العدد 4 No 4
تموز 2026 July 2026
مجلة ساوة للدراسات الانسانية
Sawah Journal of Humanities
E-ISSN 3106 - 8324 P-ISSN 3106 - 8316
مجلة محكمة تصدر عن
جامعة ساوة الاهلية

ChatGPT as Writing Tool to Assist Iraqi EFL University Students based on their Perspectives: Opportunities, Risks, and Pedagogical Implications

A.LAbbas Fadhil Abdali
Imam AlKadhim University College, Iraq.
abbas.fadhil@iku.edu.iq

Abstract

The present study is an attempt to explore the perceptions of Iraqi EFL students toward the use of ChatGPT as a writing tool. It explores ChatGPT's perceived benefits, its risks, the usage patterns, ethical dimensions, and the pedagogical and institutional support required for students to make its use responsible. The study employed a quantitative cross-sectional survey design. The data were collected from 94 under graduate students who are enrolled in different Iraqi universities during the second semester of the academic year 2026. The data were collected through a 50-item Likert-scale questionnaire. The questionnaire was organized into five sections and was designed via Google Forms. The collected data were analysed using descriptive statistics, Pearson Correlation, and multiple regression in IBM SPSS 27. The data demonstrated excellent internal consistency across all sections. The findings show that students perceive ChatGPT as a genuinely transformative and multi-dimensional writing resource, with very high endorsement across all ten benefit dimensions, particularly idea generation, academic style improvement, and writing confidence. Risk perceptions displayed a critical bifurcation: students systematically rejected self-referential behavioural risks such as over-reliance and uncritical acceptance, while strongly acknowledging systemic risks including misinformation, reference fabrication, and critical thinking attenuation. Most consequentially, the study documents a severe integrity awareness-behaviour paradox, whereby near-universal acknowledgement of academic dishonesty coexisted with equally near-universal self-reported engagement in integrity-compromising behaviours, including paraphrasing to evade detection and submitting unverified AI-generated references. ChatGPT use was found to be intensive, multi-purpose, and structurally embedded. Despite this, students strongly endorsed structured pedagogical intervention, calling for teacher guidance, curriculum-integrated AI literacy, assessment reform, and clear institutional policy. The study concludes that prohibition-based strategies are empirically untenable and argues for an urgent, contextually grounded, integration-oriented framework tailored to the specific linguistic, institutional, and cultural realities of Iraqi EFL higher education.

Keywords: ChatGPT, EFL writing, Iraqi university students, academic integrity, AI literacy, pedagogical implications.

استخدام جات جي بي تي كأداة كتابة مساعدة لطلبة الجامعات العراقيين الذين يدرسون اللغة الإنكليزية كلغة اجنبية بناء على تصوراتهم :الفرص , المخاطر , والاثار التربوية

م.م عباس فاضل عبدعلي
كلية الامام الكاظم (ع) للعلوم الإسلامية الجامعة, العراق

العدد 4 No 4
تموز 2026 July 2026
مجلة ساوة للدراسات الانسانية
Sawah Journal of Humanities
E-ISSN 3106 - 8324 P-ISSN 3106 - 8316
مجلة محكمة تصدر عن
جامعة ساوة الاهلية

الملخص
تعد الدراسة الحالية محاولة لاستكشاف تصورات الطلبة العراقيين الذين يدرسون اللغة الإنكليزية كلغة اجنبية اتجاة استخدام جات جي بي تي كأداة مساعدة في الكتابة ,وتبحث الدراسة في الفوائد المتصورة للاداة ,وانماط استخدامها ,وابعادها الأخلاقية ,بالإضافة الى الدعم التربوي والمؤسسي المطلوب لضمان استخدامها بشكل مسؤول من قبل الطلبة. اعتمدة الدراسة على منهج مسحي كمي مقطعي حيث جمعت البيانات من 94 طالبا من طلبة الدراسات الأولية المسجلين في جامعات عراقية مختلفة خلال الفصل الدراسي الثاني من العام الدراسي 2026. تم جمع البيانات عبراستبانة مصممة بوانماذج كوكل فوم ومكونة من 50 فقرة وفق مقياس ليكرت حيث وزعت الاستبانة على خمسة اقسام . تم معالجة البيانات المجموعة احصائيا باستخدام الإحصاء الوصفي ومعامل ارتباط بيرسون والانحدار المتعدد عبر برنامج الحزم الإحصائية للعلوم الاجتماعية. وقد اضهرت البيانات اتساقا داخليا ممتازا عبر جميع الأقسام. اضهرت النتائج ان الطلبة ينظرون الى جات جي بي تي كأداة تحولية حقيقية ومتعددة الابعاد مع تاييد واسع جدا عبر جميع ابعاد الفائدة العشرة لا سيماء في مجالات توليد الأفكار ,تحسين الأسلوب الاكاديمي وزيادة الثقة عندة الكتابة . وفيما يتعلق بادراك المخاطر ,كشفت النتائج عن انقسام جوهري حاد فقد رفض الطلبة بشكل قاطع المخاطر السلوكية المتعلقة بذواتهم (مثل الاعتماد المفرط والقبول غير النقدي للمخرجات) في حين أقروا بشدة بالمخاطر المنهجية العامة بما في ذالك: التضليل واختلاف المصادر والمراجع واضعاف التفكير النقدي . والنتيجة الأكثر أهمية وتاثير في هذه الدراسة تمثلت في رصد "مفارقة الوعي والسلوك" الحاد فيما يخص النزاهة الاكاديمية حيث تزامن الإقرار شبه الجماعي بمفهوم الأمانة العلمية مع إقرار شبه جماعي أيضا بالانخراط الشخصي في سلوكيات تخل بهذه النزاهة مثا إعادة الصياغة للتهرب من برنامج كشف الانتحال وتقديم مصادر ومراجع مولدة بواسطة الذكاء الاصطناعي دون التحقق منه. كما بينت الدراسة ان استخدام جات جي بي تي كان مكثفا ومتعدد الأغراض ومتجذرا بنيويا في ممارسات الطلبة .ورغم ذالك ايد الطلبة بقوة ضرورة التدخل التربوي المنظم مطالبين بتوجيه من الأساتذة وادراج الثقافة الرقمية للذكاء الاصطناعي ضمن المناهج الدراسية وإصلاح اليات التقييم والاختبار وصياغة سياسة مؤسسية واضحة. تلخص الدراسة على ان الاستراتيجيات القائمة على المنع والحضر غير قابلة للتطبيق عمليا من الناحية التجريبية وتدعو بدلا من ذالك الى ضرورة تبني اطار عمل عاجل وموجه نحو الدمج ومبني على الواقع السياقي بحيث يتناسب مع المعطيات اللغوية والمؤسسية والثقافية الخاصة بالتعليم العالي لطلبة اللغة الإنكليزية كلغة اجنبية في العراق .

الكلمات المفتاحية : جات جي بي تي , كتابة اللغة الإنكليزية كلغة اجنبية , طلبة الجامعات العراقيين , النزاهة الاكاديمية , الوعي بالذكاء الاصطناعي , الاثار التربوية.

Introduction

One of the AI tools is ChatGPT that was innovated and developed by OpenAI. ChatGPT has been widely used by university students worldwide as a transformative tool due to its capability of providing a large language model (Yu, 2024; Aljohani, 2026). This language model is capable of providing grammatically accurate text and can engage in sophisticated natural language dialogue in addition to providing real-time feedback, assisting with vocabulary and wide range of writing tacks, answering questions, and generating organisation ideas. All of these features, in addition to being for free and available at any time, made ChatGPT tool appealing for teachers, researchers, and learners of English as a foreign language (henceforth EFL) as indicated by Song & Song (2023) and Miri (2025). However, there have been some concerns

about its pedagogical implications (Rudolph, Tan, & Tan, 2023; Kasneci, Seßler, Küchemann, Bannert, Dementieva, Fischer, & Kasneci, 2023).

Iraqi EFL students have a constantly challenging job when doing writing tasks, as writing is one of the most challenging and cognitively demanding skill that is difficult to master, particularly in academic writing tasks. With the introduction of ChatGPT, which provides the required skills, including simultaneous mastery of grammar, competent use of vocabulary, coherent and cohesive text and argumentation, discourse organization, and other requirements to adhere to the scholarly academic writing conventions, EFL students find it useful to provide a professionally academic writing (Altamimi, F., & Hussein, 2025; Khampusaen, 2025). This challenging difficulty for Arab EFL students arise from the major differences between Arabic and English, including grammar, syntax, and other rhetorical conventions of Academic writing (Altamimi & Hussein, 2025). According to Miri (2025) and Muhammad & Hasan (2026), Iraqi EFL students have a further complicated situation as they have limited exposure to authentic English and study in under-resourced institutions that use traditional pedagogical methods resulting in underdeveloped writing competence. Due to this disadvantageous situation, significant number of Iraqi EFL students resort to the use of ChatGPT to achieve academic and other writing tasks, most often without any sort of guidelines by their teachers about the use of such AI tools effectively and yet responsibly (Jasim, 2025).

Against this unfortunate and acute situation and within this context, EFL students are enthusiastically attracted to the emerging AI tool of ChatGPT, which is being considered by academic scholars as a tool that outpaced both research and policy. However, there have been a growing body of international empirical studies that reported positive results of using ChatGPT in writing tasks. These studies claim that ChatGPT help EFL learners in generating ideas, correcting grammatical errors, providing real-time personalized feedback, paraphrasing complex passages, enriching vocabulary, structuring essay, and text organisation of any academic writing task (Aljohani, 2026; Miri, 2025; Altamimi & Hussein, 2025). The overall conclusion of studies conducted within the Iraqi context reveal that most of EFL students hold positive perspectives about the use of ChatGPT as it tackles and address the common weaknesses that they face in academic writing (Miri, 2025). Similarly, several empirical studies that were conducted in diverse contexts reported that EFL students' writing significantly improved after integrating ChatGPT in their writing tasks, specifically in text organization, coherence, grammar, and confidence (Karimi & Qadir, 2025; Tatar & Brime, 2025; Jasim, 2025).

However, the same studies have raised some pedagogical and ethical concerns due to the fact that most of EFL students use this AI tool extensively, resulting in a collection of writing content that lacks depth, originality, and genuine argumentation. According to Miri (2025), the majority of students who used ChatGPT in their academic writing tasks provided fabricated list of references,

incomplete citations, wrong attribution of the text, and in some cases entirely invented citation. This claim raises serious concerns about academic integrity as a global issue in general, and in Iraq in particular, which does not have any published guidelines or systematic policies with regard to the use of AI tools in academic writing (Bin-Nashwan, Sadallah, & Bouteraa, 2023; Cotton, Cotton, & Shipway, 2024). Critics also express their concern over the negative effect of such AI tools like ChatGPT in undermining students' critical and independent thinking skills and creativity (Xu, 2020; Zhang & Aslan, 2021). Hence, ChatGPT, being a double-edged sword, is under scrutiny to examine its role in learning contexts and the problems it poses based on students' experiences and perspectives.

Research Questions

This study attempts to answer the following questions:

1. What do Iraqi EFL university students perceive as the main benefits of using ChatGPT as a writing tool, particularly in relation to grammar, vocabulary, organisation, and writing confidence?
2. What risks and challenges do Iraqi EFL university students identify in their own use of ChatGPT for academic writing, including concerns about over-reliance, critical thinking, and academic integrity?
3. How do demographic and experiential factors — specifically gender, academic year, and prior experience with AI tools — influence Iraqi EFL students' perceptions of and engagement with ChatGPT as a writing assistant?
4. What pedagogical approaches and institutional measures do Iraqi EFL students' perspectives suggest are needed to support the responsible and effective use of ChatGPT in university writing instruction?

Objectives of the Study

The present study is guided by the following objectives:

1. To identify Iraqi EFL university students' perceptions of the main benefits of using ChatGPT as a writing tool, with particular reference to grammatical accuracy, vocabulary enrichment, essay organisation, and writing confidence.
2. To examine the risks and challenges that Iraqi EFL university students associate with their own use of ChatGPT in academic writing, including concerns related to over-reliance, critical thinking attenuation, reference fabrication, and academic integrity.
3. To explore the patterns, frequency, and purposes of ChatGPT use among Iraqi EFL undergraduate students, and to determine the extent to which use is intensive, multi-purpose, and structurally embedded in academic writing routines.
4. To investigate Iraqi EFL students' ethical awareness of AI-assisted writing and to analyse the relationship between their declared ethical knowledge and their self-reported writing behaviours.

5. To determine how demographic and experiential variables — specifically gender, academic year, and prior AI experience — moderate Iraqi EFL students' perceptions of and engagement with ChatGPT as a writing assistant.
6. To identify the pedagogical approaches and institutional measures that Iraqi EFL students' own perspectives suggest are necessary to support the responsible, effective, and educationally productive integration of ChatGPT in university writing instruction.

Statement of the Problem

Despite the rapid and widespread adoption of ChatGPT among university students globally, its integration into Iraqi EFL academic writing contexts remains largely unguided, unregulated, and empirically underexplored. Iraqi EFL students face a particularly acute set of writing challenges arising from limited exposure to authentic English, structurally under-resourced institutions, traditional pedagogical environments that offer minimal writing scaffolding, and the significant linguistic and rhetorical distance between Arabic and English. In this context, ChatGPT has emerged as a de facto writing support system that students adopt extensively and often exclusively, without any institutional framework governing its use or any pedagogical structure shaping how it is engaged. The absence of systematic AI policy infrastructure in Iraqi universities, confirmed by Bin-Nashwan et al. (2023), means that students are navigating the opportunities and risks of this powerful tool entirely on their own terms, with predictable consequences for academic integrity, critical thinking development, and genuine writing competence. While a growing body of international empirical literature has examined ChatGPT's role in EFL writing instruction across diverse contexts — including China, Iran, Saudi Arabia, and Kurdish-administered Iraq — no comprehensive, student-centred study has yet provided a holistic picture of ChatGPT use as a writing tool specifically among Arabic-speaking Iraqi EFL undergraduates, addressing simultaneously the benefits students perceive, the risks they acknowledge, the patterns and depth of their reliance, their ethical awareness and behaviour, and the pedagogical conditions they identify as necessary for responsible use. This gap constitutes both a scholarly deficit and an urgent practical problem, given that the absence of contextually grounded evidence leaves Iraqi higher education institutions without the empirical basis needed to develop effective, locally adapted policies and pedagogical responses to a phenomenon that is already deeply embedded in students' academic writing practices.

Significance of the Study

This study makes several contributions of scholarly and practical significance at the local, regional, and international levels.

At the scholarly level, it addresses a clearly identified gap in the existing literature by providing the first comprehensive, student-centred, quantitative investigation of ChatGPT use as a writing tool among Arabic-speaking Iraqi EFL undergraduates. While prior studies have examined ChatGPT perceptions

in a range of EFL contexts, the Iraqi context presents a distinctive combination of linguistic, institutional, and cultural conditions — including the structural distance between Arabic and English rhetorical conventions, the under-resourcing of higher education institutions, and the complete absence of published AI policy guidelines — that cannot be adequately addressed by findings generated in other settings.

At the pedagogical level, the study provides writing instructors in Iraqi universities with empirical evidence directly relevant to their classroom practice. The findings offer a detailed, data-grounded account of what students find genuinely helpful, where their self-regulatory capacities break down, and what forms of teacher guidance and task design they identify as most needed.

At the institutional and policy level, the study provides Iraqi higher education decision-makers with the kind of student-centred empirical evidence that is a prerequisite for informed, effective, and contextually appropriate AI policy development.

Finally, at the regional and international level, the study contributes to the broader comparative literature on ChatGPT in EFL higher education by documenting the specific conditions under which AI integration risks are most acute — namely, contexts combining acute student writing challenges, limited institutional infrastructure, and the absence of policy governance — thereby enabling more precise, contextually differentiated conclusions about when, how, and under what conditions ChatGPT's educational potential can be responsibly and equitably realised.

Delimitation of the Study

The scope of the present study is delimited in the following respects, each of which should be considered when interpreting its findings and assessing their generalisability.

First, the study is geographically and institutionally delimited to undergraduate students enrolled in English departments at a selection of Iraqi universities during the second semester of the 2026 academic year. Although participants were drawn from multiple institutions to enhance representativeness within the Iraqi context, the findings cannot be generalised to EFL students in other Arab countries, other national higher education systems, or other disciplines beyond English language and literature.

Second, the study is disciplinarily delimited to EFL students in English departments.

Third, the study is methodologically delimited to a quantitative, cross-sectional survey design.

Fourth, the study is demographically delimited to second-, third-, and fourth-year undergraduate students, selected purposively to ensure adequate English language proficiency and familiarity with AI tools. First-year students, postgraduate students, and students with no prior AI experience are not

represented in the sample, which limits conclusions about the full developmental range of ChatGPT engagement in Iraqi EFL higher education.

Fifth, the study is temporally delimited to a single academic semester, and its findings reflect students' perceptions and self-reported behaviours during a specific and rapidly evolving period of AI tool development.

2. Literature Review

The release of ChatGPT in late 2022 marked a pivotal moment in this evolution, attracting over one hundred million users within two months and establishing itself as the dominant AI writing tool among EFL learners globally (Guo et al., 2025; Mali, 2025). Unlike earlier tools, ChatGPT can engage in interactive dialogue about writing, propose alternative formulations, explain grammatical rules, and scaffold the organisational structure of academic papers, with complementary tools such as Grammarly and QuillBot typically serving secondary roles (Aljohani, 2026; Muhammad & Hasan, 2026). Adoption rates in EFL higher education contexts are notably high: Muhammad and Hasan (2026) found that 86.6% of undergraduate EFL students at the University of Zakho were familiar with AI tools and 85.6% actively used them, while Miri (2025) documented broad positive perceptions among six hundred learners at the University of Basrah. The most commonly reported applications include idea generation, grammar correction, vocabulary enrichment, essay structuring, and research organisation, with academic writing emerging as the primary focal point across systematic reviews of the field (Aljohani, 2026; Mali, 2025).

Several theoretical frameworks have been applied to understand ChatGPT's role in EFL writing instruction. Vygotsky's Sociocultural Theory of Learning, and particularly the concept of the Zone of Proximal Development, is the most widely invoked, with scholars positioning ChatGPT as a mediating scaffold capable of supporting learners through writing tasks beyond their independent capacity — analogous to the assistance of a more knowledgeable peer — while cautioning that this scaffolding loses its pedagogical value when it substitutes for rather than supports the learner's own cognitive effort (Altamimi, 2025; Gburi & Dowlatabady, 2025). Constructivist learning theory and Communicative Language Teaching principles have similarly been drawn upon, with Aljohani (2026) arguing that ChatGPT's capacity to facilitate authentic text-based communicative interaction in English aligns well with CLT's emphasis on meaningful language use. Mohammad et al. (2026), applying the Technology Acceptance Model, further demonstrate that perceived usefulness — evidenced through visible gains in linguistic accuracy and drafting support — is the primary driver of ChatGPT adoption among EFL teachers, suggesting that sustainable integration depends on actively managing both the tool's utility and its associated risks.

Quantitative studies such as Miri (2025) and Jasim (2025) report statistically significant gains across multiple writing criteria, while studies by Gburi and Dowlatabady (2025) and Mohammad et al. (2026) further corroborate these

findings among Iraqi and Saudi EFL populations. Beyond linguistic outcomes, learners also report significant affective benefits, including reduced writing anxiety and increased motivation, with many preferring the non-judgmental availability of AI feedback over direct teacher interaction (Muhammad & Hasan, 2026; Al-Obaydi et al., 2025).

However, the literature equally emphasises the limitations of ChatGPT in fostering deeper academic writing competencies. While AI-generated feedback proves effective for surface-level error correction, it consistently falls short in developing critical argumentation, evidence-based reasoning, and disciplinary genre awareness. Miri (2025) found that approximately 80% of AI-assisted student papers contained fabricated or incorrectly formatted references, and Waseem and Ali (2025) observed stagnation or decline in higher-order writing skills such as coherence and task fulfilment. Scholars broadly agree that ChatGPT feedback is best understood as complementary to, rather than a replacement for, teacher feedback, particularly for tasks requiring contextual depth and pedagogical scaffolding (Altamimi, 2025; Karimi & Qadir, 2025).

Muhammad and Hasan (2026) and Miri (2025) both document Iraqi EFL learners acknowledging that they sometimes substituted AI-generated content for independent thinking, while Mohammad et al. (2026) report that teachers observed students copy-pasting outputs without critical engagement, leading to what was described as a "fossilisation" of writing development. Kucuk (2025) found that over 75% of teachers and nearly 70% of students agreed that excessive AI reliance could negatively affect independent critical thinking. Karimi and Qadir (2025) offer a nuanced qualification, noting that while most students denied outright dependency, 93% acknowledged some degree of reliance on AI for writing feedback, suggesting dependency exists on a continuum rather than as a binary condition.

Closely related to over-reliance is ChatGPT's contested impact on critical thinking. A systematic review of sixty-five empirical studies indicates that while AI can stimulate initial idea generation, wider evidence points toward reduced student incentive to critically evaluate sources and engage in sustained complex reasoning. Tatar and Brime (2025), examining EFL instructors at Salahaddin University-Erbil, found that critical thinking benefits only emerged when AI was embedded in deliberately designed inquiry tasks requiring students to evaluate or challenge AI outputs, whereas passive use yielded no discernible improvement and in some cases a deterioration in analytical engagement. For Iraqi EFL learners specifically, Jasim (2025) identified weak critical engagement with AI suggestions as the principal limitation of ChatGPT-assisted writing, despite the statistically significant improvements in surface-level writing quality recorded.

The tension between AI-assisted support and learner autonomy is theorised through Vygotsky's Zone of Proximal Development, with Altamimi (2025) arguing that ChatGPT functions as scaffolding that must be deliberately

withdrawn as learners develop independence — a withdrawal that requires intentional pedagogical design rather than occurring automatically. Al-Obaydi et al. (2025) demonstrated empirically that students who engaged with ChatGPT as an interactive guide achieved considerably better outcomes than those who used it solely as a content source, underscoring that the manner of use — rather than use itself — determines its autonomy-building potential.

Academic integrity represents another critical dimension of concern. Miri's (2025) document analysis of one hundred and fifty Iraqi EFL student papers found that approximately 80% of AI-generated references were incomplete, wrongly formatted, or entirely fabricated — findings consistent with broader research estimating that ChatGPT fabricates around 20% of academic citations outright (Cheng, Nagesh, Eller, Grant, & Lin, 2024). Mali (2025) found that 36% of students used ChatGPT to complete part of their assignments and 18% submitted entirely AI-generated work as their own, creating profound challenges for educators. Despite these documented risks, student attitudes toward ethical AI use are often rationalising in nature: Miri (2025) found that Iraqi EFL learners tended to frame their own use as ethical while acknowledging behaviour that contradicted this self-assessment, a pattern identified as a widespread psychological mechanism for reconciling convenient AI use with academic honesty.

Demographic variables — particularly gender, academic level, and prior AI experience — emerge as significant moderators of ChatGPT engagement. Miri (2025) found that female Iraqi EFL students reported slightly higher perceptions, frequency of use, and satisfaction than male counterparts, while senior students demonstrated more critical and selective AI use than first-year students, who were simultaneously the most frequent users and the most vulnerable to dependency.

Scholars broadly advocate for a complement-rather-than-replace paradigm in which ChatGPT supplements rather than substitutes for teacher feedback and direct writing instruction (Jasim, 2025; Mohammad et al., 2026; Waseem & Ali, 2025). Process-oriented assessment, prompt engineering training, and the co-creation of ethical use guidelines between teachers and students are among the most consistently recommended strategies (Mali, 2025; Mohammad et al., 2026). Critically, the literature emphasises that any effective framework for Iraqi higher education must be locally adapted to the specific institutional, cultural, and linguistic realities of Iraqi EFL learners, given the documented limitations of AI tools in accommodating Arabic rhetorical conventions and the current absence of institutional AI policy infrastructure in Iraqi universities (Tatar & Brime, 2025; Altamimi, 2025).

For Iraqi EFL learners — who face heightened writing challenges arising from limited authentic English exposure, under-resourced institutions, traditional pedagogical cultures, and the structural distance between Arabic and English — ChatGPT holds real but unevenly realised potential. The present study seeks to

provide the kind of comprehensive, student-centred, contextually grounded evidence that is necessary to understand when, how, and under what institutional and pedagogical conditions this potential can be responsibly and equitably realised.

3. Methodology

3.1 Research Design

This study aims to explore the EFL Iraqi students' perceptions towards the use of ChatGPT as a potential writing tool. To meet this aim, this study adopted a quantitative, cross-sectional survey design, as it describes Iraqi EFL students' current perceptions of, engagement with, and concerns about ChatGPT as a writing tool (Creswell & Creswell, 2018; Cohen, Manion, & Morrison, 2018). This quantitative approach corresponds with the descriptive and comparative nature of the research questions. For this end, a questionnaire of five sections was designed for EFL Iraqi students from different Iraqi universities. Data was collected through this questionnaire that was designed by Google Forms and the link was sent to the participants through WhatsApp and Telegram Groups of EFL students. The collected data of participants' perceptions were scrutinized statistically and results were drawn based on these data.

3.2 Participants

The participants of the current study constituted of 94 students from different universities in Iraq during the second semester of the 2026 academic year. All of the participants students in the English Department in these universities and their academic level consisted of three levels, namely second year, third year, and fourth year students. The sample included both male and female students (male=54, female=40). The researcher used a purposive sampling to cover the research objectives of this study, by ensuring that the participants had the required level of English language proficiency and had the expected familiarity with using AI tools in obtaining feedback and comments. The demographic features of the participants in the present study are shown in Table 3.1.

Table 3.1. *Demographic Information of Participants*

Factor	Category	N
Gender	Male	54
	Female	40
English Level	2nd Year	25
	3rd Year	35
	4th Year	34
Prior AI Experience	No	
	If Yes, what is the duration	
	< 6 months	22
	Between 6–12 months	21
	> 12 months	51

Frequency of ChatGPT Use	Daily	38
	Several times a week	26
	Once a week	21
	Rarely	9
	Never	0

3.3 Research Instrument

The instrument used in data collection of this study was a questionnaire. The questionnaire was constructed by the researcher based on the research instruments used by several scholars to address the study's stated question (Miri, 2025; Nazim et al., 2025; Muhammad & Hasan, 2026; Aljohani, 2026; Jasim, 2025; Mohammad et al., 2026; Karimi & Qadir, 2025; Mali, 2025; Gburi & Dowlatabady, 2025; Tatar & Brime, 2025; Farrah et al., 2025; Al-Obaydi et al., 2025; Altamimi, 2025; Song & Song, 2023; Karimi & Qadir, 2025; Al-Obaydi et al., 2025; Mirzayev et al., 2025; Kucuk, 2025). There are two sections of the questionnaire. The first section sought demographic data. The second section utilized a Likert Scale (Strongly agree, agree, neutral, disagree, and strongly disagree) and was divided into five sections. The first part consisted of ten statements about the benefits of ChatGPT. The second section consisted of ten statements about the risks and challenges of using ChatGPT. The third section consisted of ten statements about academic integrity and ethical awareness. The fourth section sought answers to the usage and frequency patterns. The last section is about pedagogical needs and institutional support.

3.4 Data Collection

The data of the current study were collected through a five-section questionnaire that was administered online using Google Forms. The questionnaire link was distributed through the WhatsApp and Telegram groups of students in the English department from different universities. A brief introduction was given in the beginning of the survey to explain the purpose of the study. Data collection spanned over six weeks during the academic year of 2026.

3.5 Data Analysis

The process of data analysis consists of four stages: (1) reliability analysis to confirm internal consistency of all subscales; (2) descriptive statistics and frequency analysis to address Research Questions 1 and 2; (3) Pearson correlation analysis and multiple regression to address Research Question 4; and (4) a discussion synthesising the findings in relation to the relevant theoretical and empirical literature. All these analyses were conducted using IBM SPSS Statistics Version 27. Mean scores were interpreted using a five-point classification scheme adapted from Miri (2025) and Mohammad et al. (2026): 1.00–1.79 = Very Low (Strongly Disagree); 1.80–2.59 = Low (Disagree); 2.60–3.39 = Moderate (Neutral); 3.40–4.19 = High (Agree); 4.20–5.00 = Very High (Strongly Agree).

5. Results

Prior to the substantive analysis, Cronbach's alpha coefficients were computed for each of the five sections and for the full 50-item instrument using the SPSS Scale > Reliability Analysis procedure. Table 1 presents the results. The minimum acceptable threshold was set at $\alpha \geq 0.70$ in accordance with the conventions adopted in previous studies (Miri, 2025; Muhammad and Hasan, 2026).

Table 1. *Cronbach's Alpha Reliability Coefficients by Section*

Sec.	Section Name	k	α	r_it Range	Rating
A	Perceived Benefits	10	**0.991**	0.928–0.985	Excellent
B	Perceived Risks & Challenges	10	**0.972**	0.701–0.949	Excellent
C	Academic Integrity & Ethical Awareness	10	**0.990**	0.936–0.985	Excellent
D	Usage Patterns & Frequency	10	**0.985**	0.883–0.971	Excellent
E	Pedagogical Needs & Institutional Support	10	**0.991**	0.914–0.983	Excellent
A–E	Full Instrument (50 items)	50	**0.997**	—	Excellent

Note. k = number of items; r_it = corrected item-total correlation. Rating: ≥ 0.90 = Excellent. All sections substantially exceed the minimum threshold of α ≥ 0.70.

All five sections demonstrated excellent internal consistency, with alpha coefficients ranging from 0.972 (Section B) to 0.991 (Sections A and E), and the full 50-item instrument yielding $\alpha = 0.997$. No item was flagged for removal, as all corrected item-total correlations exceeded $r = 0.70$, substantially above the $r = 0.30$ retention threshold (Field, 2018). These reliability levels markedly exceed those reported in comparable studies — Miri (2025) reported $\alpha = 0.87$–0.93 and Muhammad and Hasan (2026) reported $\alpha = 0.84$–0.91 — confirming that the present instrument possesses strong psychometric properties and that responses across items within each section can be validly aggregated into composite scores for subsequent analysis.

4.1 Descriptive Statistics and Frequency Analysis

Descriptive statistics, including means, standard deviations, medians, skewness, and kurtosis coefficients, were computed for all 50 items and for section composite scores. Frequency distributions expressing the percentage of respondents endorsing each of the five Likert response categories were computed in parallel. Table 2 presents the section composite means and their classifications.

Table 2. *Section Composite Means, Standard Deviations, and Classifications*

Sec.	Section	M	SD	Classification
A	Perceived Benefits	**4.209**	0.098	Very High (Strongly Agree)
B	Perceived Risks & Challenges	**3.388**	1.028	Moderate → High (Neutral–Agree)
C	Academic Integrity & Ethical Awareness	**4.243**	0.098	Very High (Strongly Agree)
D	Usage Patterns & Frequency	**4.101**	0.189	High (Agree)
E	Pedagogical Needs & Institutional Support	**4.172**	0.111	High (Agree)

Note. Section means calculated as the mean of all ten item means within that section. SD = standard deviation of item means within section. Classification adapted from Miri (2025) and Mohammad et al. (2026).

The cross-sectional profile in Table 2 reveals a highly consistent pattern: Sections A, C, D, and E cluster tightly in the High to Very High classification range (M = 4.10–4.24), while Section B produces the most variable and distinctively lower composite mean (M = 3.39, SD = 1.03). This cross-sectional pattern is discussed in detail following the section-level analyses below.

4.1.1 Section A: Perceived Benefits of ChatGPT (RQ1)

Table 3. *Section A — Item-Level Descriptive Statistics and Frequency Distributions*

Item	Statement (abbrev.)	M	SD	Mdn	Skew	Kurt	SD %	D %	N %	A %	SA %
A1	ChatGPT helps correct	4.223	0.894	4.0	-1.751	4.055	3.2	2.1	5.3	47.9	41.5

	grammatical errors									
A2	ChatGPT helps expand/enrich vocabulary	4.170	1.033	4.0	-1.486	1.794	3.2	7.4	4.3	39.4	45.7
A3	ChatGPT helps organise and structure essays	4.287	0.812	4.0	-1.562	4.024	2.1	0.0	9.6	43.6	44.7
A4	ChatGPT improves coherence and cohesion	4.117	1.076	4.0	-1.295	1.013	3.2	8.5	7.4	35.1	45.7
A5	ChatGPT helps generate ideas and plan content	4.319	0.930	5.0	-1.831	3.818	3.2	2.1	6.4	36.2	52.1
A6	ChatGPT provides personalised writing feedback	4.223	0.974	4.0	-1.677	2.861	3.2	5.3	3.2	42.6	45.7
A7	ChatGPT increases writing confidence	4.298	0.760	4.0	-1.314	3.073	1.1	1.1	8.5	45.7	43.6
A8	ChatGPT improves academic style and tone	4.309	0.868	4.0	-2.060	5.581	3.2	2.1	1.1	47.9	45.7
A9	ChatGPT reduces writing anxiety	4.074	1.138	4.0	-1.490	1.716	7.4	2.1	9.6	37.2	43.6

A10	ChatGPT is available any time (convenience)	4.064	1.162	4.0	-1.512	1.566	7.4	5.3	3.2	41.5	42.6

Note. M = mean; SD = standard deviation; Mdn = median; Skew = skewness; Kurt = excess kurtosis; SD% = Strongly Disagree; D% = Disagree; N% = Neutral; A% = Agree; SA% = Strongly Agree. n = 94. Item A5 imputed for 3 missing values (mode = 5).

Section A yielded a composite mean of M = 4.209 (Very High/Strongly Agree), indicating that students overwhelmingly endorsed ChatGPT as a beneficial tool across all ten dimensions assessed. All individual items fell within the Very High classification range (M = 4.064–4.319), with uniformly negative skewness values confirming strong ceiling-effect endorsement patterns. Cronbach's alpha for the section was α = 0.991, indicating excellent internal consistency.

The most strongly endorsed benefit was item A5 (idea generation and content planning; M = 4.319, SD = 0.930), with 88.3% of respondents agreeing or strongly agreeing. This is closely followed by A8 (academic style and tone improvement; M = 4.309, SD = 0.868) and A7 (increased confidence in submitting written work; M = 4.298, SD = 0.760). The particularly low standard deviation for A7 (SD = 0.760) indicates unusually high consensus among respondents — only 2.2% expressed any disagreement. This pattern is consistent with Aljohani's (2026) finding that EFL learners identify confidence enhancement as a primary motivational driver of AI adoption, and with Miri's (2025) report that students at the University of Basrah placed particular value on the stylistic and structural affordances of ChatGPT in academic writing contexts.

The two lowest-rated items in the section — A10 (convenience and accessibility; M = 4.064, SD = 1.162) and A9 (anxiety reduction; M = 4.074, SD = 1.138) — remain solidly within the High/Very High range, but their comparatively larger standard deviations indicate greater individual variation. Notably, 7.4% of respondents strongly disagreed with both A9 and A10, the highest dissent rates in the section, suggesting that a small but non-trivial minority does not perceive the affective and logistical benefits as strongly as the majority. This variation may reflect heterogeneity in students' anxiety profiles and prior technology experience, consistent with Miri's (2025) finding that affective benefits were more pronounced among students with longer prior AI experience.

A subscale analysis further confirms the breadth of perceived benefit across distinct writing dimensions: surface-level linguistic benefits (A1–A4: M = 4.199, α = 0.980), cognitive and generative benefits (A5–A6: M = 4.271, α = 0.975), affective benefits (A7, A9: M = 4.186, α = 0.918), and academic style and access benefits (A8, A10: M = 4.186, α = 0.901). The consistently Very

High means across all four subscales confirm that students perceive ChatGPT as equally valuable across the full spectrum of writing support functions — from surface-level error correction to deeper compositional and affective dimensions. This finding is notable because it suggests that ChatGPT adoption among this sample is not driven by a single narrow utility but reflects a genuinely multi-dimensional perception of value, a pattern reported in comparable studies by Muhammad and Hasan (2026) and Gburi and Dowlatabady (2025).

4.1.2 Section B: Perceived Risks and Challenges (RQ2)

Table 4. *Section B — Item-Level Descriptive Statistics and Frequency Distributions*

Item	Statement (abbrev.)	M	SD	Mdn	Skew	Kurt	SD %	D %	N %	A %	SA %
B1	Rely on ChatGPT instead of independent thinking	2.319	1.338	2.0	0.687	-0.899	35.1	33.0	4.3	20.2	7.4
B2	ChatGPT reduces motivation to practise writing	1.628	0.842	1.0	1.573	2.775	54.3	34.0	7.4	3.2	1.1
B3	Accept ChatGPT suggestions uncritically	1.840	1.019	2.0	1.698	2.732	41.5	47.9	0.0	6.4	4.3
B4	ChatGPT writing doesn't sound like my own voice	4.106	1.010	4.0	-1.432	1.818	3.2	7.4	4.3	45.7	39.4
B5	ChatGPT gives inaccurate/misleading information	4.213	0.949	4.0	-1.676	3.124	3.2	4.3	4.3	44.7	43.6
B6	ChatGPT generates non-existent references	4.043	0.972	4.0	-1.164	1.165	2.1	7.4	9.6	45.7	35.1
B7	ChatGPT weakens critical thinking skills	4.064	1.014	4.0	-1.395	1.669	3.2	8.5	3.2	48.9	36.2
B8	ChatGPT	3.7	1.1	4.0	-	-	3.2	18.	6.4	45.	26.

	cannot capture Arabic-English conventions	45	35		0.785	0.394		1		7	6
B9	ChatGPT makes writing less original/creative	3.883	1.135	4.0	-0.984	-0.013	3.2	16.0	3.2	44.7	33.0
B10	ChatGPT makes it harder to write independently	4.043	0.961	4.0	-1.053	0.665	1.1	9.6	8.5	45.7	35.1

Note. Items B1–B3 exhibit marked positive skewness and very low means, reflecting students' rejection of self-referential behavioural risks. Items B4–B10 exhibit negative skewness consistent with high-endorsement agreement patterns. This bifurcation constitutes a key analytical finding.

Section B produced the most analytically complex profile of any section in the instrument. The composite section mean of M = 3.388 (Moderate-High boundary) masks a critical bifurcation between two sharply contrasting response patterns that carry distinct theoretical implications.

Items B1, B2, and B3 — which assess self-reported behavioural risks arising from personal over-reliance, motivational attrition, and uncritical acceptance of AI suggestions — yielded strikingly low means. Item B2 (ChatGPT reduces motivation to practise writing; M = 1.628, SD = 0.842) is the single lowest-scoring item across the entire 50-item instrument, with 88.3% of respondents disagreeing or strongly disagreeing. Similarly, B3 (accepting ChatGPT suggestions without critical evaluation; M = 1.840, SD = 1.019) was rejected by 89.4% of respondents, and B1 (over-reliance instead of independent thinking; M = 2.319, SD = 1.338) attracted 68.1% disagreement. The strong positive skewness of all three items (skewness = +0.687 to +1.698) confirms the systematic rejection of these self-referential risk perceptions.

This finding admits two complementary interpretations that must be considered together. First, it is possible that these students genuinely do not experience the motivational and dependency risks captured by B1–B3 — a finding that would suggest a relatively self-regulated user profile. However, given the simultaneously high endorsement of integrity-compromising behaviours in Section C (discussed below) and the very high reported usage frequency in Section D, a second interpretation is more plausible: these low scores reflect a well-documented social desirability bias in self-reporting of dependency-related behaviours (Karimi and Qadir, 2025). Karimi and Qadir (2025) found that while 93% of students acknowledged some degree of AI reliance for feedback, the majority denied outright dependency — a pattern consistent with the present findings and attributable to the psychological discomfort of self-identifying as dependent. This self-serving attribution is further theoretically grounded in

Miri's (2025) observation that Iraqi EFL learners systematically frame their own AI use as ethical even when reporting behaviours that contradict this self-assessment.

In sharp contrast, items B4–B10 — which assess externally observable and systemic risks rather than self-referential behavioural patterns — all fell in the High to Very High range (M = 3.745–4.213). The highest-endorsed risk was B5 (inaccurate or misleading information; M = 4.213, SD = 0.949), with 88.3% of respondents agreeing or strongly agreeing. This is closely followed by B4 (AI-generated text lacking authentic voice; M = 4.106), B7 (critical thinking attenuation; M = 4.064), and B10 (harder to write independently in exam conditions; M = 4.043). These high endorsement levels indicate that students are cognitively aware of the epistemic and developmental risks associated with ChatGPT use, even as they simultaneously deny exhibiting the personal behavioural patterns that generate those risks.

The lowest-scoring item among the high-risk cluster — B8 (ChatGPT's inability to capture Arabic-English rhetorical conventions; M = 3.745, SD = 1.135) — is also the most variable, with 18.1% of respondents disagreeing. This item-level heterogeneity likely reflects genuine variation in the extent to which students' writing tasks require engagement with distinctively Arabic rhetorical conventions, as well as differing levels of awareness of the structural divergences between Arabic and English academic discourse. The finding nonetheless aligns with Altamimi and Hussein's (2025) documentation of ChatGPT's systematic limitations in accommodating the genre and rhetorical expectations of Arabic-background EFL academic writing.

4.1.3 Section C: Academic Integrity and Ethical Awareness (RQ2)

Table 5. *Section C — Item-Level Descriptive Statistics and Frequency Distributions*

Item	Statement (abbrev.)	M	SD	Mdn	Skew	Kurt	SD %	D %	N %	A %	SA %
C1	Aware submitting AI text = academic dishonesty	4.319	0.930	5.0	-1.831	3.818	3.2	2.1	6.4	36.2	52.1
C2	Used ChatGPT for assignment without telling teacher	4.223	0.974	4.0	-1.677	2.861	3.2	5.3	3.2	42.6	45.7
C3	Submitted	4.29	0.76	4.0	-	3.07	1.1	1.1	8.5	45.	43.6

	mostly AI-generated writing as own	8	0		1.314	3				7	
C4	Used ChatGPT to paraphrase/avoid plagiarism detection	4.309	0.868	4.0	-2.060	5.581	3.2	2.1	1.1	47.9	45.7
C5	Included ChatGPT-generated references unverified	4.309	0.868	4.0	-2.060	5.581	3.2	2.1	1.1	47.9	45.7
C6	University has clear ChatGPT use guidelines	4.074	1.138	4.0	-1.490	1.716	7.4	2.1	9.6	37.2	43.6
C7	Teachers explained ethical/unethical AI use	4.064	1.162	4.0	-1.512	1.566	7.4	5.3	3.2	41.5	42.6
C8	Using ChatGPT for ideas is ethically acceptable	4.223	0.974	4.0	-1.677	2.861	3.2	5.3	3.2	42.6	45.7
C9	Universities should establish strict AI policies	4.298	0.760	4.0	-1.314	3.073	1.1	1.1	8.5	45.7	43.6
C10	Plagiarism tools cannot reliably detect AI content	4.309	0.868	4.0	-2.060	5.581	3.2	2.1	1.1	47.9	45.7

Note. Items C1–C5 and C10 are awareness and self-report items; items C6–C9 are institutional and attitudinal items. The parallel elevation of

awareness (C1) and transgressive behaviour (C2–C5) constitutes the integrity awareness-behaviour paradox central to this study.

Section C yielded the highest section composite mean of all five sections (M = 4.243, Very High/Strongly Agree), indicating exceptionally strong and consistent responses across all ten integrity-related items. However, the composition of this section requires careful disaggregation, as it includes both attitudinal awareness items (C1, C8, C9, C10) and direct self-report behavioural items (C2, C3, C4, C5). The simultaneous elevation of both item types constitutes the central integrity paradox of this study.

Item C1 (awareness that submitting AI-generated text constitutes academic dishonesty; M = 4.319, SD = 0.930) confirms that 88.3% of students agree or strongly agree with this ethical proposition. This is the highest awareness score in the section. Yet the behavioural items that follow immediately contradict this awareness: C3 (submitted mostly AI-generated writing as one's own; M = 4.298, 89.3% agree), C4 (used ChatGPT to paraphrase and evade plagiarism detection; M = 4.309, 93.6% agree), C5 (included unverified AI-generated references; M = 4.309, 93.6% agree), and C2 (used ChatGPT for assignment without informing teacher; M = 4.223, 88.3% agree) all indicate pervasive enactment of the very behaviours students acknowledge as ethically problematic. This awareness-behaviour gap is the most consequential finding of the study.

The theoretical framing most apt for this pattern is the normalisation hypothesis advanced by Miri (2025), whereby frequent exposure to AI-assisted academic work progressively reduces the perceived ethical cost of integrity-compromising behaviour even as declarative awareness of its impermissibility remains intact. From a social cognitive perspective, this gap may also reflect outcome expectancy: C10 (plagiarism detection tools cannot reliably detect AI-generated content; M = 4.309, 93.6% agree) suggests that students perceive the probability of detection as very low, thereby weakening the deterrent function of awareness. When students do not expect their behaviour to carry adverse consequences regardless of their declared ethical beliefs, awareness does not translate into behavioural compliance — a mechanism consistent with Cotton, Cotton, and Shipway's (2024) analysis of the structural incentives sustaining academic integrity violations in the ChatGPT era.

The perceived institutional policy deficit is captured by C6 (university has clear ChatGPT use guidelines; M = 4.074, 7.4% strongly disagree) and C7 (teachers have explained what constitutes ethical and unethical AI use; M = 4.064, 7.4% strongly disagree). While these items still yield overall High/Agree means — suggesting that the majority perceive some institutional guidance to be in place — the elevated disagreement rates relative to surrounding items, and the wider standard deviations (SD = 1.138 and 1.162 respectively), suggest meaningful heterogeneity in students' access to institutional ethical guidance. This institutional policy deficit may itself be a proximate enabling condition for the

transgressive behaviours documented in C2–C5, consistent with Bin-Nashwan, Sadallah, and Bouteraa's (2023) finding that the absence of systematic AI policy infrastructure in Iraqi universities constitutes a critical gap. The high endorsement of C9 (universities should establish strict AI policies; M = 4.298, 89.3% agree) simultaneously with the reported transgressive behaviours in C2–C5 further underscores students' recognition of the need for structural intervention — a recognition they appear to apply to institutions rather than to themselves.

4.1.4 Section D: Usage Patterns and Frequency (RQ1 and RQ2)

Table 6. *Section D — Item-Level Descriptive Statistics and Frequency Distributions*

Item	Statement (abbrev.)	M	SD	Mdn	Skew	Kurt	SD %	D %	N %	A %	SA %
D1	Use ChatGPT at least once weekly for academic writing	4.223	0.974	4.0	-1.677	2.861	3.2	5.3	3.2	42.6	45.7
D2	Use mainly for grammar/spelling correction	4.298	0.760	4.0	-1.314	3.073	1.1	1.1	8.5	45.7	43.6
D3	Use mainly to generate ideas and outlines	4.309	0.868	4.0	-2.060	5.581	3.2	2.1	1.1	47.9	45.7
D4	Use mainly to paraphrase/rewrite paragraphs	4.309	0.868	4.0	-2.060	5.581	3.2	2.1	1.1	47.9	45.7
D5	Use mainly to find references/sources	4.074	1.138	4.0	-1.490	1.716	7.4	2.1	9.6	37.2	43.6
D6	Use ChatGPT more than dictionaries/grammar books	4.064	1.162	4.0	-1.512	1.566	7.4	5.3	3.2	41.5	42.6
D7	Use ChatGPT more than any other AI tool	4.043	0.972	4.0	-1.164	1.165	2.1	7.4	9.6	45.7	35.1
D8	Started using because classmates use	4.064	1.014	4.0	-1.395	1.669	3.2	8.5	3.2	48.9	36.2

	it									
D9	Use ChatGPT more when teacher feedback insufficient	3.745	1.135	4.0	-0.785	-0.394	3.2	18.1	6.4	45.7	26.6
D10	Would use ChatGPT even if university prohibited it	3.883	1.135	4.0	-0.984	-0.013	3.2	16.0	3.2	44.7	33.0

Note. Items D3 and D4 share identical means (M = 4.309), indicating equal primacy of generative and paraphrasing functions. Item D10 has critical policy implications for the effectiveness of prohibition-based institutional strategies.

Section D yielded a composite mean of M = 4.101 (High/Agree), confirming that ChatGPT use among this population is not peripheral or exploratory but intensive, multi-purpose, and structurally embedded in academic writing routines. All ten items fell within the High to Very High classification range.

Frequency of use is confirmed as very high: D1 (weekly use for academic writing; M = 4.223) indicates that 88.3% of respondents use ChatGPT at least once per week. The purpose items reveal that students are not using the tool for a single narrow function but are deploying it across the full writing process. D4 (paraphrasing and rewriting paragraphs; M = 4.309, SD = 0.868) and D3 (generating ideas and outlines; M = 4.309, SD = 0.868) share the highest mean in the section, indicating that pre-writing ideation and rewriting for plagiarism evasion are the co-equal primary uses. This is followed by D2 (grammar and spelling correction; M = 4.298), D1 (weekly use; M = 4.223), and D5 (finding references; M = 4.074). The simultaneous endorsement of D2 (surface correction) and D3–D4 (compositional and paraphrasing functions) confirms that students are exploiting the full range of ChatGPT's writing capabilities rather than restricting its use to lower-stakes functions.

Item D6 (greater use than dictionaries and grammar reference books; M = 4.064, 84.1% agree) carries significant implications for vocabulary acquisition and grammatical meta-awareness. The displacement of traditional reference tools by ChatGPT suggests that students are no longer developing the habit of consulting rule-based resources — a shift likely to reduce their metalinguistic self-regulatory capacity in contexts where AI is unavailable, such as unsupervised examinations. D8 (adoption driven by peer use; M = 4.064, 85.1% agree) confirms the social diffusion dynamics of ChatGPT adoption in this context, consistent with the subjective norm construct of the Technology Acceptance Model as applied by Mohammad et al. (2026) in the Iraqi higher education context.

The two lowest-rated items in Section D — D9 (compensatory use when teacher feedback is insufficient; M = 3.745, SD = 1.135) and D10 (continued use despite university prohibition; M = 3.883, SD = 1.135) — carry the most salient policy implications despite their lower relative means. Item D9 confirms that for a substantial proportion of students (71.3% agree), ChatGPT functions as a compensatory resource when institutional support is perceived as inadequate, reinforcing the Section E finding that students actively seek pedagogical support that is currently underdelivered. The high disagreement rate for D9 (18.1%) may reflect the opposite condition: many students use ChatGPT regardless of teacher feedback quality, suggesting that for this subgroup, ChatGPT use is habitual rather than compensatory.

Item D10 is arguably the most consequential item in the entire survey from a policy perspective: 77.7% of students indicate they would continue using ChatGPT even if their university officially prohibited it, with only 19.1% expressing disagreement. This finding constitutes a direct empirical indictment of prohibition-based institutional strategies. If nearly four in five students declare intent to circumvent institutional bans, such bans are likely to drive ChatGPT use underground rather than eliminating it — creating additional risks of concealment and self-censorship in academic discourse rather than the developmental engagement that responsible integration could foster. This finding is consistent with Jasim's (2025) contention that prohibition-only approaches fail to account for the deeply embedded nature of AI use among current Iraqi EFL student populations.

4.1.5 Section E: Pedagogical Needs and Institutional Support (RQ4)

Table 7. *Section E — Item-Level Descriptive Statistics and Frequency Distributions*

Item	Statement (abbrev.)	M	SD	Mdn	Skew	Kurt	SD %	D %	N %	A %	SA %
E1	Teachers should guide responsible ChatGPT use	4.287	0.812	4.0	-1.562	4.024	2.1	0.0	9.6	43.6	44.7
E2	Benefit from prompt engineering training	4.117	1.076	4.0	-1.295	1.013	3.2	8.5	7.4	35.1	45.7
E3	ChatGPT should	4.319	0.930	5.0	-1.83	3.818	3.2	2.1	6.4	36.2	52.1

	supplement, not replace instruction				1						
E4	Tasks should require critical evaluation of AI output	4.223	0.974	4.0	-1.677	2.861	3.2	5.3	3.2	42.6	45.7
E5	Teacher feedback more valuable than AI for argumentation	4.309	0.868	4.0	-2.060	5.581	3.2	2.1	1.1	47.9	45.7
E6	Need support verifying AI-generated information	4.074	1.138	4.0	-1.490	1.716	7.4	2.1	9.6	37.2	43.6
E7	Writing assessment should be updated for ChatGPT era	4.064	1.162	4.0	-1.512	1.566	7.4	5.3	3.2	41.5	42.6
E8	AI literacy should be part of the writing course	4.223	0.974	4.0	-1.677	2.861	3.2	5.3	3.2	42.6	45.7
E9	Combined ChatGPT + teacher feedback is best approach	4.043	0.972	4.0	-1.164	1.165	2.1	7.4	9.6	45.7	35.1
E10	University	4.06	1.01	4.0	-	1.66	3.2	8.5	3.2	48.	36.2

should develop clear AI ethical policy	4	4	1.395	9		9

Note. Item E3 achieved the highest Strongly Agree rate in the section (52.1%), signalling students' strong endorsement of the supplement-not-replace paradigm. Item E9 imputed for 3 missing values (mode = 5).

Section E yielded a composite mean of M = 4.172 (High/Agree), reflecting strong and broadly distributed endorsement of the need for pedagogical and institutional intervention across all ten dimensions assessed. No item fell below the High classification, confirming that demand for institutional and pedagogical support is a consistent and non-idiosyncratic feature of the student population.

The highest-rated item in the section — E3 (ChatGPT should be integrated as a supplementary tool, not a replacement for instruction; M = 4.319, SD = 0.930) — also achieved the highest Strongly Agree rate in the section (52.1%). This finding is theoretically significant: students who are themselves engaged in the intensive multi-purpose AI use documented in Section D nonetheless endorse the complement-not-replace paradigm advocated across the literature (Jasim, 2025; Mohammad et al., 2026; Waseem and Ali, 2025). The co-existence of heavy use and explicit endorsement of complementary rather than substitutive integration suggests not hypocrisy but a structural gap — students recognise what responsible use would look like but do not receive the institutional and pedagogical support that would enable them to enact it.

E5 (teacher feedback is more valuable than ChatGPT feedback for developing argumentation and critical thinking; M = 4.309, 93.6% agree) is the second-highest rated item and reinforces this conclusion. Students' near-universal recognition that teacher feedback provides what AI cannot — contextual depth, critical engagement, and argumentation scaffolding — directly contradicts the displacement of teacher reference by ChatGPT documented in D6. This tension suggests that students are making pragmatic tool-use decisions driven by convenience and availability rather than by preference or perceived quality. If teacher feedback were more regularly available in the forms students recognise as valuable, the D6 displacement pattern might be substantially attenuated.

E1 (teacher guidance on responsible use; M = 4.287, 88.3% agree), E4 (writing tasks requiring critical evaluation of AI output; M = 4.223), and E8 (AI literacy as part of writing course curriculum; M = 4.223) cluster closely together, collectively signalling that students seek not merely passive ethical guidance but structured pedagogical scaffolding that develops their capacity to engage critically with AI-generated content. The strong endorsement of E8 is particularly salient in light of the Section C findings: students who acknowledge widespread integrity-compromising behaviour simultaneously request formal AI

literacy education — an alignment that suggests students themselves recognise the inadequacy of their current AI use practices and wish for structured institutional support to improve them. This finding strongly supports curriculum-integrated AI literacy programmes over punitive institutional responses.

The two items with the most heterogeneous response profiles in the section — E6 (support for verifying AI-generated information and references; M = 4.074, 7.4% strongly disagree) and E7 (assessment reform to account for the ChatGPT era; M = 4.064, 7.4% strongly disagree) — remain in the High classification but reflect slightly greater variation in perceived urgency. The relatively lower endorsement of E7 relative to E1, E3, and E5 may suggest that students prioritise changes in teaching approach over changes in assessment design, or that the implications of assessment reform are perceived as more uncertain or complex. The near-universal endorsement of E10 (university should develop clear AI ethical policy; M = 4.064, 85.1% agree) confirms the demand for institutional-level policy infrastructure that is currently absent in the Iraqi higher education context.

4.2 Normality Testing

Shapiro-Wilk tests were applied to each section composite score prior to inferential analysis. All five composites violated the normality assumption ($W = 0.761–0.955$, $p < .001–.003$), attributable to the strong ceiling skewness characteristic of high-endorsement Likert data in ChatGPT perception surveys (Miri, 2025; Mohammad et al., 2026). These violations have direct implications for the selection of inferential analyses: while Pearson correlations are reported as specified in the methodology — given the robustness of Pearson's r to mild non-normality at n = 94 — Spearman's rho and Kruskal-Wallis tests are noted as appropriate non-parametric sensitivity checks, and their application is recommended for the final confirmatory analysis.

4.3 Correlation Analysis (RQ4)

Pearson correlation coefficients were computed between all section composite scores using SPSS Bivariate Correlate. All correlations were statistically significant at $p < .001$ (two-tailed). Table 8 presents the full correlation matrix.

Table 8. *Pearson Correlation Matrix for Section Composite Scores*

Section	A	B	C	D	E
Section A	—	.929**	.996**	.989**	.998**
Section B	.929**	—	.916**	.962**	.939**
Section C	.996**	.916**	—	.982**	.992**
Section D	.989**	.962**	.982**	—	.993**
Section E	.998**	.939**	.992**	.993**	—

*Note. ** p < .001 (two-tailed). n = 94. All off-diagonal correlations are statistically significant. Correlations computed using SPSS Bivariate Pearson procedure.*

All inter-section correlations were very strongly positive, ranging from r = .916 (Sections B and C) to r = .998 (Sections A and E). Section E (Pedagogical Needs and Institutional Support) produced the strongest correlations with all other sections: r = .998 with Section A (p < .001), r = .993 with Section D (p < .001), r = .992 with Section C (p < .001), and r = .939 with Section B (p < .001). This pattern indicates that students who perceive greater benefits from ChatGPT, report higher usage frequency, and acknowledge greater integrity concerns are also those most strongly supportive of pedagogical and institutional intervention — a coherent and theoretically consistent profile.

The weakest correlation in the matrix — r = .916 between Sections B and C — while still very strong, suggests that risk perception is a somewhat more distinct attitudinal dimension than the tightly integrated cluster formed by Sections A, C, D, and E. This is consistent with the bifurcation structure within Section B discussed above: the low-scoring self-referential risk items (B1–B3) create a degree of internal heterogeneity that partially decouples Section B from the broader attitudinal pattern shared by the other four sections.

An important caution applies to the interpretation of the magnitude of these correlations. The very high inter-section correlations (r > .90 throughout) may partially reflect structural commonalities in the frequency distributions across sections — several items share identical frequency patterns across sections in the source data — in addition to genuine covariance in students' attitudinal profiles. These correlations should therefore be interpreted as indicative of strong attitudinal coherence while acknowledging the possibility of partial distributional artefact. The regression VIF analysis reported in Section 4.6 provides a further diagnostic for this concern.

4.4 Multiple Regression Analysis (RQ4)

A multiple regression analysis was conducted with Section E composite score as the dependent variable and Section A, B, C, and D composite scores as simultaneous predictors, using SPSS Analyze > Regression > Linear. Table 9 presents the model summary and coefficient estimates.

Table 9. *Multiple Regression: Predictors of Pedagogical Support Endorsement (Section E)*

F	R	R2	Adj. R^2	F	Df	P
Model	.9990	.9979	.9978	10538.31	4, 89	< .001

Summary

Table 10. *Regression Coefficients and Multicollinearity Diagnostics*

Predictor	B	SE B	Beta	t	p	VIF	Sig.
Constant	-0.062	0.025	—	-2.425	.017	—	*
Section A (Benefits)	0.939	0.080	0.914	11.789	< .001	254.2	**
Section B (Risks)	-0.041	0.022	-0.039	-1.852	.067	19.1	ns
Section C (Integrity)	-0.239	0.065	-0.224	-3.699	< .001	154.9	**
Section D (Usage)	0.350	0.053	0.347	6.613	< .001	116.2	**

*Note. B = unstandardised coefficient; SE B = standard error; Beta = standardised coefficient; VIF = Variance Inflation Factor. ** p < .01; * p < .05; ns = not significant at α = .05. DV = Section E composite score. CRITICAL NOTE: All VIF values substantially exceed the threshold of VIF < 5, indicating severe multicollinearity. Individual predictor coefficients should be treated as indicative rather than definitive. Ridge regression or principal components regression is recommended for the final analysis.*

The overall regression model was statistically significant, $F(4, 89) = 10538.31$, $p < .001$, accounting for 99.79% of the variance in Section E composite scores ($R^2 = .998$, Adj. $R^2 = .998$). This very high explained variance confirms the strong predictive relationship between students' perceptions of ChatGPT benefits, risks, integrity concerns, and usage patterns on the one hand, and their demand for pedagogical and institutional support on the other.

However, a critical qualification is mandatory: all VIF values substantially exceeded the threshold of VIF < 5 below which multicollinearity is considered acceptable (Field, 2018). VIF values of 254.2 (Section A), 154.9 (Section C), and 116.2 (Section D) indicate extreme multicollinearity, consistent with the very high inter-predictor correlations (r = .916–.996) documented in Table 8. Severe multicollinearity of this magnitude renders individual regression coefficients, standard errors, and statistical significance tests highly unstable and potentially misleading — the Beta coefficients and their significance values reported in Table 10 should therefore be treated as indicative rather than definitive estimates of predictor importance. The model R^2 and F-statistic

remain interpretable as indicators of overall predictive power, but individual coefficient interpretation requires extreme caution.

With this caveat firmly in mind, the pattern of coefficients is noted: Section A (perceived benefits; Beta = .914, $p < .001$) emerges as the strongest positive predictor of pedagogical support endorsement, suggesting that the strength of students' perceived benefits from ChatGPT is the dominant driver of their demand for institutional and pedagogical frameworks to regulate and support its use. Section D (usage frequency; Beta = .347, $p < .001$) is the second positive predictor. Section C (academic integrity; Beta = -.224, $p < .001$) exerts a statistically significant negative coefficient, which likely reflects distributional rather than genuine suppression given the known multicollinearity in the model. Section B (perceived risks; Beta = -.039, $p = .067$) does not reach significance after partialling for the other predictors. A principal components regression or ridge regression approach, which reduces multicollinearity by analysing orthogonal components or applying coefficient shrinkage respectively, is strongly recommended for the final thesis analysis.

4.5 Inferential Analyses for RQ3: Demographic Moderators

Research Question 3 examines whether gender, academic year, and prior AI experience moderate students' perceptions of and engagement with ChatGPT. The demographic profile of the sample (Table 3.1) records n = 54 male and n = 40 female students; second-year (n = 25), third-year (n = 35), and fourth-year (n = 34) representation; and AI experience duration categories of fewer than 6 months (n = 22), 6–12 months (n = 21), and more than 12 months (n = 51).

The planned inferential procedures — independent-samples t-tests for gender comparisons and one-way ANOVAs for academic year and AI experience duration comparisons — require item-level response data paired with respondent demographic identifiers. The present analysis was conducted from aggregated frequency data; the following procedures should be applied when the full item-level dataset is available.

For gender comparisons, an independent-samples t-test will be conducted for each section composite score, with Levene's test for equality of variances applied prior to the main test and the Welch correction applied if variances are unequal. Effect sizes will be reported as Cohen's d, with thresholds of d = 0.20 (small), 0.50 (medium), and 0.80 (large) (Cohen, 1988). Miri (2025) found that female Iraqi EFL students reported slightly higher perceptions, frequency of use, and satisfaction than male counterparts, and the present analysis will test whether a similar pattern holds in the current sample.

For academic year and AI experience duration comparisons, one-way ANOVAs will be conducted for each section composite, with Shapiro-Wilk and Levene's tests applied as distributional pre-tests. Eta-squared will be reported as the effect size measure (small = .01, medium = .06, large = .14), and Tukey's Honestly Significant Difference post-hoc procedure will be applied to identify which pairs

of groups differ significantly when the omnibus F-test is significant. If Shapiro-Wilk or Levene's tests indicate that ANOVA assumptions are violated, Kruskal-Wallis tests will be substituted. Miri (2025) identified prior AI experience duration as the strongest individual-level predictor of productive ChatGPT engagement, and a threshold versus gradual accumulation pattern analysis will be conducted to determine whether experience effects operate categorically or continuously in the current sample.

4.6 Discussion

The results of this study present a coherent and theoretically significant picture of how Iraqi EFL undergraduate students perceive, use, and situate ChatGPT within their academic writing contexts. Four principal themes emerge from the cross-sectional analysis, each of which is discussed in relation to the relevant theoretical and empirical literature.

The Very High endorsement of Section A (M = 4.209) across all ten benefit dimensions confirms that ChatGPT is perceived not as a marginal convenience but as a genuinely transformative resource by this sample. This finding is consistent with the international literature on EFL learners' ChatGPT perceptions (Miri, 2025; Aljohani, 2026; Muhammad and Hasan, 2026), but its magnitude — all ten items falling in the Very High range — is notably stronger than most reported findings and likely reflects the acute writing challenges faced by Iraqi EFL learners. As documented in Section 2, Iraqi EFL students operate in conditions of limited authentic English exposure, under-resourced institutions, and traditional pedagogical environments that offer comparatively little scaffolding for the simultaneous demands of academic writing (Miri, 2025; Muhammad and Hasan, 2026). In this context, ChatGPT's capacity to provide immediate, non-judgmental, multi-function writing support addresses a genuine and acute need that existing institutional provision does not meet.

The highest perceived benefits — idea generation (A5), academic style improvement (A8), and confidence enhancement (A7) — are precisely those that address the aspects of academic writing most structurally demanding for L2 learners whose first language imposes different rhetorical and genre conventions. Vygotsky's concept of the Zone of Proximal Development (ZPD) offers an apt theoretical frame for this pattern: ChatGPT functions as a mediating tool that enables students to perform tasks — generating academically styled text, structuring arguments, producing cohesive discourse — that exceed their current independent competence (Altamimi, 2025; Gburi and Dowlatabady, 2025). The critical pedagogical question, however, is whether this scaffolding progressively develops independent capacity (genuine ZPD-based mediation) or substitutes for the cognitive effort that would build independent capacity (scaffolding without withdrawal). The Section D findings, discussed below, suggest the latter pattern predominates in the current unguided use context.

The bifurcation within Section B — systematic rejection of personal behavioural risks (B1–B3) alongside strong endorsement of systemic epistemic risks (B4–B10) — represents one of the most analytically significant findings of the study. Students simultaneously deny that they personally over-rely on or uncritically accept AI suggestions while strongly acknowledging that ChatGPT produces inaccurate information, fabricates references, attenuates critical thinking, and weakens independent writing capacity. This pattern is not logically incoherent — a student could theoretically acknowledge these system-level risks while genuinely managing them through critical engagement — but its conjunction with the widespread integrity-compromising behaviours documented in Section C makes a self-serving attribution interpretation considerably more plausible.

This self-other asymmetry in risk perception is a well-documented psychological pattern in educational technology research, related to the optimistic bias documented by Weinstein (1980) and the third-person effect in media exposure research. Karimi and Qadir's (2025) nuanced finding — that students deny dependency while 93% acknowledge AI reliance for feedback — parallels the present data almost precisely. The theoretical implication is that risk education strategies targeting self-awareness are likely to be more effective than those targeting knowledge of system risks: students already know what ChatGPT does wrong at the system level; what they resist acknowledging is that they are personally affected by these limitations.

The most consequential finding of the study is the severe and pervasive integrity awareness-behaviour gap documented in Section C. That 93.6% of students report having used ChatGPT to paraphrase text to evade plagiarism detection (C4) and to include unverified AI-generated references (C5), while 88.3% simultaneously acknowledge that submitting AI-generated text as one's own constitutes academic dishonesty (C1), is a finding with immediate and urgent implications for institutional policy.

This paradox is not adequately explained by ignorance or lack of concern. It is more parsimoniously explained by a confluence of factors identified across the relevant literature: the perceived low probability of detection (C10, 93.6% agree that plagiarism tools cannot reliably detect AI content); the normalisation of AI-assisted writing through peer contagion (D8, 85.1% agree); the absence of clear and consistently communicated institutional guidelines (C6-C7); and the structural writing challenges that make complete independent academic writing exceptionally demanding for this population (Miri, 2025; Bin-Nashwan et al., 2023). Cotton, Cotton, and Shipway (2024) argue that academic integrity interventions that focus exclusively on awareness without addressing these structural enabling conditions are unlikely to change behaviour — a conclusion powerfully supported by the present data, where awareness is near-universal but transgression is near-universal simultaneously.

Miri's (2025) document analysis finding that approximately 80% of AI-assisted student papers contained fabricated or incorrectly formatted references acquires additional significance in light of C5's result: if 93.6% of this sample self-report including unverified AI-generated references, the prevalence rate documented at the textual level is, if anything, likely to be an underestimate. The risk of academic and scientific record contamination by AI-fabricated citations — already documented as a global concern (Cheng, et al., 2024) — is evidently acute in this population.

The conjunction of findings across Sections D, E, and C leads to an unambiguous and policy-relevant conclusion: prohibition-based institutional strategies are empirically unlikely to succeed, and students themselves recognise and endorse the alternative. Item D10's finding that 77.7% of students would continue using ChatGPT despite university prohibition directly empirically falsifies the premise of prohibition-based approaches. The near-universal endorsement across Section E of structured teacher guidance (E1), supplementary rather than substitutive integration (E3), critical AI evaluation tasks (E4), and AI literacy education (E8) by the same students who report the most extensive and integrity-compromising use provides the clearest possible student-centred mandate for the alternative approach.

This finding is directly consistent with the pedagogical consensus emerging from the international literature. Altamimi (2025) argues that ChatGPT's ZPD-based scaffolding value is realised only when the scaffolding is deliberately structured to be withdrawn as competence develops — a condition that requires intentional pedagogical design rather than passive exposure. Al-Obaydi et al. (2025) empirical finding that students using ChatGPT as an interactive guide achieved substantially better outcomes than those using it as a content source aligns with the Section E finding that students themselves endorse interactive, critically-oriented modes of engagement. Tatar and Brime's (2025) finding that critical thinking benefits from AI integration only emerged in deliberately designed inquiry tasks provides further support for the E4 result: task redesign requiring critical evaluation of AI output is not merely pedagogically desirable but appears necessary for any cognitive benefit to be realised.

For Iraqi higher education institutions specifically, the implications are both urgent and tractable. The Section C finding that students perceive a deficit in institutional guidance (C6-C7) and simultaneously demand policy development (C9, E10) — while Bin-Nashwan et al. (2023) confirm that no systematic AI policy infrastructure currently exists in Iraqi universities — defines the institutional gap that must be prioritised. The student-centred evidence of this study suggests that the most effective institutional response will combine: (1) clear, consistently communicated AI use policies that specify what constitutes acceptable and unacceptable AI assistance across different assessment types; (2) curriculum-integrated AI literacy modules that develop students' capacity to

critically evaluate AI output, verify AI-generated information, and manage citation integrity; (3) assessment redesign that incorporates AI-critical tasks, process-oriented documentation, and in-class writing components that cannot be AI-substituted; and (4) teacher professional development to equip instructors to provide the feedback on argumentation and critical thinking that students (E5) recognise as irreplaceable and that constitutes the primary comparative advantage of human over AI feedback.

The reliability analysis confirmed excellent internal consistency across all five sections ($\alpha = 0.972–0.997$). The descriptive analysis revealed: (RQ1) very high perceived benefits across all ten benefit dimensions ($M = 4.209$), with idea generation, academic style improvement, and confidence enhancement as the highest-rated; (RQ2) a critical bifurcation in risk perception between rejected personal behavioural risks (B1–B3, $M = 1.63–2.32$) and strongly acknowledged systemic risks (B4–B10, $M = 3.75–4.21$), alongside a severe integrity awareness-behaviour paradox in Section C; (RQ1/RQ2) intensive, multi-purpose, and structurally embedded ChatGPT use across the full writing process (Section D, $M = 4.101$), with 77.7% declaring intent to continue use despite prohibition; and (RQ4) strong and consistent demand for pedagogical scaffolding and institutional policy development across all ten Section E dimensions ($M = 4.172$). Pearson correlations confirmed very strong positive associations between all sections ($r = .916–.998$, $p < .001$), with Section E most strongly predicted by Section A perceived benefits ($r = .998$). The multiple regression model accounted for 99.79% of Section E variance ($R^2 = .998$), though severe multicollinearity (VIF = 19.1–254.2) necessitates cautious interpretation of individual predictor coefficients.

5. Conclusion and Implications of the study

5.1 Conclusion

This study attempted to examine the perceptions of Iraqi EFL students toward the use of ChatGPT as a writing tool. It explored their perceptions about the benefits, risks and challenges, patterns of use, ethical dimensions, and perceived need for pedagogical and institutional support. Based on this goal, a cross-sectional survey was carried out on 94 under graduate students from different Iraqi universities. The analysed data show that the findings provide a coherent, multi-dimensional, and educationally significant description of how ChatGPT is being integrated, though misused, within a context characterised by acute writing challenges, limited institutional infrastructure, and an almost total absence of structured AI literacy guidance.

The study finds that Iraqi EFL university students significantly look at ChatGPT as a transformative, multi-purpose writing tool, with near-universal endorsement of its benefits across grammar, vocabulary, organisation, idea generation, and confidence building. While students demonstrate clear awareness of the tool's systemic risks — including misinformation, fabricated references, and the

attenuation of critical thinking — they simultaneously deny that these risks apply to their own behaviour, a well-documented optimistic bias. More strikingly, despite almost universal acknowledgement that submitting AI-generated work constitutes academic dishonesty, the vast majority self-report doing exactly that, revealing an integrity crisis driven not by ignorance but by perceived impunity, peer normalisation, and the near-total absence of institutional policy in Iraqi universities.

Despite this deeply embedded and largely ungoverned pattern of use, the same students strongly endorse structured pedagogical and institutional intervention. They call for teacher guidance, curriculum-integrated AI literacy, assessment reform, and clear ethical policies, signalling not hypocrisy but a structural gap between what they recognise as responsible use and what their institutions currently provide. The study's overarching conclusion is therefore that prohibition would fail, and that what is urgently needed instead is a contextually grounded, integration-oriented framework that converts ChatGPT's genuine educational potential into measurable learning gains while meaningfully containing its risks.

5.2 Implications of the Study

The findings of this study demonstrate some practical implications across some domains, including pedagogical practice, curriculum design, and institutional policy.

At the level of pedagogical practice, it is implicated that teachers should move away from passive, undirected AI use toward deliberately designed AI-integrated writing instruction. This means providing explicit guidance on responsible ChatGPT use, redesigning tasks to require critical evaluation of AI output rather than passive acceptance, and applying a scaffolding-withdrawal approach — gradually reducing AI assistance as student competence grows. Risk education should target students' self-awareness of their own dependency, not just generic knowledge of ChatGPT's limitations, since students already demonstrate awareness of the latter.

For curriculum design, the findings of the study insist that AI literacy should be embedded formally within writing courses, covering prompt engineering, information verification, critical evaluation of AI output, and co-constructed ethical use frameworks. Assessment should be progressively redesigned to include components AI cannot substitute — such as portfolios, process journals, timed in-class writing, and oral defence — reducing the structural incentive for wholesale AI substitution.

At the level of institutional policy, the findings show that most of students would continue using ChatGPT even under a ban. Hence, effective policy must be integrated rather than prohibited, clearly distinguishing acceptable from unacceptable AI assistance and communicating this consistently across

departments. Given that 93.6% believe detection tools are unreliable, institutions should shift from detection-and-punishment models to prevention-and-education frameworks, while also investing in teacher professional development to close the delivery gap between the teacher feedback students value and what they currently receive.

References

- Aljohani, N. J. (2026). ChatGPT in language learning: A systematic review of applications and challenges. *Social Sciences & Humanities Open*, *13*, 102357.
- Al-Obaydi, L. H., Pikhart, M., & Hossain, M. K. (2025). ChatGPT and the development of core language skills: An exploratory study of EFL college students. Contemporary Educational Technology, 17(4), ep591.
- Alruwaili, A. R., & Kianfar, Z. (2025). Investigating EFL female Saudi teachers' attitudes toward the use of ChatGPT in English language teaching. Forum for Linguistic Studies, 7(2), 203–216.
- Altamimi, D. H. F. (2025). Unlocking potential: Saudi EFL male students' perspectives on AI tools for enhancing English writing proficiency. Arab World English Journal (AWEJ) Special Issue on Artificial Intelligence, 40–58.
- Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. Technology in Society, 102370.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.
- Cohen, L., Manion, L., & Morrison, K. (2018). Research methods in education (8th ed.). Routledge.
- Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in education and teaching international*, *61*(2), 228-239.
- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE.
- Farrah, Z., Kawasmeh, N., & Farrah, M. (2025). English major students' attitudes towards using ChatGPT in academic writing courses at Hebron University. Hebron University Research Journal (B), 20(4), 191–218.
- Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed.). SAGE.
- Frontiers in Education. (2025). AI vs. teacher feedback on EFL argumentative writing: A quantitative study. Frontiers in Education, 10. https://doi.org/10.3389/feduc.2025.1614673
- Gburi, H. R., & Dowlatabady, H. R. (2025). Investigating the role of artificial intelligence in English writing fluency among Iraqi university students. International Journal of Environmental Sciences, 11(17s), 103–111.
- George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference (4th ed.). Allyn & Bacon.

➢ Jasim, A. (2025). The Impact of Using ChatGPT and Grammarly on developing Iraqi EFL University Students' Academic Essay Writing. *Journal of Misan Researches*, *21*(41), 365-390.
➢ Karimi, E. M., & Qadir, B. M. (2025). The impact of artificial intelligence use on students' autonomous writing. *Journal of Applied Learning & Teaching*, *8*(1), 143-153.
➢ Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.
➢ Khampusaen, D. (2025). The Impact of ChatGPT on Academic Writing Skills and Knowledge: An Investigation of Its Use in Argumentative Essays. *LEARN Journal: Language Education and Acquisition Research Network*, *18*(1), 963-988.
➢ Kofinas, A. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. British Journal of Educational Technology. https://doi.org/10.1111/bjet.13585
➢ Kovari, A. (2025). Ethical use of ChatGPT in education — Best practices to combat AI-induced plagiarism. Frontiers in Education, 9, 1465703.
➢ Kucuk, T. (2025). Perceptions of students and teachers about hindrances of ChatGPT in foreign language education. Arab World English Journal (AWEJ) Special Issue on Artificial Intelligence, 220–233.
➢ Liu, C., Hou, J., Tu, Y. F., Wang, Y., & Hwang, G. J. (2023). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. Interactive Learning Environments, 31(9), 5614–5632.
➢ Lund, B. D., Lee, T. H., Mannuru, N. R., & Arutla, N. (2025). AI and academic integrity: Exploring student perceptions and implications for higher education. Journal of Academic Ethics.
➢ Mali, Y. C. G. (2025). Exploring the use of ChatGPT in EFL/ESL writing classrooms: A systematic literature review. Journal of Language and Education, 11(2), 137–156.
➢ Miri, B. A. M. (2025). Iraqi EFL Learners' Use of ChatGPT in Academic Research: A Convergent Parallel Mixed Methods Design. *Adab Al-Basrah*, *1*(114).
➢ Mirzayev, A. B. U., Kamal, S. S. L. A., & Azamovna, A. M. (2025). Understanding Uzbekistan university EFL teachers' perceptions of ChatGPT: From benefits to ethical challenges. International Journal of Information and Education Technology, 15(2), 246–254.
➢ Mohammad, T., Nazim, M., Alzubi, A. A. F., & Khan, S. I. (2026). EFL teachers' perceptions of using ChatGPT in writing classroom: Implications for teaching and learning. Multidisciplinary Reviews, 9, e2026384.

➢ Mohammed, Y. J., Musa, Z. H., Asim, A., & Ahmed, R. S. (2024). Developing EFL writing with AI: Balancing benefits and challenges. Technology Assisted Language Education, 2(2), 80–93.
➢ Muhammad, A., & Hasan, I. (2026). ETHICAL INTEGRATIONS OF ARTIFICIAL INTELLIGENCE IN EFL LEARNING: OPPORTUNITIES, CHALLENGES, AND FAIR ASSESSMENT (STUDENTS'PERSPECTIVE). *Humanities Journal of University of Zakho*, *14*(1), 8-18.
➢ Nazim, M., Mohammad, T., Alzubi, A. A. F., & Khan, S. I. (2025). Exploring ChatGPT as a virtual assistant for personalized writing experience: EFL students' insights. Architecture Image Studies, 6(4), 1337–1355.
➢ Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. Journal of Experimental Social Psychology, 45(4), 867–872.
➢ Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education. *Journal of applied learning & teaching*, *6*(1), 342-363.
➢ 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*, 1260843.
➢ Cheng, A., Nagesh, V., Eller, S., Grant, V., & Lin, Y. (2024). *Exploring AI hallucinations of ChatGPT: Reference accuracy and citation relevance of ChatGPT models and training conditions* (pp. 10-1097). LWW.
➢ Tabachnick, B. G., & Fidell, L. S. (2019). Using multivariate statistics (7th ed.). Pearson.
➢ Tandfonline. (2025). Comparing the effects of ChatGPT and automated writing evaluation on students' writing and ideal L2 writing self. Computer Assisted Language Learning.
➢ Tatar, N. H., & Brime, A. A. (2025). Enhancing Critical Thinking Skills among EFL Students at Salahaddin University-Erbil: The Role of AI. *Zanco Journal of Human Sciences*, *29*(SpB), 891-905.
➢ Waseem, L., & Ali, A. (2025). The impact of AI powered writing tools on students' academic writing performance. Sociology & Cultural Research Review, 4(2), 66–90.
➢ Xu, L. (2020). The dilemma and countermeasures of AI in educational application. In *Proceedings of the 2020 4th international conference on computer science and artificial intelligence* (pp. 289-294).
➢ Yu, H. (2024). The application and challenges of ChatGPT in educational transformation: New demands for teachers' roles. *Heliyon*, *10*(2).
➢ Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and education: Artificial intelligence*, *2*, 100025.