose systematic review identifies both the academic-writing potential of ChatGPT and the need for critical evaluation of generated information. Across disciplines, similar cautions have been raised concerning the reliability and disciplinary appropriateness of AI-generated content (Nieto-Márquez et al., 2026; Su et al., 2026; Trčková et al., 2024). The present findings add an important pedagogical dimension: source verification should not be treated merely as an external rule about accuracy. It can itself function as a reasoning mechanism because it obliges students to compare, evaluate, and decide whether an AI-generated claim deserves inclusion.

The finding that justification and revision requirements promoted metacognitive engagement further develops this argument. Participants became more conscious of their decisions when they were required to explain why they accepted, modified, or rejected ChatGPT suggestions. This pattern is consistent with research suggesting that AI-supported learning is most productive when learners retain responsibility for evaluation and decision making (Faisal, 2026; Li et al., 2026). It also aligns with the study's TBLT framework, in which tasks involving competing perspectives, evidence constraints, and iterative decisions create greater opportunities for deeper cognitive processing. Rather than positioning revision as a purely linguistic stage, the results suggest that revision can become a metacognitive task when students must articulate the rationale behind changes. This extends previous research by showing that the pedagogically significant unit is not simply the prompt submitted to ChatGPT, but the sequence of decisions surrounding the generated output.

The second research question further demonstrated that critical engagement was strengthened by explicit teacher mediation. Verification requirements, justification prompts, modeling, dialogic questioning, structured comparison, peer discussion, and guided autonomy all contributed to reducing passive acceptance. These findings are consistent with the emerging view that teachers' roles in AI-supported learning are shifting from information providers toward instructional designers and facilitators of AI literacy. Adamakis and Rachiotis (2025), Arroyo-Bello et al. (2025), and Naznin et al. (2025) emphasize the need for educators to support responsible and critical AI use rather than assume that students naturally possess the skills required to evaluate generated content. The current study provides more specific evidence of how such mediation can operate in practice. Teacher prompts such as asking learners to justify a claim, identify supporting evidence, or consider alternative explanations appeared to preserve learner responsibility rather than transfer authority from ChatGPT directly to the teacher.

Teacher modeling was also important because learners reported becoming more skeptical when teachers demonstrated errors, weak evidence, fabricated references, or overly generalized AI responses. This is particularly significant because generative AI can produce fluent and rhetorically persuasive language even when its content requires verification. The manuscript itself highlights the risk that learners may interpret linguistic fluency as evidence of factual reliability. The findings therefore support the need for critical AI literacy that goes beyond prompt engineering. Teaching students how to obtain better prompts is insufficient if they are not also taught how to question the resulting responses. This interpretation parallels Bollaert and Coppin's (2024) emphasis on explainability and intelligibility in human-AI interaction and Coman et al.'s (2026) call for responsible, human-centered AI integration.

The effectiveness of structured comparison and peer discussion further suggests that reasoning in AI-assisted writing remains socially mediated. When students compared AI outputs with peer texts, scholarly sources, or alternative responses, they became less likely to treat ChatGPT as a singular authority. This result resonates with Wang et al. (2025), whose work on tutor-peer role reconfiguration demonstrates that the value of AI-supported learning can emerge through interaction among multiple human and technological actors. It also complements TBLT principles by highlighting the importance of purposeful interaction,

negotiation, and comparison. In this respect, ChatGPT need not replace collaborative learning; properly designed tasks may actually turn AI-generated content into an object around which collaboration and critique are organized.

One finding that deserves qualification is that complex tasks did not automatically lead to critical AI use. Some students remained dependent on ChatGPT even when completing more cognitively demanding assignments, while others independently verified information in relatively simple tasks. This nuance is important because it contradicts any deterministic interpretation that certain task labels inherently produce reasoning. Instead, the findings suggest an interaction among task complexity, learner disposition, AI literacy, previous experience, and teacher mediation. Fu and Wang (2026) emphasize the importance of task-based instructional design, but the current findings indicate that the cognitive potential of a task may remain unrealized unless accountability mechanisms are built into its implementation. Thus, reasoning-rich design should be understood not simply as choosing argumentative rather than descriptive writing, but as constructing tasks that require evidence, comparison, uncertainty management, revision, and justification.

The study consequently extends previous research by proposing a more differentiated account of ChatGPT-assisted academic writing. Existing studies have commonly examined usefulness, perceptions, performance, ethical risks, or AI literacy (Harunasari, 2023; Tian, 2024; Naznin et al., 2025). The present findings shift attention toward the architecture of learner-AI interaction. The key distinction is not AI-assisted versus non-AI-assisted writing, but whether the instructional environment preserves the learner's epistemic responsibility. Reasoning-rich TBLT appears to do so by requiring students to interrogate claims, compare alternatives, verify evidence, revise deliberately, and account for their decisions.

Practically, the findings suggest that educators should avoid assigning open-ended instructions such as "use ChatGPT to write an essay" without embedding reasoning requirements. More productive task sequences would require learners to formulate an initial position independently, consult ChatGPT for counterarguments or feedback, verify its claims using credible sources, compare alternatives, revise their writing, and justify final decisions. Assessment criteria should likewise reward evidence quality, reasoning, transparency, and revision rationale rather than only linguistic accuracy or surface fluency. Such an approach is consistent with the manuscript's argument that teachers should function as designers of cognitive conditions and facilitators of AI literacy. In this sense, the contribution of ChatGPT to academic writing depends less on what the technology can generate than on what the task requires learners to think about, question, defend, and ultimately decide for themselves.

CONCLUSION

This study demonstrates that the role of ChatGPT in academic writing is strongly shaped by the architecture of TBLT tasks. Among the 21 EFL learners participating in the qualitative investigation, relatively closed, descriptive, summarizing, grammar-focused, and reproduction-oriented tasks were associated with more transactional uses of ChatGPT, including requesting complete drafts, paraphrases, and linguistic corrections. In contrast, argumentative, comparative, problem-based, source-based, and revision-oriented tasks encouraged more dialogic patterns characterized by follow-up questioning, counterargument generation, evidence checking, source comparison, iterative prompting, and justification. The findings therefore challenge a simplistic interpretation of ChatGPT as either inherently beneficial or inherently detrimental to reasoning. Instead, they indicate that AI-assisted academic writing operates along a continuum from reasoning-poor to reasoning-rich engagement. The decisive issue is the extent to which task conditions preserve learners' responsibility for evaluating information, judging evidence, making rhetorical decisions, and explaining the rationale behind revisions. In this respect, evidence verification, comparison, and justification emerged

as particularly important mechanisms for transforming ChatGPT from a source of ready-made text into a resource for cognitive engagement.

The findings also demonstrate that teacher mediation is integral to sustaining critical engagement with AI. Learners responded more critically when teachers required independent verification, demanded explanations for accepting or rejecting AI suggestions, modeled the evaluation of problematic AI outputs, used dialogic questioning, facilitated peer comparison, and established clear expectations for responsible AI use. These practices suggest that teachers' roles in AI-mediated TBLT are shifting from controlling access to information toward designing cognitive conditions in which students remain intellectually accountable. However, the study's qualitative design and purposively selected sample of 21 EFL learners from private educational settings in Bangladesh limit statistical generalization, and the reliance on interview accounts may be affected by recall and social-desirability biases. Future studies should therefore triangulate interview findings with students' actual ChatGPT interaction logs, writing drafts, revision histories, classroom observations, and longitudinal evidence. Experimental or mixed-method studies could also compare reasoning-rich and reasoning-poor task conditions more directly. Nevertheless, the present study contributes to research on GenAI-assisted language education by showing that the central pedagogical challenge is not simply whether students use ChatGPT, but whether instructional design requires them to remain active evaluators, decision makers, and authors throughout the writing process.

ACKNOWLEDGEMENT

The authors would like to express their appreciation to the EFL learners who voluntarily participated in this study and shared their experiences of using ChatGPT in academic writing. Their reflections and perspectives provided the empirical basis for understanding how different task conditions influence reasoning and learner-AI interaction. The authors also acknowledge the academic colleagues and experts who provided feedback on the interview protocol and methodological procedures, which contributed to improving the clarity and relevance of the research instrument. The authors are also grateful to the participating educational institutions for facilitating the research process and supporting access to the study participants. Any interpretations, conclusions, and remaining limitations presented in this article are solely the responsibility of the authors.

CONFLICT OF INTEREST

The authors declare that they have no financial, professional, personal, institutional, or other competing interests that could have influenced the design, conduct, analysis, interpretation, or reporting of this study. No relationships, activities, or circumstances existed that could reasonably be perceived as affecting the objectivity, integrity, or independence of the research presented herein.

FUNDING

This research received no specific grant or financial support from any public, commercial, governmental, or not-for-profit funding agency. All activities related to the study, including data collection, analysis, manuscript preparation, and publication, were independently supported by the authors without external financial involvement or influence on the research process.

INFORMED CONSENT STATEMENT

Participation in this study was entirely voluntary. Before taking part, all participants were informed about the purpose, procedures, potential risks, and benefits of the research. Participants were assured that their identities and personal information would remain confidential and that all information provided would be used solely for research purposes. They were also informed of their right to withdraw from the study at any time without penalty or

loss of benefits. By agreeing to participate and continuing with the study procedures, all participants provided their informed consent under the conditions described.

DATA AVAILABILITY STATEMENT

The data used in this study are not publicly available because of privacy, confidentiality, and ethical obligations intended to protect the participants. This restriction ensures compliance with applicable ethical research standards and data protection requirements. Researchers or other interested parties may request access to the dataset for purposes such as validation or further analysis. Such requests will be considered on a case-by-case basis and must be deemed reasonable and appropriate. Data will only be shared after approval from the relevant institutional ethics review board and in accordance with ethical guidelines, data protection regulations, and the terms of participant consent.

DECLARATION OF USING AI TOOLS

Generative artificial intelligence tools were used during the preparation of this manuscript to support language editing, improve clarity and organization, and assist in refining selected sections of the text. All AI-assisted content was critically reviewed, verified, and revised by the authors. The authors take full responsibility for the accuracy, originality, interpretation of the findings, citations, and final content of the article.

REFERENCES

Adamakis, M., & Rachiotis, T. (2025). Artificial intelligence in higher education: A state-of-the-art overview of pedagogical integrity, artificial intelligence literacy, and policy integration. *Encyclopedia, 5*(4), 180. https://doi.org/10.3390/encyclopedia5040180

Álvarez-Jiménez, M., Pomagualli, A. G. C., Gamboa, D., & Oscullo, D. S. P. (2026). Generative AI in technology-oriented higher education: A systematized review and survey on students' perceptions of performance, autonomy, and ethical implications. *Computers, 15*(7), 466. https://doi.org/10.3390/computers15070466

Ardyansyah, A., Yuwono, A. B., Rahayu, S., Alsulami, N. M., & Sulistina, O. (2024). Students' perspectives on the application of a generative pre-trained transformer (GPT) in chemistry learning: A case study in Indonesia. *Journal of Chemical Education, 101*(9), 3666–3675. https://doi.org/10.1021/acs.jchemed.4c00220

Arroyo-Bello, E., González-Toledo, B., Abad-Valle, J., Rodríguez-Gómez, P., & Ramon, M. (2025). Exploring the challenges and opportunities of ChatGPT in university teaching for health sciences: A narrative review. *Cureus*. https://doi.org/10.7759/cureus.94259

Balacheff, N. (2024). Situaciones para el aprendizaje de la prueba en matemática: Estado de la investigación y cuestiones abiertas. *Recherches en Didactique des Mathématiques*. https://doi.org/10.46298/rdm.12905

Bollaert, M., & Coppin, G. (2024). Vers l'XAI sémantique – la troisième vague de l'intelligence artificielle explicable. *Ingénierie Cognitique, 7*(2), 27–50. https://doi.org/10.21494/iste.op.2024.1147

Coman, C., Gherheș, V., Bucs, A., Coman, E., Cioca, V.-R., Bódi, D.-C., & Șerbănescu, R.-M. (2026). Navigating opportunities and challenges of generative AI in higher education: Towards responsible, equitable, and human-centered integration. *Frontiers in Artificial Intelligence, 9*. https://doi.org/10.3389/frai.2026.1750978

Decelles, S. (2023). Intelligence artificielle et technologie en éducation. *Apprendre et Enseigner Aujourd'hui: Revue du Conseil Pédagogique Interdisciplinaire du Québec, 13*(1), 25–28. https://doi.org/10.7202/1107539ar

Faisal, E. (2026). Complementary AI in higher education: Behavioral, cognitive, and ethical implications of ChatGPT and DeepSeek. *Frontiers in Psychology, 17*. https://doi.org/10.3389/fpsyg.2026.1699114

Fu, Y., & Wang, H. (2026). Task-based teaching method in distance English teaching. *International Journal of Web-Based Learning and Teaching Technologies, 21*(1), 1–16. https://doi.org/10.4018/ijwltt.403093

Harunasari, S. Y. (2023). Examining the effectiveness of AI-integrated approach in EFL writing: A case of ChatGPT. *International Journal of Progressive Sciences and Technologies, 39*(2), 357. https://doi.org/10.52155/ijpsat.v39.2.5516

Houdé, O. (2023). Le rôle clé de l'inhibition cognitive dans l'éducation du cerveau. *Revue Suisse de Pédagogie Spécialisée, 13*(04), 29–35. https://doi.org/10.57161/r2023-04-05

Li, Y., Cheng, W., Chen, Z., He, Y., Liu, D., & Chen, Z. (2026). The improvement or hindrance of the new generation of artificial intelligence. *International Journal of Information Technologies and Systems Approach, 19*(1), 1–24. https://doi.org/10.4018/ijitsa.406761

Naznin, K., Mahmud, A. A., & Chua, C. (2025). ChatGPT integration in higher education for personalized learning, academic writing, and coding tasks: A systematic review. *Computers, 14*(2), 53. https://doi.org/10.3390/computers14020053

Nieto-Márquez, N. L., Madrigal-Cerezo, R., Ramos-Marcos, L., Rueda-Díaz, N., Martín, T. G., & López-Muñoz, F. (2026). From AI tool use to instructional design: Development and validation of the AID-CTQ in higher education. *Education Sciences, 16*(6), 982. https://doi.org/10.3390/educsci16060982

Ostos, R., Félix, V. G., Mena, L. J., Toral-Cruz, H., Ochoa-Brust, A., González-Potes, A., Félix, R. A., Pacheco, J. C. R., Flores, V., & Martínez-Peláez, R. (2026). Pedagogical transformation using large language models in a cybersecurity course. *AI, 7*(1), 25. https://doi.org/10.3390/ai7010025

Su, C. S., Hsu, C. C., & Campos, J. G. (2026). ChatGPT as a resource to strengthen mathematics teaching and learning: A systematic review. *Journal of Computer Assisted Learning, 42*(3). https://doi.org/10.1002/jcal.70241

Tian, C. (2024). The influence of generative AI technologies on academic writing in EFL education. *Journal of Education, Humanities and Social Sciences, 28*, 575–584. https://doi.org/10.54097/xv0rbq11

Trčková, K., Maršálek, R., & Teplá, M. (2024). Why do fishes die? Lab assignment created using ChatGPT. *Journal of Chemical Education, 101*(8), 3171–3178. https://doi.org/10.1021/acs.jchemed.4c00130

Wang, R., Liu, C., & Yi, F. (2025). Reconfiguring tutor–peer roles to foster critical thinking in medical PBL assisted by tutor-managed ChatGPT. *Medical Education, 60*(8), 894–907. https://doi.org/10.1111/medu.70100

Zhang, B. (2024). Exploring pedagogical ideologies and strategies for college English writing instruction from a generative artificial intelligence perspective. *Journal of Contemporary Educational Research, 8*(1), 173–178. https://doi.org/10.26689/jcer.v8i1.5951