STUDENTS' PERCEPTIONS OF GENERATIVE AI IN ACADEMIC WRITING FOR ESP COURSES

Vuong Thi Hai Yen

Faculty of Foreign Languages, Hanoi Metropolitan University

vthyen@hnmu.edu.vn

ORCID: https://orcid.org/0009-0005-9163-5128

https://doi.org/10.63506/jilc.1002.438

(Received: 05/05/2026; Revised: 27/07/2026; Accepted: 17/8/2026)

Abstract: This descriptive survey examines English-major undergraduates' perceptions of generative artificial intelligence (AI) as a tool to support academic writing in ESP courses. Data were collected from 123 students at a public university in Hanoi through a Vietnamese questionnaire combining five-point Likert items and open-ended questions. The analysis draws on Technology Acceptance Model, AI literacy, academic ethics, and ESP genre-based pedagogy. Students generally perceived generative AI as useful for idea generation, language checking, terminology support, and academic genre awareness. However, they also reported concerns about inaccurate outputs, limited confidence in evaluating AI-generated content, and unclear ethical boundaries. The findings therefore suggest the need for cautious pedagogical integration. ESP instruction may benefit from explicit support for critical evaluation, discipline-sensitive prompting, responsible use, and clearer institutional guidance on acceptable AI-assisted academic writing practices.

Keywords: Generative AI; academic writing; English for Specific Purposes; Technology Acceptance Model; AI literacy

NHẬN THỨC CỦA SINH VIÊN VỀ TRÍ TUỆ NHÂN TẠO TẠO SINH TRONG VIẾT HỌC THUẬT Ở CÁC HỌC PHẦN TIẾNG ANH CHUYÊN NGÀNH

Tóm tắt: Nghiên cứu khảo sát mô tả này tìm hiểu nhận thức của sinh viên ngành Ngôn ngữ Anh về trí tuệ nhân tạo tạo sinh (AI) như một công cụ hỗ trợ viết học thuật trong các học phần tiếng Anh chuyên ngành. Dữ liệu được thu thập từ 123 sinh viên tại một trường đại học công lập ở Hà Nội bằng bảng hỏi tiếng Việt gồm các mục theo thang Likert năm mức độ và câu hỏi mở. Phân tích dựa trên Mô hình Chấp nhận Công nghệ mở rộng, năng lực AI, đạo đức học thuật và giảng dạy tiếng Anh chuyên ngành dựa trên thể loại văn bản. Nhìn chung, sinh viên đánh giá AI tạo sinh hữu ích trong gợi ý ý tưởng, kiểm tra ngôn ngữ, hỗ trợ thuật ngữ và nhận biết các đặc điểm của thể loại văn bản học thuật. Tuy nhiên, họ cũng lo ngại về thông tin thiếu chính xác, hạn chế trong khả năng đánh giá nội dung do AI tạo ra và ranh giới đạo đức chưa rõ. Kết quả cho thấy việc tích hợp AI vào giảng dạy cần được thực hiện một cách thận trọng, đi kèm với hướng dẫn về đánh giá phản biện, xây dựng câu lệnh phù hợp với đặc thù chuyên ngành, sử dụng AI có trách nhiệm và hướng dẫn rõ ràng từ cơ sở đào tạo về cách thức sử dụng AI được chấp nhận trong viết học thuật.

Từ khóa: Trí tuệ nhân tạo tạo sinh; viết học thuật; tiếng Anh chuyên ngành; Mô hình Chấp nhận Công nghệ; năng lực AI

1. Introduction

The rapid emergence of generative artificial intelligence tools such as Copilot, Gemini, ChatGPT, and DeepSeek has changed the environment in which academic writing is taught and practiced in higher education. Students increasingly use such tools to brainstorm, check their language, develop text structure, and identify relevant terminology (Crompton et al., 2024; Sanz-Tejeda et al., 2026). The issue is especially relevant to academic writing in ESP courses. In these courses, students work not only with English but also with the genres, purposes, and discourse conventions attached to specialized academic tasks. The central question, therefore, is not simply whether students use AI. It is how they use it, how they evaluate it, and how that use may affect their development as academic writers.

This question has clear practical relevance in Vietnamese higher education. ESP has long been a required component in many degree programs, yet it continues to face familiar constraints: large class sizes, limited contact hours, scarce authentic discipline-specific texts, and a sizeable gap between students' general English proficiency and the demands of professional genres (Nguyen, 2017). Under such conditions, generative AI can serve as a convenient support tool, functioning as a dictionary, tutor, proofreader, and source of models. However, an increasing number of studies warn that excessive reliance on AI, without adequate pedagogical scaffolding, could limit opportunities for critical thinking. This could also create ethical concerns and result in students copying textual patterns produced by AI without proper evaluation (Gayed et al., 2022; Li et al., 2025). At the same time, empirical evidence on the benefits and risks of AI-assisted writing is far from conclusive. Barrett and Pack (2023) examined student and teacher perceptions of acceptable GenAI use across the writing process. They found broad acceptance of several uses related to revision and teacher feedback, while also identifying a need for clearer classroom and institutional guidance. Other studies have found that second language writers can improve their writing quality and motivation while reducing their language anxiety (Derakhshan, 2025; Song & Song, 2023). Yet, along with these apparent advantages, scholars have also raised concerns about overdependence, superficial rhetorical engagement, unreliable content, and inconsistent transparency in the use of AI (Li et al., 2025; Mo & Crosthwaite, 2025; Wu et al., 2026). These concerns carry particular weight for ESP teaching, since students must often manage linguistic accuracy, academic structure, and specialized vocabulary simultaneously.

The present study examines this issue from the viewpoint of English-major students at a university in Vietnam. It focuses on academic writing tasks commonly encountered in ESP-oriented courses, including essays, outlines, reports, abstracts, case analyses, and terminology-supported academic discussion. The study does not claim to represent all ESP settings in Vietnam. Its narrower aim is to describe how one group of students perceives generative AI as a support tool for academic writing in their course context. This focus situates the study within applied language education and ESP writing pedagogy rather than general technology adoption. The study is guided by the following research questions.

1) How do English-major students report using generative AI for academic writing in ESP courses?

2) How do students perceive the usefulness, ease of use, and future use of generative AI for academic writing in ESP courses?

3) What challenges do they report in using generative AI for academic writing in ESP courses?

2. Literature review

2.1 Generative AI and academic writing

Earlier writing-support technologies tended to operate at the surface of the text, checking spelling, flagging punctuation problems, or proposing local grammar corrections. Generative AI works on a broader scale. It is capable of producing long texts, reorganizing ideas, proposing genre-specific structures, and providing feedback that mimics academic writing conventions (Crompton et al., 2024). This represents a move from basic editing towards engagement with planning, drafting, revising, and reflection across the writing process. Empirical investigations into the reactions of students and teachers to this transition are only now beginning to appear. Barrett and Pack (2023) found that students and teachers generally viewed several uses of GenAI in revision and feedback as acceptable, although their study examined perceptions of appropriate use rather than measured writing outcomes. In the intervention reported by Song and Song (2023), the ChatGPT-assisted group showed gains in writing performance and motivation. Generative AI might alleviate language anxiety, freeing up cognitive resources for second-language writers to concentrate on arguments and analysis (Derakhshan, 2025). Such findings bear directly on ESP students, who often need support with language, structure, and specialized expression all at once.

Research has also revealed certain limits. Li et al. (2025) warn that students may miss out on developing their writing skills when AI takes over much of the writing. Mo and Crosthwaite (2025) show how large language model-generated texts can appear polished even when they rely on an impoverished set of stance and engagement resources. Wu et al. (2026) find an apparent contradiction between support for transparency in principle and the practice of disclosing AI use. Taken together, these studies suggest that generative AI is not a neutral writing aid, but a pedagogical object that calls for guidance, verification, and ethical reflection.

2.2 Academic writing in ESP and genre-based support

ESP is commonly understood as a needs-driven approach to language teaching. Within ESP, pedagogical decisions are influenced by the learners' current competence, the target communication contexts and their professional or disciplinary goals (Dudley-Evans & St. John, 1998; Hutchinson & Waters, 1987). In the present study, the term ESP does not refer to a single professional discipline. It refers to academic writing tasks in English language courses that require students to work with academic genres, discipline-related vocabulary, and genre conventions relevant to specialized communication. In a Vietnamese ESP needs analysis, Ho (2017) reported students' priorities for workplace-oriented language skills and course materials. The study concerns curriculum needs rather than AI-assisted academic writing, but it reinforces the importance of relating technological support to specific learning tasks.

Genre theory helps to explain this focus. Swales (1990) and Bhatia (1993) argue that genres are not fixed templates but rather are patterned forms of social action. Similarly, Belcher (2006) argues that ESP needs analysis has shifted from analysis of surface language forms to institutional context, discourse community and learner positioning. Within this perspective, writing instruction involves helping students notice the rhetorical moves of academic genres and use them with

control. In the present context, such genres include academic essays, outlines, reports, abstracts, and case-based writing tasks.

Generative AI enters this context in a specific way. It can provide quick access to examples, alternative phrasings, structural suggestions, and terminology support. These affordances may be useful where students have limited access to authentic academic texts. Yet the same affordances may also weaken learning if students accept AI outputs as ready-made genre knowledge. For ESP pedagogy, the central issue is therefore not whether AI can generate models. The issue is whether learners can evaluate, adapt, and internalize those models through guided practice.

2.3 AI literacy and academic ethics

Long and Magerko (2020) define AI literacy as a set of competencies enabling individuals to evaluate AI technologies critically, to communicate and collaborate with them productively, and to put them to use across contexts. In academic writing, this literacy involves at least three elements. The first is a basic understanding that AI may fabricate information or produce language that does not suit a particular academic purpose. The second is the ability to design prompts matching writing goals. The third is the ability to evaluate, revise, and integrate AI output without ceding authorial control. Ngo and Hastie (2025) introduce a pragmatic pedagogical model entitled "AI for Academic Purposes" (AIAP). The ten-week EAP module incorporated AI literacy into writing instruction, not as an optional technical add-on. The results of this study indicate that deliberate instruction can boost students' confidence in AI tools, enhance their ability to evaluate AI output, and expand their understanding of ethical use. Accordingly, the present study uses AI literacy to refer consistently to these competencies.

Ethical concerns have also moved to the center of this research area. Kohnke (2024) found that EAP students in Hong Kong valued generative AI more than traditional grammar checkers, despite expressing concerns about over-reliance, data use, and unequal access to premium features. Chen and Gong (2025) similarly report concerns about over-reliance, ethical issues, technical constraints, and the unreliability of AI-generated content in academic writing in Chinese as a second language. Recent research in Vietnam likewise presents a mixed picture. Lai and Trinh (2025) report perceived benefits for academic vocabulary, grammatical accuracy, idea generation, and motivation among English-major undergraduates. Nguyen et al. (2025) found that students valued ChatGPT for grammar, vocabulary, writing organization, and time saving, while also expressing concerns about dependence, reduced critical thinking, and academic originality. Tran (2025) further identifies usefulness, ease of use, social influence, and facilitating conditions as factors associated with ChatGPT acceptance and argues for AI-literacy initiatives and ethical guidance in academic writing.

2.4 The technology acceptance model and its limits

The Technology Acceptance Model (TAM), proposed by Davis (1989), explains technology acceptance through two core variables: perceived usefulness and perceived ease of use. These variables shape attitude and behavioral intention. TAM remains useful because it offers a clear language for examining why students may continue or discontinue using a tool. However, TAM was not designed for generative AI or for academic writing, where questions of authorship, institutional rules, and ethical legitimacy are central.

For this reason, the present study uses TAM as an organizing frame rather than a complete explanatory model. Recent research on AI acceptance in education has expanded TAM with social, normative, and contextual variables (Kılıç & Çelik, 2025). In this study, usefulness, ease of use, and intention to continue using AI are therefore read together with AI literacy, genre support, challenges, academic ethics, and institutional policy. This extended frame is better suited to the study because a student may find AI useful and easy to use while still feeling uncertain about whether a particular form of use is academically legitimate.

2.5 Empirical studies of AI-assisted writing in language education

Prior studies offer different kinds of evidence about AI-assisted writing. Barrett and Pack (2023) examined student and teacher judgments about acceptable generative AI use across writing activities. Song and Song (2023) investigated writing performance and motivation in an EFL intervention. These studies address ethical judgments and measured outcomes, respectively. Neither design can establish how students in the present ESP-oriented courses evaluate support for specialized academic writing. Kohnke (2024) examined EAP students' perceptions of generative AI and grammar-checking tools, including reservations about overreliance and access. Ngo and Hastie (2025) investigated the integration of AI literacy into an EAP module. Their instructional context provides a useful reference for teaching evaluation of AI output, although the present survey did not examine the effects of an AI-literacy intervention.

Vietnamese studies provide a more immediate comparison. Lai and Trinh (2025) addressed ChatGPT and English-major students' academic writing; Nguyen et al. (2025) reported perceived assistance with vocabulary, grammar, organization, and time saving, alongside concerns about dependence and originality. More directly, Do et al. (2026) studied AI chatbots in self-regulated EFL writing using survey and interview data. Participants reported using chatbots during the writing process, especially at pre-writing and post-writing stages, and identified needs for support. These studies show that AI use and students' writing-related perceptions have already been examined in Vietnam. In a related but distinct ESP context, Ho (2017) documented learners' needs for particular language skills and materials. That study did not examine generative AI or academic writing, but it supports attending to the tasks and disciplinary purposes of ESP teaching. Kılıç and Çelik (2025) studied technology acceptance among English language instructors; their findings cannot be taken as evidence about the present undergraduate sample.

3. Research methodology

3.1 Research design

The study follows a descriptive survey design. Quantitative and qualitative data were collected through a single questionnaire. This design was selected because the study aimed to map broad patterns in students' perceptions while also allowing students to explain how they used AI, what difficulties they encountered, and what kinds of institutional support they expected.

3.2 Context and participants

The study took place in the second semester of the 2025–2026 academic year at the foreign-language faculty of a public university in Hanoi. The participants were students enrolled in courses that included ESP-oriented academic writing tasks such as essays, outlines, reports, abstracts, and case-based writing. Participation was voluntary, anonymous, and participants could

withdraw at any time. In total 123 responses were kept for analysis. The sample was predominantly female with 82.9% (n = 102) identifying as female, 16.3% (n = 20) as male and 0.8% preferring not to specify their gender. Most of the respondents were third year students (68.3%) followed by second year students (22.8%) and fourth year students (8.9%). For self-reported English proficiency, 87.0% placed themselves at B1–B2, 8.1% at A1–A2, and 4.9% at C1–C2. Most had prior experience with generative AI: 46.3% had used it for more than one year, 26.0% for six months to one year, and 22.0% for less than six months. Seven students (5.7%) reported no prior AI use. Their treatment in experience-dependent analyses is specified in Section 3.4.

3.3 Data collection instruments

The questionnaire was developed by the research team in Vietnamese in order to minimize linguistic ambiguity for respondents. English labels such as "Technology Acceptance Model," "Perceived Usefulness," and "AI literacy" were retained where they identified the analytical constructs. Part A collected demographic information and prior experience with AI. Parts B through I contained 45 five-point Likert-scale items: reported AI use for writing activities (6 items), perceived usefulness (6 items), perceived ease of use (5 items), continued use and training-related perceptions (4 items), genre support (5 items), immediate feedback and linguistic confidence (6 items), challenges and AI literacy (6 items), and ethics and institutional policy (7 items). All 45 Likert items were rated from 1 (strongly disagree) to 5 (strongly agree). The questionnaire ended with three optional open-ended questions about phases of AI use, difficulties encountered, and recommendations to the institution and instructors. Two instructors experienced in ESP and academic writing reviewed the questionnaire for construct relevance, clarity of the Vietnamese wording, and correspondence with writing tasks encountered in the courses. A formative pilot involving five students outside the main sample was then used to check item clarity and the response format. Given this limited purpose, the pilot was not treated as psychometric validation. Feedback from the expert review and pilot led to simpler wording in several items, clearer examples of writing tasks, and less leading language in items concerning ethics and institutional policy.

3.4 Data analysis

Quantitative analysis was conducted at the respondent level. The analytic cohort remained the 123 responses. For the 26 items in Parts C to G, the seven students who reported no prior AI use were excluded because these items presupposed experience with AI. The other Likert items retained the full analytic sample. The respondent-level spreadsheet also showed a small number of cases in which more than one Likert category had been selected for a single item. Because the questionnaire intended one response per item, these records were treated as item-level missing values rather than being recoded. This affected 63 of 5,353 eligible item-level records, or 1.2%, across 33 respondents. Valid N therefore varies slightly by item. For every Likert item, the analysis calculated the mean, standard deviation, valid N, and the combined percentage at levels 4 and 5 as an agreement indicator. Table 1 summarizes the ranges across each questionnaire section. Selected item-level percentages, M, and SD are reported in the Findings only where they directly substantiate claims about stronger or weaker endorsement and address the research questions; the full item-level tables are not reproduced. No inferential tests were conducted because the research questions are descriptive and exploratory. Internal consistency was assessed

separately for each multi-item section using Cronbach's alpha based on listwise complete cases within that section. For each section, the *N* used to calculate Cronbach's alpha was the number of respondents with complete responses to all items in that section; it could therefore differ from the item-level valid *N* ranges reported in Table 1. The coefficients were .912 for Part B, .920 for Part C, .901 for Part D, .882 for Part E, .910 for Part F, .908 for Part G, .891 for Part H, and .895 for Part I. The training-request item in Part E is interpreted as an expressed support need rather than as a direct behavioral-intention indicator.

The three open-ended questions were optional, and response counts therefore varied: 43 students responded to J1, 40 to J2, and 33 to J3. The responses were analyzed thematically following Braun and Clarke (2006). The analysis involved close reading of all available responses, open coding, grouping of related codes into themes, and selection of representative excerpts. Excerpts in the Findings were translated from Vietnamese into English for reporting. To protect anonymity and make the qualitative evidence traceable, responses were identified by question number and response order, for example J1-R06 or J3-R25. The translations were intended to preserve the meaning of the original responses rather than reproduce spelling or grammar errors.

3.5 Research ethics

Participants were informed about the purpose of the research, the anonymous nature of the data, and the fact that their responses would not affect their academic results. All data were used solely for research purposes and were stored in an account accessible exclusively to the research team.

4. Findings

4.1 Generative AI use in academic writing for ESP courses

Responses to the writing-use items indicate the activities for which students reported drawing on generative AI, rather than providing a direct measure of time spent in each writing phase. Grammar, spelling, and expression checking received the strongest endorsement (76.2% agreement, $M = 4.04$, $SD = 0.85$). Seeking feedback on drafts before submission was also reported, although its endorsement was lower (68.0%, $M = 3.89$, $SD = 0.83$). Across the six items, agreement ranged from 68.0% to 76.2% (Table 1). Taken together, these responses describe perceived use for language checking, planning, terminology-related support, and review; they do not establish the relative frequency of AI use at each stage of writing.

The responses to the optional open-ended question add a temporal dimension to these reported purposes. Students described using AI during preparation, when encountering difficulty developing ideas, and when reviewing a draft. One response referred to use “during outlining and the search for vocabulary and writing structures” (J1-R06). Another described using AI “before writing to build an outline and expand ideas, and after writing to check grammar and expression” (J1-R31). A further student turned to AI when “short of ideas” or uncertain whether a draft was appropriate (J1-R28). These examples help locate the reported activities within the writing process. They are illustrative accounts from the open-ended responses, not quantitative estimates of phase-specific use.

4.2 Usefulness, ease of use, and future use intentions

Students with prior AI experience generally evaluated generative AI positively. The strongest selected usefulness item concerned faster and more efficient task completion (78.4% agreement, $M = 4.00$, $SD = 0.77$), while the general ease-of-use item received 79.1% agreement ($M = 3.99$, $SD = 0.73$). Regarding future use, 78.1% indicated that they would use AI more often if the institution provided clear guidance ($M = 4.08$, $SD = 0.80$), which suggests a conditional intention to increase use. Additionally, 81.0% expressed interest in further training on effective and responsible AI use ($M = 4.12$, $SD = 0.78$). This indicates a need for additional support rather than an intention to use AI more frequently (Table 1).

Responses concerning genre support provided further detail about perceived usefulness in ESP writing. Across the five items, agreement ranged from 67.5% to 71.6%. Students most strongly endorsed AI as a source of useful examples of academic sections (71.6%, $M = 3.91$, $SD = 0.81$), while agreement was somewhat lower for its support in understanding distinctive genre structures (67.5%, $M = 3.83$, $SD = 0.81$). These responses suggest that students perceived AI as helpful for accessing examples and understanding genre organization (Table 1).

Open-ended responses further illustrated these perceptions. One student described using AI "during the stage of organizing the layout" (J1-R26), while another characterized AI-generated outlines for academic essays as "detailed and logical" (J1-R38). Both comments highlighted AI's perceived value for planning and organizing academic writing.

Students also reported positive perceptions of AI-supported feedback and linguistic confidence. Across the six items, agreement ranged from 75.0% to 86.2% (Table 1). The strongest selected item concerned students' perceived ability to express academic ideas in English (86.2%, $M = 4.10$, $SD = 0.70$). Other items addressed error noticing, revision, confidence, and anxiety. Together, these findings suggest that students valued AI for its perceived support in writing efficiency, genre organization, feedback and academic expression. However, these self-reported perceptions do not establish actual improvements in writing performance or confidence over time.

4.3 Reported challenges in using generative AI for academic writing

In response to the third research question, students described difficulties in using and evaluating AI for academic writing. The most strongly endorsed selected challenge item concerned the need for more instructor guidance on critical and effective use (72.4% agreement, $M = 3.95$, $SD = 0.88$). Concern about inaccurate or outdated AI-generated content received 69.4% agreement ($M = 3.90$, $SD = 0.85$). Uncertainty about producing effective discipline-specific prompts received 59.3% agreement ($M = 3.63$, $SD = 0.99$) (Table 1). These items describe different reported obstacles, rather than a single measured level of AI literacy.

The open-ended responses further illustrate these concerns. One student wrote that "AI sometimes gives false information so subtly that it is very hard to detect, even after several prompts, so I have to check it manually" (J2-R04). Another wrote that AI "often produces information without supporting evidence" (J2-R27). Several students focused on language and terminology, noting that AI may use "less common vocabulary" (J2-R02), "too many complicated words" (J2-R16), or terms that are too general for the disciplinary context (J2-R32). These comments illustrate students' concerns about the accuracy and reliability of AI-generated

information, as well as the appropriateness of its language and terminology for discipline-specific writing.

Ethical and policy uncertainty formed another group of reported challenges. Some 53.3% agreed with the item describing current school or faculty rules as unclear or inconsistent (M = 3.52, SD = 0.98), and 57.0% reported uncertainty about the boundary between acceptable AI assistance and an academic ethics violation (M = 3.61, SD = 1.03). These are perceptions of the guidance available to respondents; they do not independently establish whether institutional rules exist or how they are implemented. Beyond these reported challenges, responses to other ethics and policy items provide additional context: 71.9% agreed that AI could be used ethically with appropriate guidance (M = 4.02, SD = 0.88), and 69.1% indicated willingness to declare AI use if required (M = 3.93, SD = 0.90) (Table 1). In the optional recommendations, one student proposed "clear regulations on the scope of AI use" and instruction in "discipline-specific prompting," alongside oral examinations or draft comparison (J3-R25). Another described AI as "a supporting tool, not a source of truth" and emphasized student responsibility (J3-R28). Together, these findings identify three main areas of reported difficulty: evaluating the accuracy and appropriateness of AI-generated content, using AI effectively for discipline-specific writing, and understanding ethical and institutional expectations. Students' recommendations further highlight their perceived need for clearer guidance and support.

Table 1

Descriptive Summary of Likert-Scale Sections

Section	Items	Valid *N*	Agreement (%)	*M*	Cronbach's α
Reported AI writing activities (B)	6	118–123	68.0–76.2	3.84–4.04	.912
Perceived usefulness (C)	6	113–116	68.1–78.4	3.91–4.05	.920
Perceived ease of use (D)	5	114–116	66.7–79.1	3.78–3.99	.901
Conditional use and training (E)	4	112–116	67.2–81.0	3.90–4.12	.882
Genre support (F)	5	114–116	67.5–71.6	3.81–3.91	.910
Feedback and confidence (G)	6	112–116	75.0–86.2	3.96–4.10	.908
Challenges and AI literacy (H)	6	119–123	59.3–72.4	3.63–3.95	.891
Ethics and policy (I)	7	120–123	53.3–71.9	3.52–4.02	.895

Note. Agreement is the combined percentage of valid responses at levels 4 and 5. For Parts C–G, the seven respondents with no prior AI use were excluded. Valid *N* varies slightly because item-level multiple selections were treated as missing. *M* = mean. *SD* = standard deviation; item-level *SD* values are reported selectively in the Findings. The Part B items report writing-related uses; phase-specific descriptions come from the optional J1 responses. Part E includes a training-needs item that is not interpreted as a direct measure of continued-use intention.

5. Discussion

5.1 Reported writing purposes

In response to the first research question, students reported using generative AI to support several writing activities in ESP courses. The item on language checking received the highest level of agreement among the surveyed writing purposes, while open-ended accounts also referred to outlining, vocabulary searches, idea development, and draft review. This pattern is broadly

consistent with Nguyen et al. (2025), whose participants valued assistance with grammar, vocabulary, and writing organization. Do et al. (2026) likewise reported AI chatbot use at several points in the EFL writing process, particularly before and after writing. The studies are not directly comparable in their measures: the present Likert items address purposes, while the optional comments illustrate when individual students used AI. Consequently, the present data cannot establish which writing phase was most frequent.

The resemblance across these studies concerns students' accounts of assistance, not demonstrated improvement in writing. Song and Song (2023) examined writing outcomes in an intervention, whereas this survey contains neither assessed writing products nor an intervention comparison. The reported use of AI for language checking and feedback may therefore inform support for academic writing in ESP courses, but it should not be interpreted as evidence of changes in students' writing ability.

5.2 Perceived usefulness, ease of use, and future use intentions

The second research question concerns usefulness, ease of use, and future use intentions, alongside writing-specific support. Positive ratings for task efficiency and general ease of use are compatible with the perceived advantages reported in EAP and EFL settings (Kohnke, 2024; Nguyen et al., 2025). However, this study also identifies perceived assistance with examples of academic sections, genre organization, feedback, and the expression of academic ideas. These writing-related perceptions belong to the discussion of usefulness because ESP writing requires attention to rhetorical purpose and discipline-sensitive expression, not merely grammatical accuracy (Bhatia, 1993; Swales, 1990). The survey records students' judgments of such assistance. It does not test whether exposure to AI-generated examples improved genre knowledge.

The conditional intention to use AI more often under clearer guidance (78.1%) differs from the request for further training (81.0%). The first item concerns anticipated use under a stated condition; the second indicates a perceived support need. Kılıç and Çelik (2025) examined technology acceptance among instructors using an extended TAM model, but their explanatory findings cannot be transferred to the present student sample. Here, TAM provides terms for describing perceptions rather than a model tested through associations or causal paths. More specifically, the interest in genre examples suggests a classroom question: whether students can evaluate an AI-generated structure against an authentic academic model before adapting it for their own writing. This is a possible pedagogical response, not an outcome measured by the survey.

The interpretation requires caution because the evidence concerns perceptions and reported practices, not observed writing behavior. AI-generated examples may provide material for examining academic genres, but the survey does not show whether students learned from those examples. One possible instructional response is to compare AI-generated sections with authentic academic texts, discuss rhetorical choices and terminology, and ask students to revise the suggested wording for a specific writing purpose. This is a pedagogical implication rather than a tested intervention.

5.3 Reported challenges in using generative AI for academic writing in ESP courses

The third research question concerns students' reported difficulties in checking AI-generated information, choosing discipline-appropriate terminology, developing prompts, and interpreting acceptable use. Similar concerns about reliance and reliability appear in EAP research (Kohnke, 2024) and in work with Vietnamese English-major students (Nguyen et al., 2025). Do et al. (2026) also documented concerns about AI-output quality, ethics, and support in EFL self-regulated writing. These studies provide points of comparison, but their samples and instruments differ. The present results describe endorsement of specific survey items and open-ended accounts within ESP-oriented academic writing; they do not measure actual AI errors or students' AI-literacy competence.

Uncertainty about acceptable assistance is another relevant concern. Barrett and Pack (2023) reported differing judgments about appropriate AI use across writing activities. Some respondents in the present study perceived institutional rules as unclear or inconsistent, but their answers cannot establish whether rules exist or how they are implemented. Their requests for clearer boundaries and training are consistent with the support needs discussed by Do et al. (2026). For ESP instructors, possible responses include source verification, comparing AI-generated and authentic genre examples, checking specialized terminology, and specifying when AI use should be disclosed. These suggestions follow from reported needs; their effectiveness remains untested in this study.

6. Pedagogical implications and conclusion

The findings point to several cautious implications for ESP language education. At the curriculum level, AI literacy could be integrated into ESP instruction as a formal learning component. Such integration may include basic knowledge of how generative AI works, its common limitations, discipline-sensitive prompting, evaluation of terminology and factual accuracy, and ethical reflection on when and how AI support should be declared. One possible reference point is the AIAP framework proposed by Ngo and Hastie (2025), which would need to be adapted to Vietnamese classroom conditions.

At the classroom level, AI is perhaps best treated as an object of scrutiny rather than simply a shortcut to finished work. Students might set AI-generated abstracts alongside authentic ones, weigh their relative strengths and weaknesses, check terminology for accuracy, and then rework the drafts into their own voice. Assessment could also move beyond the finished product itself, incorporating process-oriented tasks such as brief oral explanations of a draft or reflective notes on how AI was used along the way. These suggestions respond to students' reported needs and concerns, but their effectiveness requires further investigation.

At the institutional level, the findings point to a need for clearer guidance. Policy should spell out what forms of AI use are permitted and what forms are not, the conditions under which such use must be declared, and where responsibility lies for both students and instructors. Professional development for instructors could support the implementation of this guidance, particularly in task design, assessment of AI-assisted writing, and consistent responses to student use of AI.

The study has several limitations. The sample is restricted to 123 English-major students at a single university, so the findings cannot be generalized to all ESP settings. The data are self-reported and therefore reflect perceptions and reported practices rather than directly observed writing behavior or measured writing improvement. The study also relies on one questionnaire as its sole data-collection instrument. Section-level Cronbach's alpha coefficients ranged from .882 to .920, but internal consistency does not remove the limitations of self-report or establish construct validity on its own. A small proportion of eligible item-level records contained more than one Likert selection and were treated as missing for the affected item. Future research should use survey formats that enforce a single response for each Likert item and should combine questionnaires with interviews, classroom observations, writing samples, or longitudinal evidence. Within these limits, the study describes how students in the surveyed context report using and evaluating generative AI for academic writing in ESP courses.

References

Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, *20*, Article 59. https://doi.org/10.1186/s41239-023-00427-0

Belcher, D. D. (2006). English for specific purposes: Teaching to perceived needs and imagined futures in worlds of work, study, and everyday life. *TESOL Quarterly*, *40*(1), 133–156. https://doi.org/10.2307/40264514

Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Longman. https://www.taylorfrancis.com/books/mono/10.4324/9781315844992/analysing-genre-bhatia

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

Chen, C., & Gong, Y. (2025). The role of AI-assisted learning in academic writing: A mixed-methods study on Chinese as a second language students. *Education Sciences*, *15*(2), 141. https://doi.org/10.3390/educsci15020141

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, *55*(6), 2503–2529. https://doi.org/10.1111/bjet.13460

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Derakhshan, A. (2025). Tracking the effects of Gemini as a GenAI tool on L2 learners' writing proficiency and anxiety: Latent growth curve modelling approach. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-025-01042-5

Do, H. M. L., Nguyen, T. T. N., Nguyen, L. M. T., Nguyen, L. T. L., & Cao, T. X. L. (2026). Using AI chatbots for self-regulated writing: EFL students' practice, perceptions, and support needs. *Journal of Inquiry into Languages and Cultures*, *10*(1), 1–17. https://doi.org/10.63506/jilc.1001.414

Dudley-Evans, T., & St. John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press. https://assets.cambridge.org/97805215/96756/excerpt/9780521596756_excerpt.pdf

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*, article 100055. https://doi.org/10.1016/j.caeai.2022.100055

Ho, T. Q. N. (2017). English learning needs of business administration students: Factors to consider in ESP curriculum development. *Journal of Inquiry into Languages and Cultures*, 1(1), 59–71. https://jilc.vn/index.php/jilc/article/view/260

Hutchinson, T., & Waters, A. (1987). *English for specific purposes*. Cambridge University Press. https://www.academia.edu/35231434/English_for_Specific_Purposes_A_learning_centred_approach

Kılıç, E., & Çelik, H. E. (2025). Factors affecting acceptance of ChatGPT-4o by English language instructors: The extended TAM approach. *Computers and Education: Artificial Intelligence*, *9*, article 100452. https://www.sciencedirect.com/science/article/pii/S2666920X2500092X

Kohnke, L. (2024). Exploring EAP students' perceptions of GenAI and traditional grammar-checking tools for language learning. *Computers and Education: Artificial Intelligence*, *7*, article 100279. https://doi.org/10.1016/j.caeai.2024.100279

Lai, T. K. L., & Trinh, H. L. (2025). The effects of ChatGPT on academic writing proficiency among English-majored undergraduates at a university. *VNU Journal of Foreign Studies*, *41*(3), 92–109. https://doi.org/10.63023/2525-2445/jfs.ulis.5514

Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence*, *9*, article 100445. https://doi.org/10.1016/j.caeai.2025.100445

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). ACM. https://doi.org/10.1145/3313831.3376727

Mo, Z., & Crosthwaite, P. (2025). Exploring the affordances of generative AI large language models for stance and engagement in academic writing. *Journal of English for Academic Purposes*, *75*, article 101499. https://doi.org/10.1016/j.jeap.2025.101499

Ngo, T. N., & Hastie, D. (2025). Artificial intelligence for academic purposes (AIAP): Integrating AI literacy into an EAP module. *English for Specific Purposes*, *77*, 20–38. https://doi.org/10.1016/j.esp.2024.09.001

Nguyen, T. C. N. (2017). *Aligning English for specific purposes (ESP) curriculum with industry needs: Language practices for Vietnam's globalised workplaces* [Doctoral dissertation, Queensland University of Technology]. https://eprints.qut.edu.au/110536/1/Thi%20Chau%20Ngan__Thesis.pdf

Nguyen, T. H. H., Pham, T. Q. H., Nguyen, T. H., & Le, T. N. D. (2025). Ứng dụng ChatGPT khi viết học thuật của sinh viên ngành Ngôn ngữ Anh trong các học phần giảng dạy bằng tiếng Anh: Nghiên cứu tại Đại học Kinh tế Quốc dân. *Tạp chí Giáo dục*, *25* (số đặc biệt 10), 406–411. https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/4566

Pennycook, A. (1997). Vulgar pragmatism, critical pragmatism, and EAP. *English for Specific Purposes*, *16*(4), 253–269. https://doi.org/10.1016/S0889-4906(97)00019-7

Sanz-Tejeda, A., Domínguez-Oller, J. C., Baldaquí-Escandell, J. M., Gómez-Díaz, R., & García-Rodríguez, A. (2026). The impact of generative AI on academic reading and writing: A synthesis of recent evidence (2023–2025). *Frontiers in Education*, *10*, article 1711718. https://doi.org/10.3389/feduc.2025.1711718

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press. https://archive.org/details/genreanalysiseng0000swal

Tran, T. N. M. (2025). Factors affecting the use of ChatGPT in academic writing perceived by English-majored students. *VNU Journal of Foreign Studies*, *41*(3), 110–125. https://doi.org/10.63023/2525-2445/jfs.ulis.5519

Wu, C., Moorhouse, B. L., Wan, Y., & Wu, M. (2026). Exploring PhD students' utilization of generative AI in academic writing for publication purposes: Insights for EAP. *Journal of English for Academic Purposes*, *79*, article 101612. https://doi.org/10.1016/j.jeap.2025.101612