

Hyland, K. (2026). Rethinking AI in writing pedagogy: From fluency to rhetorical competence. Discourse on Culture.

Rethinking AI in writing pedagogy: From fluency to rhetorical competence

Ken Hyland

Highlights

- Five studies comparing ChatGPT and student essays reveal that AI excels in fluency and coherence but underperforms in stance, engagement, and rhetorical depth.
- I position ChatGPT as a scaffold, highlighting its strengths while recognising its limitations.
- I Propose tasks such as contrastive analysis, revision of AI drafts, and prompt design to foster rhetorical awareness and critical AI literacy.
- I call for recalibrated rubrics and process-oriented practices to safeguard construct validity to assess student learning accurately.
- I emphasise transparency in AI use and the need to guide students to reflect on authorship, agency, and integrity

Abstract

Generative AI tools such as ChatGPT are reshaping academic writing in higher education, offering students unprecedented fluency, coherence, and accessibility. Yet these seductive characteristics, together with serious rhetorical limitations, raise concerns about authorship, integrity, and the development of critical thinking. The similarities and differences between AI generated texts and human writing, and the implications of these for writing instruction, however, remain underexplored. This study aims to bridge this gap by synthesizing five studies conducted by Jiang and Hyland, which systematically compare ChatGPT-generated essays with those authored by British undergraduates. The studies reveal a consistent pattern: while ChatGPT excels in producing structurally polished texts, it underperforms in rhetorical depth, interpersonal engagement, and authorial presence. Building on these findings, the paper proposes a framework for integrating AI into writing pedagogy. This positions ChatGPT as a supportive tutor rather than a proxy author, stressing both its advantages and limitations. It offers practical pathways for educators and underscores the importance of cultivating critical AI literacy, enabling students to navigate new writing contexts where human creativity and algorithmic fluency intersect.

Keywords

ChatGPT; Argumentative essays; AI literacy; written interaction; Writing pedagogy

1 Introduction

The rapid emergence of generative artificial intelligence (AI) tools such as ChatGPT has provoked widespread debate in higher education. On the one hand, these systems can produce fluent, coherent, and lexically sophisticated texts at remarkable speed, offering students accessible models of organisation and style. On the other hand, their rhetorical limitations - weak stance-taking, limited engagement, and formulaic cohesion - raise concerns about plagiarism, diminished critical thinking, and the erosion of academic literacy. The Guardian newspaper, for example, recently reported a 'freefall in public confidence' in US higher education with the potential for AI to increasingly substitute independent thought. Instructors at US Ivy League universities in the article believed that reliance on artificial intelligence was 'fundamentally antithetical to the development of human intelligence' (Speri, 2026).

As universities grapple with questions of integrity, pedagogy, and assessment, however, much of the discussion of AI in writing seems overly instrumental, preoccupied with detection, regulation, or efficiency (Gao et al., 2025). What risks being overlooked is the extent to which AI forces us to confront assumptions about the relationship of writing to learning and disciplinary participation. The main challenge, it seems, lies not in resisting AI outright but in understanding its capacities and limitations, and in integrating it responsibly into classrooms and other academic contexts.

This paper therefore examines AI in academic writing, focusing on student contexts and exploring how the distinctive features of AI generated texts can inform how we use these tools to develop literacy skills. It draws on five corpus-based studies conducted by Jiang and Hyland in the last 3 years, which systematically compare ChatGPT-generated essays with those written by British undergraduates. These studies analyse key rhetorical dimensions - stance and voice, cohesion and coherence, engagement and interaction, argumentation and evidence, and genre awareness to provide a comprehensive empirical base for evaluating AI's rhetorical performance. Together, they reveal a consistent pattern: while ChatGPT excels in producing structurally clear and fluent prose, it lacks rhetorical penetration and interpersonal engagement and so fails to produce effective academic writing.

Building on these findings, I outline a conceptual framework for integrating AI into academic writing pedagogy. Rather than treating ChatGPT as a automated writer or a threat to integrity, the framework positions it as a scaffold for teaching rhetorical awareness and fostering critical AI literacy. By juxtaposing AI outputs with student texts, revising drafts to add rhetorical features, and embedding process-oriented assessment practices, educators can harness AI's strengths while tackling its limitations. The framework thus offers a roadmap for using AI creatively and critically in higher education, ensuring that students develop the metalinguistic and rhetorical skills necessary to navigate hybrid writing environments where human creativity and algorithmic fluency intersect.

2 Friend or foe? AI in Higher Education

Since its release in late 2022, ChatGPT has rapidly become one of the most debated tools in higher education, particularly in the domain of academic writing. Scholars have noted its ability to generate grammatically accurate, coherent, and lexically sophisticated texts, raising both opportunities and challenges for pedagogy and assessment. For many learners, particularly those writing in a second language, AI offers accessible models to assist learners in organising their texts and polishing their grammar, functioning as a supportive assistant for developing academic literacy (Barrot, 2023; Deep & Chen, 2025). Studies also suggest that AI can support productivity by streamlining drafting, suggesting revisions, and assisting with citation formatting, thereby reducing the mechanical burdens of writing (Imran & Almusharraf, 2023). In this sense, then, ChatGPT acts as a helpful prompt or “writing assistant” capable of modelling academic genres and helping students visualise the rhetorical structures they might otherwise struggle to produce.

Yet alongside these opportunities, scholars have raised significant concerns. There is a risk, for example, that AI-assistance remains superficial, enabling students to meet formal requirements without gaining deeper control over disciplinary meaning-making (Li & Wu, 2025; Siddiqui, et al, 2025). One recurring theme is the apparent rhetorical shallowness of AI-generated texts, so while ChatGPT produces structurally polished essays, it consistently underperforms in authorial presence and interpersonal dimensions. This absence of rhetorical complexity not only produces emotionally bland texts but encourages students to rely on surface fluency rather than developing the critical thinking and persuasive skills that we value in academic writing. Ethical issues compound these concerns: questions of plagiarism,

transparency, and authorship remain unresolved, with institutions struggling to balance innovation with integrity (Das & Chen, 2025). These issues are compounded, and even stimulated, by assessment practices which use rubrics that privilege clarity and coherence and which, as a result, inadvertently reward AI outputs over student learning. This has prompted calls to recalibrate descriptors to emphasise reasoning, adaptability, and responsiveness to audience (e.g. Hyland, 2026).

Unfortunately, much of the discussion surrounding AI reduces a complex pedagogical issue to a question of rule compliance. From the perspective of L2 writing, the more consequential issue concerns authorship and responsibility (Li & Wu, 2025). Writing is not simply the production of text that meet external criteria; it is an act through which writers take responsibility for meanings, positions, and claims and which can engage and persuade readers. L2 writing specialists, then, are beginning to recognise that the challenge is to harness AI in ways that genuinely expand participation rather than merely smooth over inequality at the level of text. What this means is that whether AI works to narrow or widen existing proficiency and literacy gaps will depend less on the technology itself than on how we make use of it: on the pedagogical and institutional contexts in which it is embedded.

Recent scholarship has begun to explore how these challenges might be addressed pedagogically. Pack and Maloney (2023), for example, argue for teaching “prompt literacy,” helping students understand how input shapes output and encouraging them to reflect critically on AI’s rhetorical limitations. Similarly, Imran and Almusharraf’s (2023) systematic review concludes that while ChatGPT can enhance academic writing, sustainable integration requires embedding critical AI literacy into the curriculum. This involves not only recognising what AI can do well but also identifying what it cannot: constructing persuasive arguments, engaging readers, and signalling authorial stance. By positioning AI outputs as objects of analysis and revision, educators may be able to help students develop metalinguistic awareness and rhetorical competence.

In sum, the current literature exhibits an overall ambivalence towards AI in higher education. On one hand, ChatGPT offers powerful tools for modelling fluency and supporting students in the early stages of writing, while on the other, its rhetorical limitations and ethical risks demand careful pedagogical responses. This tension provided the rationale for the studies conducted by Kevin Jiang and myself examining AI’s role in academic writing, drawing on

parallel corpora to determine where its strengths and weaknesses lay. The results of these studies can help develop frameworks that integrate AI creatively and critically into university classrooms.

3 The need for comparison studies

The increasing sophistication of AI-generated texts raises pressing issues about academic integrity and potential effects on critical thinking skills. If students can easily access AI-generated essays that are indistinguishable from human-authored work, it not only undermines academic ethics (Barrot, 2023) but might impede the development of students' analytical and argumentative skills, which are crucial for their academic and professional growth (Adeshola & Adepoju, 2023; Kim et al, 2025). The issue is becoming critical as increasing numbers of students turn to AI to for assistance with their writing. A global survey by the Digital Education Council (Kelly, 2024), for example, found that 86% of students use AI in their studies, with 54% using it weekly, a rate of adoption mirrored in the Chegg survey of nearly 12000 undergraduates across 15 countries (Chegg, 2025).

Recent studies have therefore sought to identify features that can distinguish between AI-generated and human-authored texts. For example, AI texts seem to exhibit higher informational density with more noun and preposition-based structures (Herbold et al., 2023; Markey et al., 2024), while human-authored texts exhibit a higher frequency of clause-related bundles and more pronounced use of epistemic stance markers (Jiang & Hyland, 2024). Yang et al. (2024) also found that ChatGPT-generated texts rely more on repetitive connective signals and constant thematic progression, suggesting less careful planning for text development, which leads to redundancy and a simplistic list-like structure. This can be attributed to the fact that the model leverages statistical analyses, large data collections, and machine learning to determine the likelihood of text sequencing rather than an understanding of language (Sarrion, 2023; Wolfram, 2023).

Texts created by ChatGPT's statistical algorithms therefore appear to reinforce a view of writing where it might be seen as a fixed structure following a system of grammatical rules (Vee, 2023). It tends to objectivize academic writing and so "fails to take account of the beliefs and knowledge writers assume readers will draw on in reading their texts" (Hyland, 2026, p. 6). The critical analyses of ChatGPT generated texts can therefore be helpful in providing linguistic insights into the limitations of the LLM in terms of how it understands

textual interaction. This is particularly the case with argumentative writing, where writers seek to establish a coherent disciplinary context by drawing on their background knowledge and logical reasoning, and attempt to persuade readers to accept their position. Success is therefore a rhetorical outcome of what authors include in the text and the stance they take towards it (Hyland & Jiang, 2016; Aull & Lancaster, 2014; Lee, & Deakin, 2016).

Academic literacy is not simply a matter of producing grammatically accurate or structurally coherent texts; it demands an awareness of appropriate genre conventions and how these might be manipulated as well as the understanding and use of interactive and interpersonal aspects of argument that involve writers taking a stance and seeking to imagine and persuasively engage an audience. These are precisely the areas where generative AI tools such as ChatGPT remain underexplored. The corpus-based studies conducted by Jiang and Hyland provide a comprehensive foundation for addressing this gap, offering detailed analyses of rhetorical features across student and AI essays¹.

Each of these studies examines a distinct rhetorical dimension of writing, including stance and voice, cohesion and coherence, engagement and interaction, argumentation and evidence, and genre awareness. Together, they offer a systematic comparison of ChatGPT-generated essays and student-authored texts across the core dimensions of academic literacy. The findings reveal consistent patterns: while ChatGPT excels in fluency and coherence, it underperforms in rhetorical depth, interpersonal engagement, and authorial presence. This evidence base is critical for moving beyond anecdotal debates about AI in education toward a grounded understanding of its rhetorical capacities. The rationale for focusing on these five studies is therefore twofold. First, they provide robust empirical evidence of how AI-generated texts differ from student writing in ways that are pedagogically significant. Second, they collectively point toward the need for a framework that integrates AI into academic writing instruction to help model rhetorical awareness and foster critical AI literacy. By situating AI within this framework, the present paper contributes to ongoing debates about the future of academic writing, offering a pathway for educators to harness AI's advantages while addressing its limitations.

4 The data and analyses

¹ These studies are listed in the references as Jiang and Hyland 2024, 2025a, 2025b, 2026, and in press.

All five studies are based on the same two balanced corpora of argumentative essays designed to enable systematic comparison between student and AI-generated texts. Table 1 shows the characteristics of the corpora in terms of texts, words (tokens), and different words (types).

Table 1 Corpus characteristics

Corpus	No. of texts	Tokens	Types
British student essays	145	78,060	6,975
ChatGPT essays	145	72,819	7,164

The first corpus comprised 145 essays written by British university students aged 18–22, drawn from the Louvain Corpus of Native English Essays (LOCNESS) (Granger, 2013, 2017). These texts were produced in untimed, coursework-like settings, reflecting a developmental stage in academic writing where rhetorical awareness and authorial identity are emerging. The topics - Britain’s relationship to the EU, transport, boxing, and the lottery - are typical prompts in English for General Academic Purposes (EGAP) courses, encouraging students to take a stance and provide supporting arguments while also reflecting specific cultural and institutional perspectives on academic argumentation.

The second corpus mirrors the student one with 145 essays generated by ChatGPT in response to prompts following the student tasks as closely as possible. Each prompt specified a role (“competent university-student writer”), a realistic academic context, and a persuasive writing style. Prompts were refined to reduce ambiguity but deliberately avoided extensive optimisation, ensuring that outputs reflected ChatGPT’s “default” rhetorical style under classroom-like conditions rather than artificially enhanced performance. For example, the transport prompt instructed:

“You are a competent university-student writer of English texts for academic purposes. Write 16 argumentative essays with a persuasive writing style on the topic of [transport]. Each essay is about 500 words long.”

Each AI essay was generated independently to maximise lexical and syntactic diversity and avoid any replication effects. This approach captures how learners might realistically engage with generative AI tools using broad thematic prompts.

The analyses mainly draw on Hyland's (2005) frameworks of stance and engagement and metadiscourse, which see academic writing as a form of interpersonal negotiation where writers project credibility, evaluative judgments, and alignment with readers. This framework has provided the basis for numerous studies across a range of genres, such as academic blogs (Zou & Hyland, 2020) and doctoral confirmation reports (Jiang & Ma, 2018) but also include argumentative essays (Aull & Lancaster, 2014; Lee & Deakin, 2016). The study of metadiscourse has gained particular traction in comparing different groups of writers, which reveal the challenges of meeting academic norms (Hyland, 2005a; Hyland & Jiang, 2018). These have revealed contrasts in high- and low-scoring essays (e.g., Ho & Li, 2018), native and non-native speaker essays (Lee & Deakin, 2016) and genres (Hyland & Jiang, 2022), but none have examined rhetorical performance in human and machine generated texts.

In this study all texts were standardised to a base of 1,000 words to enable direct comparison. Corpus linguistics tools were used to identify and quantify relevant features, with statistical significance tested using Rayson's (2016) log-likelihood calculator at the threshold of $LL \geq 3.8$ ($p < .05$). To complement significance testing, effect sizes were calculated using %DIFF (Gabrielatos, 2018), which measures the magnitude of difference between normalised frequencies. This dual approach ensured that findings reflected both statistical robustness and practical relevance. Qualitative analysis supplemented quantitative results, illustrating how rhetorical choices shaped argumentation, authorial presence, and reader engagement.

Like all corpus-based research, there are certain limitations to these studies. The reliance on LOCNESS restricts the student sample to British undergraduates, which may not capture the diversity of academic writing across cultural and disciplinary contexts. Similarly, the ChatGPT texts reflect outputs from a specific version of the model at a particular point in time, meaning that findings may shift as AI systems evolve. The focus on short argumentative essays also limits generalisability to other genres, such as research reports or reflective assignments. Finally, while corpus methods reveal patterns of language use, they cannot fully account for the cognitive or social processes underlying writing, nor the pedagogical contexts in which AI is deployed. Despite these limitations, the methodological design provides a robust basis for comparison, offering empirical evidence of the rhetorical strengths and weaknesses of AI-generated texts.

4 Rhetorical differences: what do the results show?

The five corpus-based studies conducted by Jiang and Hyland provide a detailed picture of how ChatGPT-generated essays compare with those written by undergraduate students. Taken together, they reveal a strikingly consistent configuration: while AI excels in producing fluent, coherent, and structurally polished texts, it falls short in the rhetorical dimensions that are central to academic literacy.

The first set of findings concerns *stance and voice*. Student essays frequently employ hedges, boosters, and attitude markers to signal commitment, evaluation, and epistemic positioning. These features help construct a persuasive voice and establish authorial presence. By contrast, ChatGPT essays contain far fewer stance markers, resulting in a more impersonal tone and weaker rhetorical identity. This absence of evaluative language underscores the model's tendency to produce text that is grammatically correct but rhetorically flat (Hyland & Jiang, 2023). So, while LLMs offer considerable support for surface-level linguistic production, they do not yet simulate the complex rhetorical work that stance performs in mediating epistemic authority and academic identity. For educators, this opens up possibilities for pedagogical intervention: by positioning ChatGPT as a critical object of analysis, teachers can foster learners' understanding of how language choices shape argumentation, credibility, and engagement in academic discourse.

In terms of *cohesion and coherence*, Jiang and Hyland explored regular, high frequency three-word sequences or 'lexical bundles'. Here, ChatGPT demonstrated strong structural clarity, often relying on additive transitions such as "moreover" and "in addition", as well as explicit stage-labelling phrases like "this essay will...". These devices ensure logical sequencing but are generally more formulaic than human writing and include very few devices which reformulate material for readers. The AI generated texts also exhibited greater lexical variation than the student essays although contained a lower frequency of bundles overall, suggesting that its bundles are more rigid and standardised. Clearly, ChatGPT is proficient at recognizing and generating fixed structures due to their high frequency and regularity in the language data it was trained with. Student essays, in contrast, while less consistently fluent, employed a wider range of cohesive strategies, including reformulation and topic-shifting markers, which contribute to rhetorical flexibility and nuanced argumentation.

The studies also highlight differences in *engagement and interaction*. Student essays frequently use questions, personal asides, and direct appeals to readers, creating a dialogic relationship and anticipating counter-arguments. ChatGPT essays, on the other hand, rarely employ such engagement markers, resulting in reduced interpersonal connection. A striking example of this difference is the near absence of personal asides in ChatGPT essays, where writers break off to make a personal comment on the ongoing text. This finding serves as a clear counterpoint to any tendency to anthropomorphize or “personify” the model. Despite generating coherent text, ChatGPT rarely inserted personal comments, questions or digressions as the human writers often did, underscoring its nature as a statistical model rather than a sentient entity. Although it was able to generate a series of extended and coherent responses to our prompts, it lacked the ability to adopt a strong perspective on a topic and to engage in persuasive interactions to carry it through, thus neglecting aspects of argument that are highly valued in academic writing. These limitations reinforce the model’s tendency to produce monologic texts that lack responsiveness to audience,

Analysis of *argumentation and evidence* further illustrates the gap. The Lexical bundle analysis shows that ChatGPT essays are dominated by noun + preposition structures (e.g., “*the role of*”, “*the advantage of*”), which provide abstract descriptions but little argumentative force. Student essays, by contrast, employ more clause-based bundles (e.g., “*it is important to note*”), which signal reasoning, evaluation, and authorial presence. This distinction once more highlights something of the rigidity of AI argumentation compared to the persuasive coherence of student writing. Again, this almost certainly is a result of the machine learning mechanism that gives ChatGPT its power. The massive dataset of texts which provides its learning input almost certainly contains a high density of noun and preposition-related bundles as these are common in many registers of writing.

Finally, the studies reveal differences in *genre awareness*. As I noted earlier, ChatGPT consistently marks stages and transitions, producing structurally clear texts. Yet it lacks the interpersonal nuance provided by hedges, boosters, and attitude markers, which balance objectivity with authorial presence. Student essays, though less polished, displayed a wider repertoire of organisational strategies, including consequential linking, topic shifts, and reformulation, which are indicative of purposeful reasoning and sensitivity to audience needs. They also showed greater genre awareness by combining organisation with interpersonal positioning, thereby demonstrating a richer understanding of academic discourse.

These findings are summarised in Table 2.

Table 2: Comparison of AI-Generated vs. Student Essays

Dimension	ChatGPT essays	student essays
Stance and Voice	Fewer hedges, boosters, and attitude markers Weaker authorial presence More impersonal tone	More stance markers Stronger authorial identity and persuasive voice
Cohesion and Coherence	Heavy reliance on additive transitions ("moreover," "in addition") and stage-labelling ("this essay will...") Formulaic cohesion	Wider range of cohesive devices, including reformulation and topic shifts Greater rhetorical flexibility
Engagement and Interaction	Limited use of questions, personal asides, and direct appeals to readers Limited reader connection	Frequent engagement strategies More dialogic interaction and anticipation of objections
Argumentation and Evidence	Dominated by noun + preposition bundles ("the role of") Abstract, rigid argumentation	More clause-based bundles ("it is important to note") Greater coherence using reasoning, evaluation, and persuasive devices
Genre Awareness	Strong structural clarity, Consistent stage marking Reduced interpersonal nuance	Combine organisation with hedges, boosters, and attitude markers Richer rhetorical awareness

Across all five dimensions, the findings converge on a clear conclusion: ChatGPT produces fluent and coherent texts but lacks the rhetorical depth, interpersonal engagement, genre awareness and authorial presence that define effective academic writing. Our findings indicate that AI-generated texts seem to have little difficulty in satisfying surface expectations of coherence, but fail to demonstrate the deeper competencies that written assessments are designed to measure. This means that if our assessment criteria overemphasise structural clarity or linear cohesion, they risk privileging machine-generated writing over the developmental evolution of student learning. Understanding these distinctions is therefore essential for preserving construct validity in writing assessment and for supporting principled responses to the increasing presence of AI in academic contexts.

While algorithmic coherence by LLMs facilitates logical flow, it cannot currently replicate the reader-driven process of constructing meaning that is characteristic of human-written texts. This limited rhetorical flexibility highlights the importance of interaction in crafting persuasive and engaging arguments. This distinction between surface fluency and rhetorical competence provides the empirical foundation for the pedagogic framework proposed in the next section. This framework employs AI as a scaffold for teaching rhetorical awareness and fostering critical AI literacy, rather than as an automated author. Overall, then, the results suggest that AI literacy should become an integral component of writing instruction.

5. Implications: Towards an AI-informed pedagogical framework

Essentially, the framework positions ChatGPT as a writing assistant. Its ability to deliver texts with grammatical accuracy, structural clarity, and lexical sophistication make it a valuable model for students struggling with fluency and organisation. However, its obvious limitations in interactivity and rhetorical nuance mean that it cannot replace the intellectual and interpersonal work of academic writing. Instead, AI outputs might better be treated as pedagogical objects of analysis, allowing students to identify rhetorical gaps and reflect on how language choices shape persuasion and credibility. The framework therefore carries, I think, significant implications for pedagogy, assessment, ethics, and future research in writing instruction in higher education.

Pedagogically, it suggests that ChatGPT should be integrated into writing instruction as a scaffold for developing rhetorical awareness. Rather than treating AI outputs as finished products, instructors can frame them as texts to be interrogated, critiqued, and revised. In this way, students are encouraged to recognize the rhetorical gaps that distinguish fluent prose from persuasive academic arguments. Classroom activities might include contrastive analysis, where learners compare AI-generated essays with their own or with published exemplars to foreground differences in, say, stance and engagement. Teachers may use the model's outputs, as we have in these studies, as a contrasting backdrop to human writing, focusing on salient differences to foster skills of critical analysis, enabling students to discern between the mechanical organisation of ideas and the nuanced, persuasive expression that characterises good argumentative writing. Students can also be encouraged to improve upon these AI-generated drafts by adding personalized and contextually appropriate rhetorical features, thereby developing their skills in crafting more effective academic arguments.

In addition, revision tasks, where students enhance AI drafts by inserting stance markers, hedging devices, and engagement features, can also offer valuable opportunities for learners to cultivate authorial presence and metalinguistic awareness. Prompt design experimentation by students further encourages exploration in how input shapes output, fostering metalinguistic awareness and critical AI literacy. Through these kinds of activities AI becomes a catalyst for teaching rhetorical competence, foregrounding the interpretive and interactive dimensions of academic writing rather than reducing it to surface fluency. Clearly, teacher education programmes should include explicit training in understanding the linguistic and rhetorical capacities of LLMs, so that instructors are better prepared to guide students in their use. Such training could include genre awareness, stance construction, prompt engineering, and critical discourse analysis of AI-generated texts.

The framework also has important implications for *assessment*. Current rubrics that privilege clarity, coherence, and grammatical accuracy risk inadvertently rewarding AI outputs over student learning. To safeguard construct validity, assessment practices must be recalibrated to emphasise reasoning, rhetorical adaptability, and responsiveness to audience expectations. Embedding process-oriented approaches, including staged drafting, annotated reasoning, and reflective commentary, ensure that assessments capture the intellectual work of students rather than the polished surface fluency of machine-generated text. Transparency in AI use should be encouraged, with students asked to reflect explicitly on how they employed tools such as ChatGPT and what rhetorical limitations they identified. Such practices not only protect academic integrity but also foster critical literacy by making students more aware of their own rhetorical choices and the limitations of algorithmic fluency.

Ethically, the integration of AI into academic writing raises pressing questions of authorship, agency, and accountability. Here institutions must establish clear policies that balance innovation with integrity, guiding students to use AI responsibly and reflectively.

Encouraging learners to acknowledge AI's role in their writing processes helps demystify its contribution while reinforcing the value of human rhetorical competence. We need, then, as teachers, to find ways to help students in understanding that while AI can assist in generating ideas and drafting texts, key areas of academic writing such as critical thinking, personal engagement, and creativity remain, for the moment, outside of the abilities of LLMs. In this way, AI can be positioned as a partner in learning rather than a threat to authenticity,

supporting the development of reflective writers who understand both the affordances and the constraints of machine assistance.

Finally, the framework points toward several directions for *future research*. Longitudinal studies are needed to track how evolving AI models develop rhetorical capacities over time and how students' engagement with them changes. Cross-cultural investigations can explore how AI-generated texts align with diverse academic traditions and rhetorical expectations, while classroom-based studies will be particularly valuable in capturing how students interact with AI in practice and how pedagogical interventions shape their development of rhetorical awareness. Lines of inquiry such as these will be essential for refining frameworks of AI integration and ensuring that higher education remains responsive to technological change. It also contributes to a teaching-learning context that balances technological innovation with an enduring commitment to cultivating academic literacy

Figure 1 summarises these implications, indicating the framework, its supporting arguments and future directions for research into how we might use AI productively and ethically.

Strengths	Limitations
Fluency and grammatical accuracy	Weak stance and authorial presence
Structural clarity and coherence	Limited engagement with readers
Lexical sophistication	Formulaic cohesion and bundle use
Accessible models of organisation	Lack of rhetorical depth and nuance

Pedagogical Uses	Ethical & Assessment Implications
Contrastive analysis tasks	Risk of privileging AI outputs if rubrics overemphasise clarity and coherence
Revising AI drafts to add stance	Need to recalibrate descriptors to value reasoning, adaptability, and audience
Prompt design experimentation	Combine with process-oriented assessment
Support rhetorical awareness	Transparency in AI use
Developing critical AI literacy	

Emerging research Directions

Longitudinal tracking of AI's evolving rhetorical capacities

Integration of AI literacy into writing pedagogy

Positioning AI as scaffold, not surrogate author

Figure 1: Reasons for and implications of using ChatGPT as a writing assistant

6 Conclusion

This study has explored the role of generative AI in academic writing, situating ChatGPT within the broader debates about fluency, rhetoric, and integrity in higher education. I have highlighted both the promise and the risks of AI integration and summarised the results of five recent corpus-based studies by Jiang and Hyland which provide empirical evidence of the rhetorical gaps that distinguish AI-generated texts from student essays. Across stance, cohesion, engagement, argumentation, and genre awareness, the findings underscore a consistent pattern: showing how ChatGPT produces structurally polished texts but lacks the rhetorical depth and interpersonal nuance that define effective academic literacy.

Following a line from this comparative corpus evidence to the classroom, I have proposed a conceptual framework that treats AI outputs as objects of analysis, revision, and reflection. Taking this approach seriously means that writing teachers can harness AI to strengthen what they do best: assisting learners to produce fluent, clear and accessible texts. The framework therefore addresses current limitations of ChatGPT in creating rhetorically competent essays while also emphasising the need to rethink assessment practices to more transparently reveal AI use and to cultivate critical AI literacy, prompting students not only to use AI responsibly but also to recognise and compensate for its rhetorical shortcomings.

Beyond pedagogy, institutions must guide students in navigating questions of authorship and accountability, while scholars should continue to investigate how AI evolves and how learners engage with it across contexts. Ultimately, this paper is a small contribution to the growing work seeking to reframe AI not as a threat to academic integrity but as a catalyst for pedagogical innovation. By embedding AI within a critically informed framework, higher education can prepare students to thrive in hybrid writing ecologies where human creativity and algorithmic fluency intersect. In doing so, we safeguard the intellectual and rhetorical

dimensions of academic writing while embracing the opportunities of a digital age. The challenge, then, is not whether AI belongs in the classroom, but how we design practices that ensure students remain the authors of their own learning.

References

- Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, 1–14.
<https://doi.org/10.1080/10494820.2023.2253858>
- Aull, L., & Lancaster, Z. (2014). Linguistic markers of stance in student writing: Corpus-based comparisons across disciplines. *Written Communication*, 31(3), 271–305.
<https://doi.org/10.1177/0741088314532518>
- Barrot, J. (2023). Exploring generative AI in language education: Opportunities and challenges. *Language Education Review*, 12(2), 45–62.
- Chegg Global Student Survey (2025). Chegg, Inc., accessed March 12, 2026
<https://investor.chegg.com/Press-Releases/press-release-details/2025/Chegg-Global-Student-Survey-2025-80-of-Undergraduates-Worldwide-Have-Used-GenAI-to-Support-their-Studies-But-Accuracy-a-Top-Concern/default.aspx>
- Das, P. D., & Chen, Y. (2025). The role of AI in academic writing: Impacts on writing skills, critical thinking, and integrity in higher education. *Societies*, 15(9), 247.
<https://doi.org/10.3390/soc15090247>
- Deep, P. D., & Chen, Y. (2025). The role of AI in academic writing: Impacts on writing skills, critical thinking, and integrity in higher education. *Societies*, 15(9), 247.
<https://doi.org/10.3390/soc15090247>
- Gabrielatos, C. (2018). Keyness: Appropriate metrics and practical issues. *International Journal of Corpus Linguistics*, 23(2), 166–210. <https://doi.org/10.1075/ijcl.16034.gab>
- Granger, S. (2013). The Louvain Corpus of Native English Essays (LOCNESS). Louvain-la-Neuve: Université Catholique de Louvain.
- Granger, S. (2017). Academic writing development: Insights from learner corpora. *International Journal of Learner Corpus Research*, 3(1), 5–24.
- Gao, R., Yu, D., Gao, B., Hua, H., Hui, Z., Gao, J., & Yin, C. (2025). Legal regulation of AI-assisted academic writing: Challenges, frameworks, and pathways. *Frontiers in Artificial Intelligence*, 8, 1546064. <https://doi.org/10.3389/frai.2025.1546064>

- Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports*, *13*(1), 18617.
- Ho, V., & Li, C. (2018). The use of metadiscourse and persuasion: An analysis of first year university students' timed argumentative essays. *Journal of English for Academic Purposes*, *33*, 53–68.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London: Continuum.
- Hyland, K. (2026). *Teaching and researching writing* (5th edition). Routledge
- Hyland, K., & Jiang, F. (2016). Change of Attitude? A Diachronic Study of Stance. *Written Communication*, *33*(3), 251–274.
- Hyland, K., & Jiang, F. (2018). Academic lexical bundles: How are they changing? *International Journal of Corpus Linguistics*, *23*(4), 383–407.
- Hyland, K. & Jiang, K. (2022) Metadiscourse across languages and genres: An overview. *Lingua* *252*.
- Hyland, K., & Jiang, F. (2023). Interaction in written texts: A bibliometric study of published research. *Studies in Second Language Learning and Teaching*, *13*(4), 903–924.
<https://doi.org/10.14746/ssl.40220>
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, *15*(4), ep464. <https://doi.org/10.30935/cedtech/13605>
- Jiang, F. & Hyland, K. (2024). Does ChatGPT argue like students? Bundles in argumentative essays. *Applied Linguistics*.
- Jiang, F. & Hyland, K. (2025a). Rhetorical distinctions: Comparing metadiscourse in essays by ChatGPT and students. *English for Specific Purposes*. *79*: 17–29
- Jiang, F. & Hyland, K. (2025b). Does ChatGPT write like a student? Engagement markers in argumentative essays. *Written Communication*. *42.3*
- Jiang, F. & Hyland, K. (2026). Taking a stance: AI generated vs student written argumentative essays. *Applied Linguistics Review*
- Jiang, F. & Hyland, K. (in Press). AI, argumentation and assessment: An analysis of interactive metadiscourse in ChatGPT and student writing. *Journal of Pragmatics*
- Jiang, F., & Ma, X. (2018). 'As we can see': Reader engagement in PhD candidature confirmation reports. *Journal of English for Academic Purposes*, *35*, 1–15.
<https://doi.org/10.1016/j.jeap.2018.05.003>

- Kelly, R. (2024). Survey: 86% of Students Already Use AI in Their Studies. *Campus Technology*. <https://campustechnology.com/articles/2024/08/28/survey-86-of-students-already-use-ai-in-their-studies.aspx>
- Kim, J., Lee, S., Detrick, R., Wang, J., & Li, N. (2025). Students–Generative AI interaction patterns and their impact on academic writing. *Journal of Computing in Higher Education*, 38, 504–525.
- Lee, J. J., & Deakin, L. (2016). Interactions in academic writing: A corpus study of rhetorical moves and stance. *Journal of English for Academic Purposes*, 24, 92–107. <https://doi.org/10.1016/j.jeap.2016.09.002>
- Li, H., & Wu, X. (2025). The use of generative AI tools in academic writing: A systematic review of research trends and thematic insights. *AI & Society*. <https://doi.org/10.1007/s43681-025-00827-0>
- Markey, B., Brown, D. W., Laudenbach, M., & Kohler, A. (2024). Dense and disconnected: Analyzing the sedimented style of ChatGPT-generated text at scale. *Written Communication*, 1–30.
- Pack, A., & Maloney, J. (2023). Prompting literacy: Teaching students to write with AI. *Computers and Composition*, 67, 102–118. <https://doi.org/10.1016/j.compcom.2023.102118>
- Rayson, P. (2016). Log-likelihood and effect size calculator. Lancaster University. Retrieved from <http://ucrel.lancs.ac.uk/llwizard.html>
- Sarrion, J. (2023). Prompt engineering for generative AI: Strategies and implications. *AI in Education Journal*, 5(1), 33–49.
- Siddiqui, M. N., Feliciano, V., Pea, R., & Subramonyam, H. (2025). *AI in the writing process: How purposeful AI support fosters student writing*. In *Artificial Intelligence in Education* (pp. 190–203). Springer.
- Speri, A. (2026) I wish I could push ChatGPT off a cliff’: professors scramble to save critical thinking in an age of AI. *The Guardian*. <https://www.theguardian.com/technology/ng-interactive/2026/mar/10/ai-impact-professors-students-learning>
- Vee, A., Laquintano, T., & Schnitzler, C. (2023). TextGenEd: Teaching with text generation technologies. *The WAC Repository*, 1(1), 1–100.
- Yang, S., Chen, S., Zhu, H., Lin, J., & Wang, X. (2024). A comparative study of thematic choices and thematic progression patterns in human-written and AI-generated texts. *System*, 126, 103494.

Zou, H. & Hyland, K. (2020). "Think about how fascinating this is": Engagement in academic blogs across disciplines. *Journal of English for Academic Purposes*. 43. <https://doi.org/10.1016/j.jeap.2019.100809>