ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

LEARNING GAINS AND PEDAGOGICAL IMPLICATIONS OF GENERATIVE AI IN ACADEMIC WRITING: A SYSTEMATIC REVIEW OF BS ENGLISH STUDENTS IN PAKISTAN

Qaisra Honey
qaisrahoney@gmail.com

Abstract

The paper is a systematic literature review of the utilization of generative artificial intelligence (GenAI) and specifically ChatGPT in academic writing in higher education, focusing on BS English learners in Pakistan and other ESL/EFL settings. The analysis identified three important themes, including (1) academic writing learning gains, (2) ethical and academic integrity issues, and (3) implications in pedagogy and assessment. The results show that GenAI may positively impact grammar, coherence, sentence structure, and paragraph organization, especially when students actively use AI in brainstorming, revising, and reflecting. Nevertheless, there are still some issues of authorship, plagiarism, over-reliance, and the quality of outputs produced by AI. To adequately integrate GenAI, there must be a clear instruction of AI literacy, ethical guidance, and adjustment of assessment practices so that writing competence can be truly demonstrated. The paper identifies the conditional possibility of GenAI to aid academic writing, underlining that pedagogical intentionality, institutional frameworks, and responsible use are essential to the meaningful educational outcomes. Teaching and assessment implications and future research implications on the development of long-term writing in ESL/EFL settings are discussed.

Keywords: Generative artificial intelligence, ChatGPT, academic writing, ESL/EFL, higher education, AI literacy, pedagogical implications, academic integrity

Introduction

Generative artificial intelligence (GenAI) is transforming teaching, learning, and assessment in higher education, especially in writing-intensive courses. ChatGPT is one of the modern GenAI applications that have garnered significant academic interest due to its capacity to create, edit, summarize, and scaffold written text in real-time. When used in English as a Second Language (ESL) and English as a Foreign Language (EFL) learning, these tools are also beginning to be examined as means of helping learners in the idea generation, drafting, revision and linguistic refinement processes. Current studies indicate that GenAI could positively affect grammar, coherence, cohesion, lexical choice, and structural composition in student writing and provide prompt and personal feedback, which can be quite challenging to maintain in a traditional classroom environment (Chanpradit, 2025; Mahapatra, 2024; Ogunleye et al., 2024). Although these tools help improve the accuracy and the organization of language, the issues of authorship, over-dependence, and shallow learning still persist.

GenAI has a pedagogical promise especially in situations where students have a consistent challenge in academic writing in English and where teacher feedback is limited due to the size of the class, institutional resources, and time constraints. English as an academic tool in Pakistani higher education is a vehicle of academic engagement and success, but still, in the undergraduate level, numerous students are still grappling with argumentation, grammar, textual coherence, vocabulary, and disciplinary writing norms. The consequences of these kinds of challenges are particularly more significant to the students who pursue BS English programs since academic writing is not only a communication skill, but a primary intellectual and disciplinary means of interaction. In this regard, writing assistance is potentially a valuable tool, like ChatGPT, which expands learning opportunities, offers feedback, self-correction, and independent learning opportunities in the classroom (Ajmal et al., 2025; Yasmin et al., 2025).

Meanwhile, the increased use of GenAI in academic writing has spawned a lot of pedagogical, ethical, and epistemological controversy. Even though some studies have documented an improvement in the quality of student writing, researchers have also been

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

concerned about plagiarism, overdependency, cognitive offloading, hallucinated content, decreased authorship awareness, and blindly accepting language generated by machines. These issues are especially important in writing pedagogy, where the purpose is not just to produce refined texts but to build a capacity to think and write independently, to make rhetorical decisions, to become critically literate and to establish a voice of one's own. This means that a significant difference should be drawn between the enhanced writing performance and real learning achievement, the enhanced final performance does not necessarily imply the enhanced learning and the enhancement of writing skills (Chanpradit, 2025; Radtke & Rummel, 2025; Shi et al., 2025).

This difference is particularly critical when GenAI is framed as a teaching tool, as opposed to a substitute of the human intellect. Recent research shows that the educational benefits of AI-assisted writing can be significantly influenced by the levels of AI literacy, self-regulated learning, evaluative judgment, and strategic engagement among learners. The application of AI tools as a scaffold to aid the process of reflection, revision, and problem-solving is most beneficial to students, and not as a content generating tool. In line with this, to be responsible and pedagogically sound, the integration of GenAI should be accompanied by both access to the technological tools, as well as by an organized guidance, institutional policy, and direct instruction on ethical and critical use of AI (Mahapatra, 2024; Ogunleye et al., 2024; Shi et al., 2025).

Although the study of GenAI in the education field has been growing very fast, the current evidence-base is uneven, geographically focused, and contextually dispersed. A lot of the literature has been produced in an international or a technologically advanced environment, and relatively little systematic focus has been put on Pakistani ESL/EFL higher education and especially on BS English students, whose academic progress is directly linked to high levels of writing skills. Even though there is limited synthesis and critical analysis of the broader implications of the application of ChatGPT on learning gains, academic integrity, assessment practices, and sustainable development of writing skills, emerging local studies indicate that Pakistani learners tend to regard ChatGPT as a helpful tool in writing support and language improvement (Ajmal et al., 2025; Ashraf et al., 2025; Yasmin et al., 2025).

This gap is significant because of a number of reasons. To begin with, Pakistan is a unique educational setting whose characteristics are influenced by linguistic diversity, uneven access to technology, developing standards of academic integrity, and institutional dissimilarity. Second, as universities start to face issues related to the use of AI in assignments, originality, authorship, plagiarism, and redesigning of assessment, context-specific evidence is urgently required on whether GenAI can promote meaningful learning or providing a means of performing procedural tasks. Third, as writing is a fundamental academic in BS English courses, the role of GenAI in developing writing requires comprehension not only in pedagogy but also in curriculum design and institutional policy. It is thus necessary to conduct a systematic review of existing evidence to establish the type of learning gains that are being reported, the conditions that they are reported to occur in, and the challenges that are associated with them in the Pakistani context (Ashraf et al., 2025; Ogunleye et al., 2024; Wang and Fan, 2025). This is an essential gap since academic writing is the main focus of disciplinary development in BS English programs, but there is still a lack of context-specific evidence.

It is against this backdrop that the current study is a systematic review of the recent literature on the application of generative AI in academic writing, especially, in the context of its application to BS English students in Pakistan. The review does not address ChatGPT as a technological innovation but, instead, views it as a pedagogical intervention, the usefulness of which should be studied in the context of the quality of writing, the development of learners, academic honesty, and assessment practice. In particular, the review aims to examine the

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

associations between the use of GenAI and the improvement of cohesion, coherence, grammar, argumentation, ethical writing behavior, and formative feedback processes. The synthesis of available evidence with the help of a Pakistan-focused lens may allow the study to be relevant to the current debates in the field of AI in education and writing pedagogy, as well as provide practical suggestions to educators, curriculum developers, and policymakers regarding the use of AI in higher education in a responsible and context-sensitive way (Khampusaen, 2025; Mahapatra, 2024; Wang & Fan, 2025). The paper may be valuable as it offers a contextualized systematic synthesis of GenAI in academic writing in Pakistani higher education. It provides information about learning benefits, ethical issues, and pedagogical adjustments required to be made to implement AI responsibly.

Problem Statement

The growing application of generative artificial intelligence (GenAI) to the educational field has established new opportunities to facilitate academic writing, especially in ESL and EFL settings. ChatGPT and similar tools are becoming a popular topic in discussions related to their potential usage in enhancing grammar, cohesion, coherence, argumentation, and revision processes in student writing. Nevertheless, regardless of the increasing interest in these technologies, the pedagogical worth of GenAI has not been carefully comprehended in relation to real learning outcomes, particularly in academic settings that rely on writing. The emerging scholarship has centered much on short-term gains in written products but significantly less has been dedicated to whether gains in written products are indicative of a deeper learning process, longer-term growth in writing, and transferable academic competence (Chanpradit, 2025; Mahapatra, 2024; Wang & Fan, 2025).

This disjuncture is most clearly seen through the prism of Pakistani undergraduate education, where English is a dominant language of academic interaction, but is frequently acquired and utilized in environments of linguistic, pedagogical, and technological limitation. Academic writing proficiency is anticipated of high standards in students enrolled in BS English programs, but many still experience difficulties in organization, clarity, critical argumentation, grammatical accuracy and disciplinary expression. GenAI can seem a viable option in such a scenario, complementing teacher-provided feedback and providing self-directed writing assistance. However, the context-specific evidence regarding the ability of these tools to promote writing development in Pakistani BS English students or simply serve to perform short-term tasks and superficially fix them (Ajmal et al., 2025; Yasmin et al., 2025) remains insufficient.

The next significant issue is that the application of GenAI to academic writing is surrounded by various ethical, pedagogical and institutional risks. Problems that have been identified through the years by scholars include plagiarism, excessive reliance on AI-generated material, hallucinatory information, less critical interaction, and a low sense of authorship. Such dangers are particularly acute in writing pedagogy since academic writing cannot be simply seen as a matter of correct sentence production; it is essentially connected to the critical thinking, originality, disciplinary voice, and intellectual responsibility. Absent of explicit pedagogical instructions, AI literacy, and institutional controls, GenAI can inadvertently undermine, instead of enhance, the independent writing skills of students (Chanpradit, 2025; Ogunleye et al., 2024; Radtke & Rummel, 2025).

Another issue is that the current body of work on GenAI and academic writing is disorganized and uneven in distribution within the contexts. Despite the fact that the studies published after 2023 and 2025 have reported some promising results regarding the quality of writing, motivation of learners, and formative feedback, the majority of research has been done in international, generalized, or non-Pakistani contexts. Consequently, existing literature fails to reflect the particular realities of Pakistani BS English students, whose lives are influenced

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

by the variables of digital inequality, unequal access to AI, and institutional uncertainty and change in norms of academic integrity. This poses a severe gap in knowledge among teachers and policymakers in need of context-sensitive data to make a well-informed decision regarding the application of AI in the writing process at higher learning institutions (Ashraf et al., 2025; Shi et al., 2025; Wang & Fan, 2025).

It is not a question of whether or not GenAI can be used to enhance writing output, but rather whether it can do so in a manner that encourages meaningful learning, ethical academic practice and sustained pedagogical growth in Pakistani higher education. Without a critical and focused synthesis of evidence at hand, any use of AI tools by institutions can be pedagogically incongruent, ethically unregulated, and superficial in its approach to education. The need to conduct a systematic review, which explores the association of GenAI (and the ChatGPT, in particular) with the learning gains in academic writing among BS English students in Pakistan, and discusses the implications of this on academic integrity, AI literacy, and assessment practices, is therefore evident (Khampusaen, 2025; Mahapatra, 2024; Ogunleye et al., 202

The recent surge in generative artificial intelligence (GenAI) in higher education has led to discussions about the need to consider both whether these tools augment the quality of written work, but also whether they yield valuable learning outcome, ethical writing practice, and sustainable academic growth. Applying to the example of Pakistani BS English courses, where academic writing is a core aspect of disciplinary formation, the importance of comprehending how tools like ChatGPT affect the writing process, performance and academic practices of learners is increasing. Current studies reveal both the possible advantages and serious issues, such as the increase in the linguistic accuracy and organization on the one hand and the danger of over-reliance, authorship, and scholarly dishonesty, on the other hand (Mahapatra, 2024; Radtke & Rummel, 2025; Wang & Fan, 2025).

Considering these issues, the current research aims to give a syntactic overview of all accessible literature on the topic of GenAI in academic writing, including the specifics of its implications to BS English learners in Pakistan. To investigate the potential of GenAI as an instructional tool in higher education writing scenarios, the objectives and research questions are hence developed not only to explore the instructional potential of GenAI, but also to explore the ethical, pedagogical, and assessment-related implications of the application of this technology (Chanpradit, 2025; Ogunleye et al., 2024; Shi et al., 2025

Research Objectives

The present study aims to achieve the following objectives:

1. To systematically review the existing literature on the use of generative artificial intelligence in academic writing within higher education.
2. To examine the reported learning gains associated with the use of ChatGPT and other generative AI tools in the academic writing development of BS English students in Pakistan.
3. To explore the pedagogical, ethical, and assessment-related implications of generative AI in academic writing instruction.

Research Questions

The study is guided by the following research questions:

1. What does the existing literature reveal about the use of generative artificial intelligence in academic writing within higher education?
2. What learning gains have been reported regarding the use of ChatGPT and other generative AI tools in the academic writing development of BS English students in Pakistan?
3. What pedagogical, ethical, and assessment-related implications emerge from the use of generative AI in academic writing?

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Literature Review

The introduction of generative artificial intelligence (GenAI) has had a profound impact on the field of higher education, especially in areas where reading, writing, and knowledge production are crucial. Artificial intelligence tools like ChatGPT are becoming more common in schools to carry out activities like idea generation, summarization, drafting, revision, and feedback. In the context of writing pedagogy, GenAI has been established as not only a technological innovation but as a possible tool in instruction that can help students engage in the writing process in more personalized and immediate forms. Recent systematic reviews suggest that GenAI is being investigated in an exceptionally broad category of education thanks to its potential to offer quick language aids, decrease revision constraints, and scaffold textual creation in real time (Chanpradit, 2025; Ogunleye et al., 2024).

Nevertheless, GenAI cannot be interpreted strictly in terms of its technological affordances, which are not its sole educational value. Academic writing in a higher education context is not confined to the linguistic correctness, but instead, it entails the process of building the argumentative reasoning, disciplinary consciousness, critical voice and the epistemic participation. Consequently, the applicability of GenAI in writing education does not necessarily lie in whether it facilitates the creation of improved text by the students, but in whether it promotes writing development based on learning. The difference has been a focus in the modern studies as the researchers become more concerned with whether AI-aided textual enhancement is indicative of more profound learning or merely a temporary boost in performance (Wang & Fan, 2025; Shi et al., 2025).

In this connection, the existing literature implies that GenAI should be considered as a pedagogical scaffold, but not an independent educational solution. It seems to be valued depending on how critically, strategically and reflectively the learners apply it, and whether the teachers and the institutions establish structures of responsible integrations. This particularly applies to academic writing teaching, where process, reflection, and revision are valued as much as the end product of writing (Mahapatra, 2024; Radtke and Rummel, 2025).

Among the most stable results of the recent literature is that GenAI can assist in various aspects of the development of academic writing, particularly ESL/EFL. In both intervention studies and review studies, ChatGPT has been linked to upgrades in grammar, cohesion, coherence, clarity, lexical appropriateness, and organization that are key to the quality of academic writing. These advancements can be especially significant in case of second-language learners, who may experience difficulties with sentence-level accuracy and text-level structure and at the same time, have to meet academic requirements (Chanpradit, 2025; Mahapatra, 2024).

However, a closer examination of the literature reveals that such gains do not cut across the board in all areas of writing. The best evidence is likely to have to do with surface and structural aspects of writing like grammar correction, sentence fluency, paragraph coherence, and revision support but the evidence about more complex aspects, like originality, disciplinary voice, rhetorical sophistication, and critical argumentation, is relatively shallow. It implies that GenAI can be a better tool to be used when scaffolding revision and refining linguistically than as an alternative to the cognitive and rhetorical moves needed to compose a higher-level piece (Khampusaen, 2025; Radtke & Rummel, 2025).

Moreover, it is shown in the literature that pedagogical effect of GenAI is significantly determined by the way it is applied. It seems to be a more productive way of developing writing when applied as a formative support tool like to brainstorm or interpret feedback or plan revision than when applied to produce finished writing with minimal learner input. This difference is essential as it implies that the AI does not enhance writing skill, but the quality of

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

the interaction with the learner writing activities that are supported by AI (Mahapatra, 2024; Shi et al., 2025).

One of the primary issues in the literature is the question of whether advances in AI-aided writing constitute real learning benefits or just better performance in the task. This difference is particularly relevant in the educational setting since a well-written product does not always imply that a learner has internalized the skills, strategies or conceptual knowledge required to repeat the same type of performance on his/her own. Recent meta-analytic and intervention-based studies indicate that GenAI has the potential of positively impacting more general education outcomes, such as learning performance, engagement, and higher-order thinking, but these are typically mediated by contextual and learner-related factors instead of being the direct result of using the tool (Wang & Fan, 2025; Shi et al., 2025).

Learning gains in writing-specific research have been reported in the domains of error reduction, argument development, better text organization and better revision awareness. But, such advantages are best persuasive in the research where learners interact with AI in the course of time and in which writing is studied as a recursive process comprising of drafting, reflection and revision. In comparison, the research that emphasizes product quality as the main factor without tracking the process of decision-making, or skills transfer among learners leaves the question as to whether GenAI is promoting learning or merely supporting performance unanswered (Khampusaen, 2025; Mahapatra, 2024).

This is an especially important issue when it comes to BS English students whose academic writing is not only a language skill, but also a practice in discipline. In such learners, improvement of meaningful learning should preferably encompass improved grammar and coherence, as well as, academic voice, argumentation, textual independence, metacognitive control and metacognitive independence. These higher-order benefits do not just happen, according to the growing body of literature; they can only be achieved when the application of AI is incorporated into guided pedagogical settings and not uncontrolled self-use (Shi et al., 2025; Wang & Fan, 2025).

Although GenAI has got pedagogical potential, it has also raised some serious issues concerning ethics and academic integrity. The most commonly mentioned risks can be found throughout the literature, with plagiarism, excessive reliance on AI-generated work, hallucinated information, the lack of originality in the author, and the uncritical acceptance of the language generated by a machine being the most prevalent. These issues are particularly acute in academic writing, where the originality, attribution, and intellectual responsibility are the key issues in both evaluation and academic practice (Chanpradit, 2025; Ogunleye et al., 2024).

One of the most crucial problems arising due to the recent study is that students might not necessarily have the evaluative skills to assess the quality, accuracy, or the appropriateness of AI-generated text. In other instances, students seem to place too much trust in AI-generated work, and spend less time and critical thinking on editing machine-generated drafts, compared to peer- or self-written work. These results are alarming, as GenAI can potentially decrease, instead of increase, critical interaction in case students do not train to question, verify and adapt AI recommendations in a critical way (Radtke & Rummel, 2025; Shi et al., 2025).

Notably, the literature alone does not help to justify simple rejection of AI based on ethical considerations. Instead, numerous researchers claim that the point of debate is not the use of GenAI but its how, why, and under what institutional circumstances students use it. Ethical integration thus involves clear policy and expectation, and pedagogical frameworks, which can make AI as an aid to learning, not an end to learning (Chanpradit, 2025; Ogunleye et al., 2024).

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

The other key theme in the literature relates to the nature of AI literacy and how it is related to effective academic writing support. There is a growing understanding of AI literacy beyond technical familiarity as the capacity to judge AIs and identify constraints, make strategic use of prompts, and make informed decisions regarding the ethical and pedagogical appropriateness. Research indicates that more AI literate students tend to apply GenAI in a manner that facilitates their learning whereas less literate students are likely to overrely on the tool or misapply it in ways that can become problematic in academics (Shi et al., 2025; Ogunleye et al., 2024).

The potential uses of GenAI in support of formative feedback and assessment are also noted in the literature. ChatGPT could also be used in writing courses with many students or other resources to help balance out the lack of teacher feedback by offering students real-time feedback on grammar, structure, clarity, and revision options. This is especially applicable in situations where one on one writing help is hard to maintain. Nevertheless, researchers also warn that the AI-generated feedback cannot be considered as pedagogically neutral and unconditionally reliable. Learners without teacher mediation can accept bad or unsuitable suggestions, thus weakening instead of enhancing the process of developing writing (Mahapatra, 2024; Ajmal et al., 2025).

GenAI is also challenging institutions in terms of assessment to redefine what evidence of learning in students means. When students are able to generate grammatically refined work using AI, then the methods of assessment should be progressively more focused on process, reflection, transparency of authorship and the thinking quality, instead of only a focus on the surface-level textual accuracy. Therefore, the literature is leaning more towards the necessity of redesigning assessment, rather than adoption of tools (Ogunleye et al., 2024; Radtke and Rummel, 2025).

Despite a rapid growth in the international literature on GenAI and academic writing, the evidence that is available is not contextualized with reference to the country of Pakistan, especially with the BS English students. The bulk of published literature has either discussed GenAI in general international contexts or addressed general groups of students as opposed to those studying a particular discipline as an undergraduate English learner. Consequently, the mechanisms of GenAI that are at work in the applied context of Pakistani higher education in terms of linguistic, pedagogical, technological, and institutional realities remain under-investigated (Ajmal et al., 2025; Yasmin et al., 2025).

This gap is significant since the educational landscape in Pakistan is unique due to an unequal distribution of digital access, inconsistent exposure to AI, institutional heterogeneity, and a changing culture of academic honesty. These circumstances not only define the ability of students to use GenAI, but the quality and impacts of doing so. Although local research indicates that Pakistani students might use AI-based writing support, it also implies the necessity of context-specific guidance, especially regarding ethical writing, AI literacy, and feedback, and access equity (Ashraf et al., 2025; Yasmin et al., 2025).

Hence, the literature indicates a definite need to have a narrow synthesis that goes beyond such generalized statements about AI and explore the use of GenAI in the development of academic writing in Pakistani BS English students. The current research fills this gap by critically surveying the literature that exists in this field through the prism of learning gains, academic integrity, and assessment and contributing to the budding scholarship on AI in higher education writing in a more contextually grounded manner.

Research Design

In this study, a qualitative systematic review with PRISMA (2020) was used to provide transparency, reproducibility, and methodological rigor. The systematic review was deemed as suitable, as the objective of the research was not to create primary data, but rather to locate,

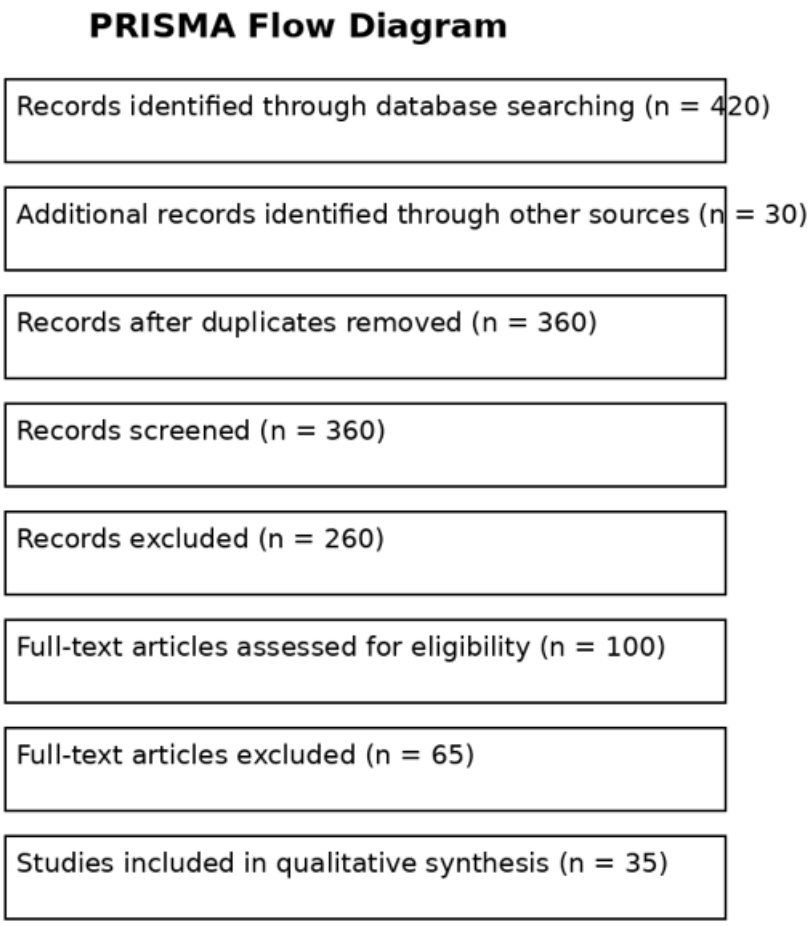
ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

analyze and critically synthesize available evidence pertaining to learning gains, academic honesty and implications of GenAI-assisted writing in post-secondary education. The approach to the review was based on the Preferred Reporting Items to Systematic Reviews and Meta-Analyses (PRISMA 2020) framework to serve as a transparency, replicability, and procedural rigor in the identification, screening, and selection of the relevant studies.

In contrast to a typical narrative review, a systematic review involves a rigorous and clearly defined procedure of finding, screening, evaluating, and synthesizing pertinent studies. Due to the swift development of the literature on ChatGPT and other similar tools since 2023, a systematic method was needed to differentiate between the general technological excitement and evidence that is specifically applicable to the development of academic writing in ESL/EFL higher education settings. The review thus narrowed down to the studies that dealt with GenAI as a learning aid in writing-related situations and how it is relevant to the three main issues of the current research: writing development, ethical application, and pedagogical practice.

A systematic search strategy was created to access studies that were applicable to the research objectives. The search was done in four significant academic databases, which include Scopus, Web of Science, ERIC, and Google Scholar. These databases have been chosen since they all index peer-reviewed literature in the fields of education, applied linguistics, language teaching, educational technology, and higher education research. Besides the database searching, a manual backward reference searching was performed by looking at the reference lists of the included articles in order to determine the potentially relevant studies that were not represented by the initial search.

The selection of the study was carried out in accordance with PRISMA guidelines, namely the identification, screening, eligibility, and inclusion steps. Records identified, screened, excluded, and included are reported to provide transparency.



ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Only works published in the past three years (since January 2023) were found. This period was specifically chosen since the mass educational use of ChatGPT and similar generative AI models became a reality only after late 2022, and the literature on the matter is not as directly applicable to the topic at hand. Limiting the review to this time frame ensured that the study was able to include the latest and conceptually pertinent literature on GenAI in the academic writing and at the same time, coherence in the scope of technology.

The search terms were constructed based on combinations of keywords and Boolean operators pertaining to the subject. The following search strings were used to represent the search:

- "generative AI" AND "academic writing"
- "ChatGPT" AND "ESL writing"
- "ChatGPT" AND "EFL writing"
- "AI-assisted writing" AND "higher education"
- "generative artificial intelligence" AND "Pakistan" AND "undergraduate"
- "ChatGPT" AND "English students" AND "academic writing"

These keywords were modified to fit slightly in databases according to the search interface requirements. The search strategy was created in such a way that it was broad enough to cover all topics and at the same time narrow enough to cover all studies relevant to the context of Pakistan. The search was performed between March-May 2025 and the search strings were modified in all databases with Boolean operators and truncation to be consistent.

Inclusion and exclusion criteria were set in advance to allow consistency and to cut down on selection bias. The studies were considered to have the following criteria:

1. They were peer-reviewed journal articles.
2. They were published in English.
3. They were published between 2023 and September 2025.
4. They examined the use of generative AI tools (e.g., ChatGPT or comparable large language model applications).
5. They focused on academic writing**,** ESL/EFL writing, or closely related writing development outcomes.
6. They addressed higher education learners, with particular attention to studies relevant to Pakistani undergraduate or BS English contexts, or to comparable EFL/ESL university-level populations.
7. They reported empirical findings, systematic evidence, or studies providing explicit analytical or empirical discussion of writing-related outcomes

Studies were excluded if they:

1. were conference abstracts, editorials, blog posts, theses, or unpublished opinion pieces;
2. focused on AI applications unrelated to writing;
3. addressed only technical or computational aspects of AI without pedagogical relevance;
4. did not involve higher education learners;
5. lacked sufficient relevance to learning outcomes, academic writing, academic integrity, or assessment**.**

These requirements were supposed to make sure that the end product of literature was methodologically relevant as well as substantively in line with the objectives of the review.

Study Selection Procedure

Any records that were retrieved were initially sorted and filtered to get rid of duplicates. After deduplication, the selection of the study was two-fold. In stage1, titles and abstracts have been filtered to show pre-relevance to the topic. In the second phase, the complete texts of the potentially eligible studies were filtered by the inclusion and exclusion criteria predefined. This

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

process of staging assisted in minimizing irrelevant records and keep possibly valuable studies to be evaluated further.

In order to enhance procedural trustworthiness, the screening process was conceptually structured based on independent review and solving of issues by consensus. In the event of an uncertainty about the eligibility of a study, decisions were made by re-reading and comparing them with the mentioned criteria. This method is in line with the established practice of systematic reviews and served to reduce arbitrary inclusion. Nevertheless, due to the synthesis of a somewhat recent and conceptually overlapping literature, close consideration was taken to the selection of studies that only mentioned GenAI as opposed to those that substantially covered writing-related learning outcomes. A total of 35 studies met the inclusion criteria and were included in the final analysis. The screening was done by two reviewers. Where there was uncertainty, the studies were re-examined to guarantee uniformity in selection.

Data Extraction

A data extraction matrix was created in order to guarantee systematical and similar registration of information in the involved studies. Pilot-testing of the extraction matrix was done on a subset of the studies to guarantee uniformity and thoroughness of data recording.

The following were extracted in each of the studies:

- author(s) and year of publication,
- country or educational context,
- study design,
- participant characteristics,
- type of GenAI tool used,
- writing-related outcomes,
- ethical or academic integrity concerns,
- assessment or feedback implications,
- key findings,
- and reported limitations.

This systematic method of extraction allowed the reviewer to not only compare studies based on their results, but also on the basis of their methodology, learning context and analysis focus. It also minimized interpretive inconsistency by making sure that all the included studies were analyzed using the same categories of analysis.

Quality Appraisal and Risk of Bias

Since the studies included were heterogeneous, qualitative appraisal approach was used, with the established criteria of methodological rigor as opposed to a unified checklist. These indicators included:

- clarity of research purpose,
- appropriateness of study design,
- adequacy of participant/sample description,
- transparency of data collection procedures,
- coherence of analysis,
- and alignment between findings and conclusions.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Articles with high levels of methodological uncertainty or low levels of relevance to the purpose of the review were handled cautiously during synthesis as opposed to being weighted to the same level as those with more substantial evidence. This was of particular concern since a significant portion of the literature on GenAI is still recent, exploratory, and uneven in terms of methodology. In this regard, the review did not consider all studies included as having a similar evidentiary value but interpreted the results based on the methodological robustness of a study and its contextual applicability.

Data Analysis and Synthesis

The inductive thematic analysis approach was used, and initial open coding was conducted, followed by the categorization into the themes in line with the research objectives. The analysis process consisted of a number of steps, including re-reading and inclusion of the included studies, concept repetition, primary coding of the relevant results, similar codes aggregated into descriptive patterns, and emergence of higher order levels of analysis. This methodology allowed the review to go beyond mere summary and have a more interpretive synthesis of patterns within the literature.

The thematic analysis was guided by the three broad analytical concerns established in the research objectives:

1. Learning gains in academic writing
2. Ethical and academic integrity implications
3. Pedagogical and assessment-related implications

These broad areas produced cross-study themes, which were inductively created based on the data. Indicatively, recurring patterns pertaining to grammar enhancement, coherence, revision aids, plagiarism issues, AI literacy, and formative feedbacks were found and subsequently arranged into interpretive themes. This analysis procedure enabled the synthesis of convergent results in addition to critical tensions in the literature, and thus generated a more balanced and analytically beneficial explanation of GenAI in scholarly writing.

Since the research is entirely a literature-based study, the research did not entail the participation of human subjects and thus did not necessitate the approval of formal ethics or consent of participants. The ethical responsibility was upheld by proper citation, true reflection of the reviewed researches, and prudent synthesis of the evidence according to the acceptable academic standards.

Results and Findings

Thematic analysis of the chosen articles revealed three overall themes about the use of generative artificial intelligence (GenAI) and ChatGPT, specifically, in academic writing by BS English students in Pakistan and other ESL/EFL settings: (1) learning benefits, (2) issues of ethics and academic integrity, and (3) implications in pedagogy and assessment. The use of GenAI tools was discovered to improve the performance in writing, especially in grammar, sentence structure, coherence, cohesion, and paragraph organization. These advantages were greatest in ESL/EFL situations and at revision phases. But the gains in higher-order skills like critical argumentation, originality and rhetorical sophistication were less uniform. The educational value was based on the interactive interaction with AI (e.g., brainstorming, revising) instead of passive generation of the content. AI can complement constrained writing assistance in Pakistan, but there is still a paucity of evidence on the long-term benefits.

AI-assisted writing creates issues of authorship, plagiarism, over-reliance, and reliability of the facts. Students can present refined text that does not reflect the intellectual effort on their own part, which could diminish critical and metacognitive ability. Inaccurate or fabricated data may also be produced by AI, which is why it is necessary to be critically literate and use AI as per institutional rules in order to be responsible. The use of GenAI is affected by

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

the teaching situation. AI literacy and critical evaluation, strategic prompting and revision are skills that students need. With formative feedback, AI can be used to assist many classes, though not to substitute the teacher judgment. The use of assessment practices has to change to focus on drafting, revision, commentary of reflection, and transparency in authorship in order to uphold validity. Teacher training, ethical instructions and systematic pedagogy are essential in achieving successful integration in Pakistan.

Generally, GenAI provides conditional assistance with scholarly writing, grammar, coherence and revision, but includes difficulties with authorship, ethical application, and assessment validity. It is maximally effective in conjunction with critical engagement, development of AI literacy, ethical guidance, and context-sensitive pedagogy.

Discussion

This current review investigated how generative artificial intelligence (GenAI) specifically ChatGPT can help ESL, EFL, or BS English students in Pakistan to write academic papers. The discussion is structured around research objectives and questions of the study. In line with the rest of the literature on the topic, the results suggest that GenAI has emerged as an even more significant academic support tool in higher education in terms of academic writing (Stanford, 2025; Holmes, 2019). In research, ChatGPT and other AI were mostly utilized to aid in shallow writing abilities, such as grammar correction, sentence composition, coherence, paragraph formatting. These results validate that GenAI is an additive, and not a substitute, to human thinking and this is in line with the previous studies that have indicated that AI is useful in aiding language learners with purposeful writing (Xiao et al., 2026). Based on the analyzed literature, it is proven that GenAI can stimulate quantifiable learning benefits, especially in such fields as accuracy in the language, clarity, and structuring of the text. The gains in higher-order skills like critical argumentation, originality, and disciplinary voice were however less uniform. This is in line with the literature that emphasizes AI as being more inclined towards mechanical and structural than cognitive or rhetorical sophistication in writing (Mekheimer, 2025; Yang and Lin, 2025). It is important to note that the usefulness of AI depended on the active engagement: students who engaged in the interactive process of brainstorming, revision, and reflection using AI demonstrated more significant improvements than those who used AI to get ready-made material. This substantiates the pedagogical claim that learning gains take place when technology supplements and does not replace active processes of writing (Chen and Chen, 2025). The results indicate a number of implications that could be made on teaching and assessment. In terms of ethics, GenAI brings up the issue of authorship, plagiarism, and over-reliance, following the previous studies on the topic of academic integrity in AI-assisted writing (Kuklick et al., 2023). Pedagogically, AI literacy, such as critical assessment, strategic prompting, and revision skills, is needed by students to be responsible and effective when utilizing AI. The assessment practice will also need to change; the evaluation of the final written work is becoming less relevant to reflect the actual learning and more in-writing processes, reflective commentary, and in-class writing should be prioritized (Holmes, 2019). These findings are of special interest in the Pakistani context where there are limited resources to support writing. AI has the potential to provide available and instant help, but its advantages have preconditions related to institutional leadership, educator training, and the incorporation into the system of pedagogies. The lack of such supports can put at risk the over- reliance, superficial engagement, and impaired academic integrity.

On balance, this review suggests that GenAI is a great tool that should be used to improve the writing process, but the ethical and pedagogical issues associated with this technology are dual-fold. Further studies are needed to explore the effects of AI use on the process of independent writing among students over the long term, learn about barriers to AI integration in higher education that are unique to various cultural practices, and find ways to

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

incorporate the use of AI in the ESL/EFL curriculum in a responsible way. The results suggest to practitioners the necessity of clear teaching of AI literacy, institutional policies delineating responsible use, and redesign assessment to guarantee genuine presentation of writing proficiency.

Conclusion

This paper is a systematic literature review of the application of generative artificial intelligence (GenAI) and, specifically, ChatGPT in academic writing among higher education students with a particular emphasis on BS English students in Pakistan. The results show that GenAI has the potential to enhance the development of the writing process by enhancing the grammar, coherence, sentence structure, and paragraph organization. These advantages can be best realized when students utilize AI more as a brainstorming, revising, and reflection tool, as opposed to a source of prepared content. Nevertheless, the review also mentions serious ethical, pedagogical and assessment-related problems. The issues of authorship, plagiarism, over-dependence, and inaccurate AI responses underline the importance of critical interaction, AI literacy, and institutional policies. In addition, the conventional methods of assessment might have to be redesigned to consider AI-assisted writing and guarantee that they show genuine competence among the students. On the whole, GenAI has a considerable yet contingent potential in academic writing teaching. It is effective when carefully incorporated into pedagogies, clearly taught how to use it responsibly and correlated with learning goals. Future studies ought to examine the long-term effects on writing, and culturally sensitive approaches to responsible AI incorporation in higher education.

References

Ajmal, A., Aslam, S., Yousaf, M. S., & Raza, M. A. (2025). USING CHATGPT TO ENHANCE ESL WRITING SKILLS IN PAKISTANI HIGHER EDUCATION. International Journal of Social Sciences Bulletin, 3(3), 666-673.

Ashraf, M. A., Alam, J., & Kalim, U. (2025). Effects of ChatGPT on students' academic performance in Pakistan higher education classrooms. Scientific Reports, 15(1), 16434.

Chanpradit, T. (2025). Generative artificial intelligence in academic writing in higher education: A systematic review. Edelweiss Applied Science and Technology, 9(4), 889-906.

Chen, Y., & Chen, J. (2025). Generative AI in EFL writing instruction for lower-intermediate learners: Do perceived affective gains translate to improved performance?. *Language Teaching Research*, 13621688251397724.

Holmes, W. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. *Center for Curriculum Redesign*.

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.

Kuklick, L., Greiff, S., & Lindner, M. A. (2023). Computer-based performance feedback: Effects of error message complexity on cognitive, metacognitive, and motivational outcomes. *Computers & Education*, *200*, 104785.

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, *11*(1), 9.

Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: a study on proficiency, revision frequency and writing quality. *Discover Education*, *4*(1), 170.

Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). A systematic review of generative AI for teaching and learning practice. *Education Sciences*, *14*(6), 636.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

Radtke, A., & Rummel, N. (2025). Generative AI in academic writing: Does information on authorship impact learners' revision behavior?. *Computers and Education: Artificial Intelligence*, *8*, 100350.

Shi, J., Liu, W., & Hu, K. (2025). Exploring how AI literacy and self-regulated learning relate to student writing performance and well-being in generative AI-supported higher education. *Behavioral Sciences*, *15*(5), 705.

Stanford, J. (2025). How disruptive is AI in students' academic writing in higher education?. *Studies in Technology Enhanced Learning*, *4*(1).

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: insights from a meta-analysis. *Humanities and Social Sciences Communications*, *12*(1), 1-21.

Xiao, F., Zhu, S., & Xin, W. (2026). Exploring the landscape of generative AI (ChatGPT)-powered writing instruction in English as a foreign language education: A scoping review. *ECNU Review of Education*, *9*(1), 20965311241310881.

Yang, W., & Lin, C. (2025). Translanguaging with generative AI in EFL writing: Students' practices and perceptions. *Journal of Second Language Writing*, *67*, 101181.

Yasmin, M., Fatima, W., & Irshad, I. (2025). Evaluating ChatGPT's effectiveness in enhancing argumentative writing: A quasi-experimental study of EFL learners in Pakistan. *Sustainable Futures*, *10*, 100809.