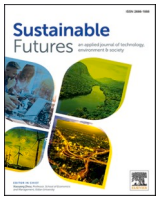
Sustainable Futures 10 (2025) 100809

Contents lists available at ScienceDirect

Sustainable Futures

journal homepage: www.sciencedirect.com/journal/sustainable-futures



Evaluating ChatGPT's effectiveness in enhancing argumentative writing: A quasi-experimental study of EFL learners in Pakistan

Musarat Yasmin *, Walees Fatima, Isra Irshad

Department of English, University of Gujrat, Hafiz Hayat Campus, Gujrat, Pakistan

ARTICLE INFO

Keywords:
ChatGPT
Argumentative writing skills
EFL education
SSC students
Pretest/post-test
Error analysis

ABSTRACT

A crucial area of ongoing research is the potential of artificial intelligence to improve learning outcomes in diverse educational settings, given its increasing global integration into education. This study assesses the effectiveness of ChatGPT, an artificial intelligence tool, in improving the argumentative writing skills of Pakistani students pursuing the Secondary School Certificate (SSC) in English as a Foreign Language (EFL) environment. The focus of this study is to assess the impact of ChatGPT on the skills of argumentative writing of SSC students to evaluate to what extent they improve the key components of argumentative writing through pre-test and post-test, and to identify the specific types of errors committed by the learners in Pakistani context. This study adopted a quasi-experimental approach, using pre- and post-tests to measure changes in students' argumentative writing abilities after three months of ChatGPT interaction. The data from the pre-test and post-test is analyzed by using key components of argumentative writing by Tolumin (1958) and the framework of error analysis by Ellis (2002). The findings demonstrate significant progress in students' argumentative writing, with a corresponding decrease in error types after implementing the ChatGPT intervention. The study's results indicate a significant improvement in the argumentative writing abilities of secondary school students using ChatGPT, suggesting its usefulness in supporting error correction and enhancing writing skills in English as a Foreign Language education.

1. Introduction

Artificial intelligence is leading a profound revolution in 21st-century education, fundamentally altering its structure. Internationally educational systems are integrating AI to improve the education quality and outcomes signifying a departure from traditional teaching approaches. This incorporation of AI opens new and unique opportunities to address the educational challenges and enhance the efficacy of pedagogies, particularly in language education.

1.1. Technological advancements in education

With a progress in digital technologies, their incorporation into education marked a shift in teaching and learning pedagogies. These developments have resulted in a more interactive, personalized, and accessible educational paradigm, substantially broadening the reach and effectiveness of education [1]. The 2018 Programme for International Student Assessment (PISA) revealed that 71 % of United States students utilized laptops in their classrooms [2]. Secondary students utilize educational technologies such as Carnegie Learning's tutoring platform (mathematics, science, and English) and ASSISTments (mathematics practice) to foster collaboration, creation, and the sharing of work. Covid-19 accelerated the trend through offering new learning tools [3]. The widespread adoption of large language model-based generative AI tools like OpenAI's ChatGPT in early 2023 has necessitated a reconsideration of educational technology by both educators and academic policy makers [4,5]. Soori et al. [6] demonstrated the efficacy of AI in enhancing decision-making capabilities and student achievement in educational settings. Utilizing predictive analytics within AI, learning patterns are statistically analyzed to identify students at risk of academic failure, thereby enabling proactive interventions [7]. They established that AI-assisted algorithms help the teacher in scoring assignment and hence reduce his workload. Conversational agents, chatbots, are found helpful as pedagogical tools and for registration [8]. Haider and Chowdhury [9] examined the incorporation of Communicative Language Teaching within a CALL framework in Bangladesh, noting positive learner responses and areas for pedagogical improvement. Dwivedi and Seema [10] surveyed various CALL approaches in

* Corresponding author.
E-mail address: musarat.yasmin@uog.edu.pk (M. Yasmin).

https://doi.org/10.1016/j.sftr.2025.100809
Received 8 April 2025; Received in revised form 17 May 2025; Accepted 7 June 2025
Available online 8 June 2025

India, highlighting technology's role in enriching language education. Jamaldeen et al. [11] explored the potential of Mobile Assisted Language Learning among Sri Lankan students, proposing a theoretical framework for its implementation. Additionally, Yeasmin and Chowdhury [12] investigated Bangladeshi teachers' perceptions of technology-enhanced blended learning environments, emphasizing the benefits for collaborative learning. These studies highlight the region's tendency to integrate technological advancements in language education.

1.2. The role of AI in language learning

Language acquisition among students serves diverse purposes, including, but not limited to, professional and academic pursuits. Artificial Intelligence (AI) has proposed different tools, such as ChatGPT, that are being used in education for the improvement of language learning. ChatGPT provides substantial advantages to those studying English as a second or foreign language. This tool has an approach to process, synthesize, and analyze speech, writing, images, etc. [13]. The impact of ChatGPT on students' learning activities is noteworthy. According to Ratnam et al. (2023), ChatGPT aids students in establishing learning objectives, reviewing assessments, and following appropriate evaluation processes, influencing their learning efficacy. They started a pilot study and found that it is systematic and very helpful in use as it helped them to write coherently linked, with full cohesion, accurately, and in a systematic way. It was suggested that some modified learning objectives will enhance the capabilities of students to think creatively and critically. They suggested that it cannot help the students in thinking and broad development of their mentality, but the students should be trained as they can easily carry their subjects through ChatGPT.

ChatGPT's function as a learning scaffold is elucidated by Vygotsky's sociocultural theory, with its emphasis on the Zone of Proximal Development (ZPD). Higher-order skills are cultivated within this framework through the guidance of more expert entities, including AI technologies such as ChatGPT. Through immediate feedback, structured argumentation modeling, and independent editing facilitation, ChatGPT serves as a digital scaffolding mechanism, fostering learner autonomy in writing. This approach is consistent with constructivist language-learning methodologies, which prioritize active student participation, reflective practice, and iterative refinement.

Researchers and scholars can create and generate their projects by using ChatGPT in their work and they concluded that some required evaluation is necessary for the development of creativity and critical thinking [14]. Green and Chen [15] argued that structures and errors are improved due to the utilization of AI. It is discussed that the committed errors are controlled with the assistance of ChatGPT. Research showed that teachers trust ChatGPT in affecting the writing skills of foreign language learners [16,17]. Semi-structured interviews were used as a tool to collect the data from the teachers of a foreign language from Indonesian Universities and analyzed to examine what the impact on students. The outcome of this study revealed that writing tools of AI positively assist the students in their writing skills [16]. They found that AI tools have a positive impact on students' writing by improving content and organization. The AI-based tools also enhanced vocabulary use, coherence, and idea generation. Instructors acknowledged the AI tools but also warned about over-reliance that hinders critical thinking. Sheikh et al. [18] found that ChatGPT assists learners in learning English as a foreign language. A questionnaire was utilized to gather data from the second language learners. Shaikh et al. [18] determined that ChatGPT facilitates vocabulary and pronunciation acquisition among learners. Cao and Zhong's [19] study revealed that ChatGPT responded to prompts rapidly, yet exhibited limitations in feedback efficacy. They posit that ChatGPT excels at enhancing the cohesion and coherence of written text. The key components of writing argumentative essays are also assisted by the employment of ChatGPT. Smith [20] concluded that Artificial Intelligence has much importance in the field of education and is used for the embellishment of writing skills. He found that the key components of argumentative writing are refined by the utilization of ChatGPT. The feedbacks of learners are also offered by ChatGPT and grounded that EFL learners are shifted towards modern technology instead of traditional methods. A comparative study by Herbold et al. [21] analyzed student-written and ChatGPT-generated argumentative essays. Argumentative essays were produced by expert educators and ChatGPT concurrently, both addressing the same topics. The study by Herbold et al. [21] demonstrated that ChatGPT produced higher-quality essays compared to human-generated texts. He compared essays written by human experts with those generated by ChatGPT. Both writings from humans and AI were quite different. It was found that ChatGPT produces better argumentative essays than human. ChatGPT performs better in producing logical structure, language complexity, vocabulary and cohesion. AI writings used complex language and sentence structures. It also made it more formal and exhibited diverse language as compared to human who used more phrases.

1.3. Challenges in EFL education in pakistan

The educational system in Pakistan, specifically EFL instruction, faces significant obstacles to effective language acquisition. Yasmin et al. [22], in this regard, found three major barricades: sociocultural, psychological, and institutional constraints. A Pakistani classroom exhibits a strong reliance on traditional teaching methodologies, characterized by learner's dependence on the instructor and showing a passive learner engagement [23]. The teachers' demeanor was perceived as authoritative, and they were reluctant to provide students with opportunities for creative expression or questioning [24]. The inadequacy of resources in Pakistani classrooms restricted student access. A detailed analysis of writing proficiency among 40 secondary schools' tenth-grade students in the Bahawalnagar District, Pakistan, was performed [25]. The assessment encompassed vocabulary, syntax, reading comprehension, grammatical accuracy, and handwriting. A sample of 440 students, representing urban and rural public and private schools, was analyzed to assess gender, density, and school type differences using mean scores, standard deviations, and *t*-tests. Research findings indicate superior student comprehension performance compared to other assessed skills, suggesting focused areas for pedagogical intervention. Researchers also focused on the area of error in writing. In this instance, Albuhairy [26] highlighted that the influence of the mother tongue in teaching and learning is the primary cause of errors. The learners also copy the essence of the structure of the first language. He also found that the learners also overgeneralize and replicate the first language structures in others. Research also showed that Pakistani students face issues in grammar and argumentation in academic writing due to the lack of instructions. A fundamental flaw in traditional teaching methodologies is the overemphasis on memorization at the expense of critical analysis. Aziz and Rawian [27] found that traditional teaching methods encourage rote memorization over critical thinking among students. He discovered that it resulted in deficient writing skills due to a lack of student training in the organization and the presentation of evidence in written form.

1.4. Rationale for using ChatGPT in EFL settings

In Pakistan, where English serves as a crucial medium for academic and professional advancement, mastery of argumentative writing is essential for Secondary School Certificate (SSC) students. Despite the increasing integration of technology in education, there is limited research on the efficacy of AI tools like ChatGPT in enhancing English writing skills among EFL learners in the Pakistani educational context. It can be anticipated by keeping in view the previous research that if Pakistani learners are provided proper guidance, and technological advancement, AI can help and improve the writing skills of Pakistani EFL learners. Present study intended to explore the utilization of

ChatGPT in language learning, especially to improve writing skills at secondary school level in Pakistan. The intervention examined the nature of student errors and the extent to which ChatGPT use influenced the frequency and types of errors in student writing. Present study addressed the following research question:

How does the inclusion of ChatGPT impact the argumentative writing of SSC students in Pakistan as EFL learners?

a. How does the use of ChatGPT influence the learning of key components of argumentative writing among Pakistani Secondary School Certificate (SSC) students as EFL learners?
b. What specific types of errors do secondary school certificate students commit in the Pakistani context in pre-test and post-test?

2. Research methodology

The current study addresses the employment of ChatGPT among EFL learners in argumentative writing. The study employed a quasi-experimental approach to analyze writing development as measured by pre- and post-tests. The study employed a one-group pre-test/post-test design due to administrative constraints in accessing parallel classes. While this limits causal attribution, the same instructional context and teacher were maintained across the intervention, and the pre-test served as a baseline.

2.1. Sample and procedure

The study utilized a purposive sample of 30 SSC-level students, equally drawn from a boys' and a girls' public-sector school in an urban district of Punjab, Pakistan. This gender-balanced selection process guaranteed representation from diverse educational backgrounds. The chosen schools adhered to the national curriculum and served students from middle-socioeconomic backgrounds, reflecting the composition of numerous urban public schools nationwide in Pakistan. Despite the limited sample size, its composition accurately mirrors the main demographic characteristics of mainstream English as a Foreign Language (EFL) students in Pakistani Secondary School Certificate (SSC) contexts. Stratification by prior proficiency was not feasible due to institutional assessment limitations, which is acknowledged as a limitation. However, the authors recognize the presence of broader regional, linguistic, and socioeconomic differences and suggest future research employing more diverse samples to improve the generalizability of findings. These students completed pre- and post-tests, as well as the intervening session. To protect the anonymity of research participants, alphabetic identifiers have been substituted for their names. The data was collected from the argumentative essays through the pre-test and the post-test among Pakistani SSC students. In the first session, students were given a short introduction to argumentative essays and then a pre-test. Subsequently, a 45-minute lecture was delivered, during which participants were instructed to utilize ChatGPT for their writing assignments and this continued for three months period. Upon completion, a post-test assessment was conducted. The pre-test and the post-test were analyzed by using key components of argumentative essays by Toulmin [28] and steps of error analysis by Ellis [29]. At the end of the analysis, both tests were compared to know the changes over the time duration from the pre-test to the post-test, see Fig.1 for complete methodological review.

Although the study adopted a quasi-experimental design without randomization or a control group, internal validity was maintained by using a within-subjects design. The same group of students was assessed before and after the intervention, with no changes in instructional content or environment except the introduction of ChatGPT. This allowed the study to attribute changes in performance primarily to the use of ChatGPT. Future studies employing randomized control groups

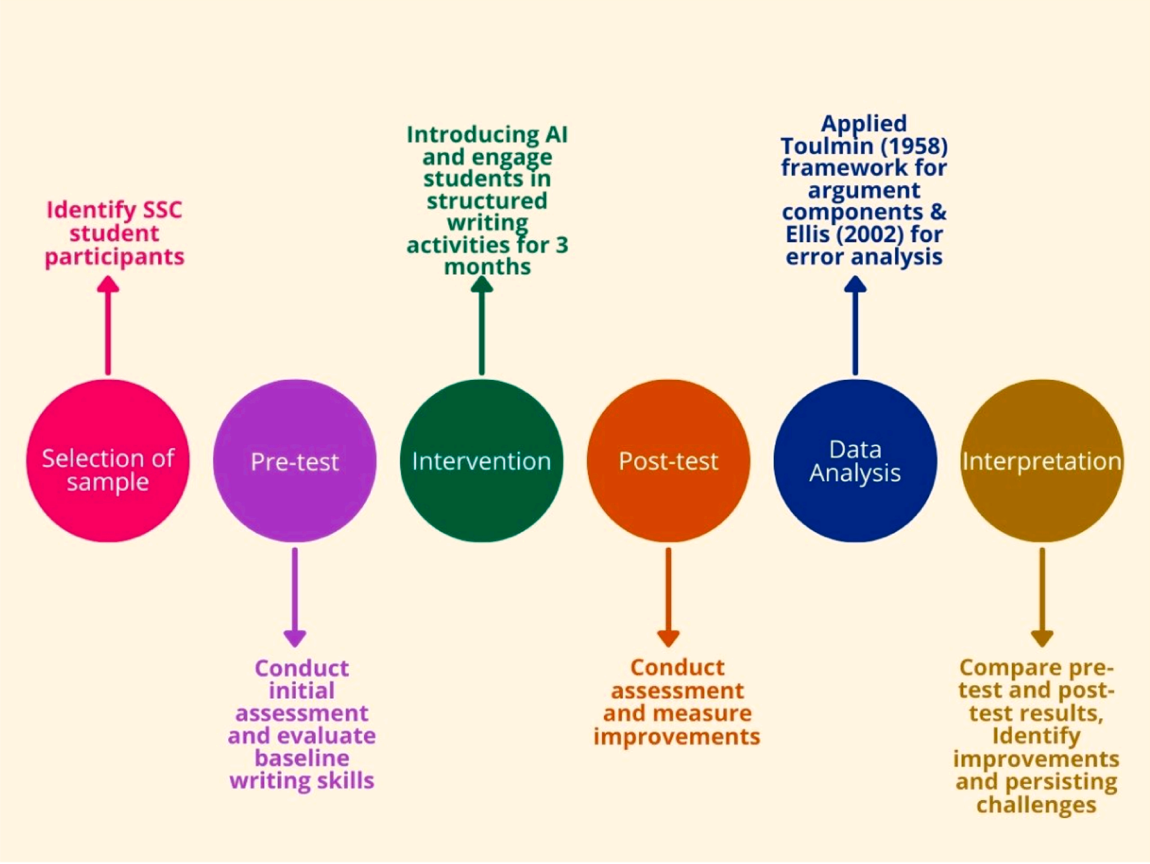


Fig. 1. Intervention Process.

would further strengthen the causal claims.

OpenAI's ChatGPT (GPT-3.5) was employed by the students. The prompts concentrate on the generation and revision of argumentative paragraphs. Examples include: "Generate a counterclaim for the argument that social media is harmful," or "How can I improve this thesis statement?" A rubric structured according to Toulmin's model was provided to students, highlighting the importance of clear claims, relevant evidence, and effective counter argumentation.

The significance of observed changes between the pre-test and post-test scores was analyzed by inferring the statistics. The specific test utilized was a paired sample T-test to determine whether the mean difference between pre-test and post-test intervention scores is statistically significant. SPSS v.26 was used for all analyses to determine statistical significance. The observed effect magnitude and precision were assessed using 95 % confidence intervals and effect size (Cohen's d).

Informed consent was obtained from all students and school authorities prior to the study. No personally identifiable information was gathered. Students received instruction on the function of artificial intelligence in composition, adhering to established institutional and ethical guidelines. The study did not involve sensitive data and complied with local AI-use policies and research ethics guidelines.

2.2. Analytical framework

The skill to articulate the ideas and persuade others is difficult to understand. Argumentative writing is used to increase problem-solving. The data is analyzed by using the framework of key components of argumentative essays by Toulmin (1958). Key components included claim, evidence, reason, and counterclaim. This study utilizes Ellis's [29] error analysis framework, a refinement of Pit Corder's [30] foundational research, to analyze the argumentative essay. This framework comprises four critical steps: identification, where common errors are categorized; description, which investigates the causes behind the errors; explanation, which applies standard rules to understand the errors; and evaluation, where both local and global errors are assessed for their impact. This approach ensures a focused and coherent analysis of errors within the essay.

3. Results and discussion

A comparative analysis of student essays, conducted before and after ChatGPT implementation, revealed a demonstrable impact of AI on student writing. Analysis revealed differential impacts of ChatGPT usage on various aspects of argumentative composition.

3.1. Impact of ChatGPT on key components of argumentative writing

The pre-test and the post-test were held to observe the impact of ChatGPT integration on the key components (claim, reason, evidence, and counterclaim) as established by Tolumin [28] in the writings of the SSC students. The influence of ChatGPT on students' argumentative writing skills is illustrated in Fig. 2. Pretest data, gathered prior to ChatGPT implementation, indicate that students demonstrated below-average writing proficiency. Post-test analysis demonstrate significant improvement, with the majority of students exhibiting above-average performance on core argumentative writing skills. This implies that employing ChatGPT as a pedagogical instrument significantly improves students' writing proficiency.

Student language learning progress, from pre-test to post-test, shows significant development in argumentative writing skills. While taking components individually, it is clear from Fig. 3 that the maximum number of students did not write the 'claim' accurately; however, post-test results indicate improved performance in claim component construction following ChatGPT utilization. It is worth mentioning here that

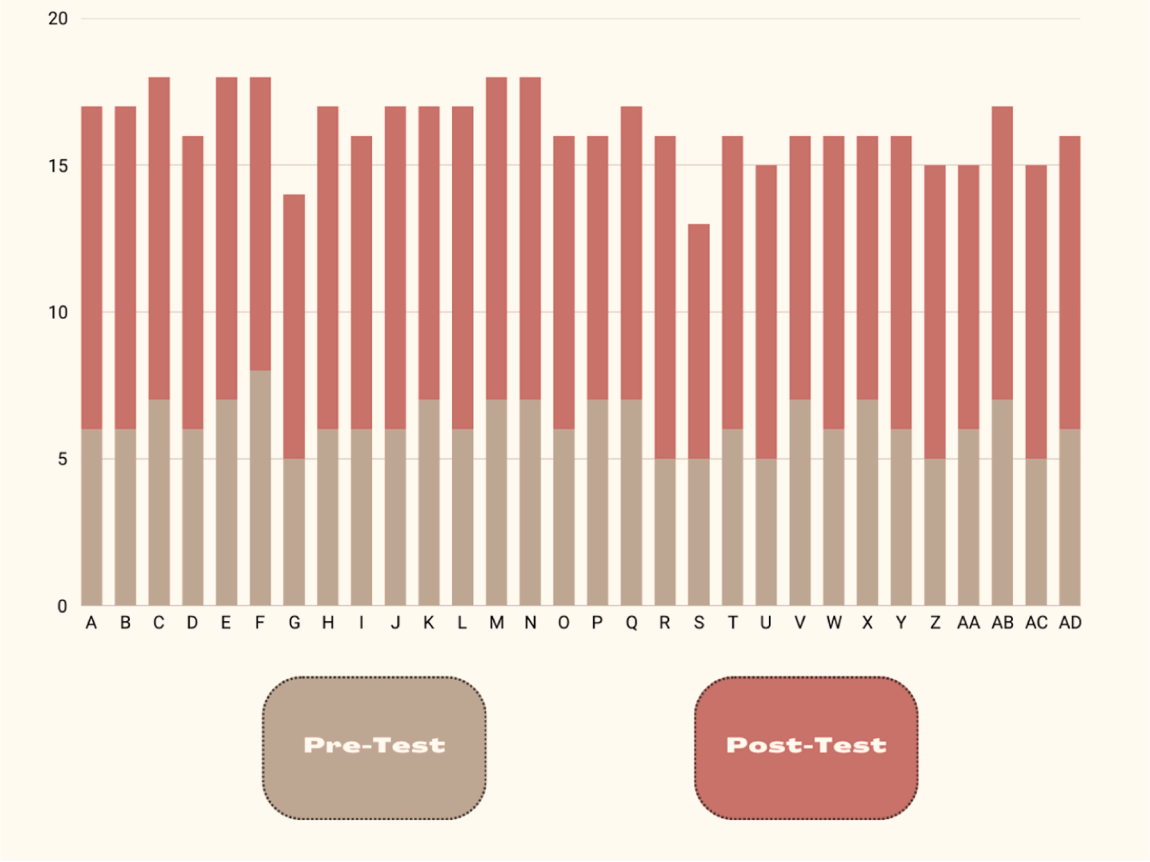


Fig. 2. Impact of ChatGPT on Key Components of Argumentative Writing.

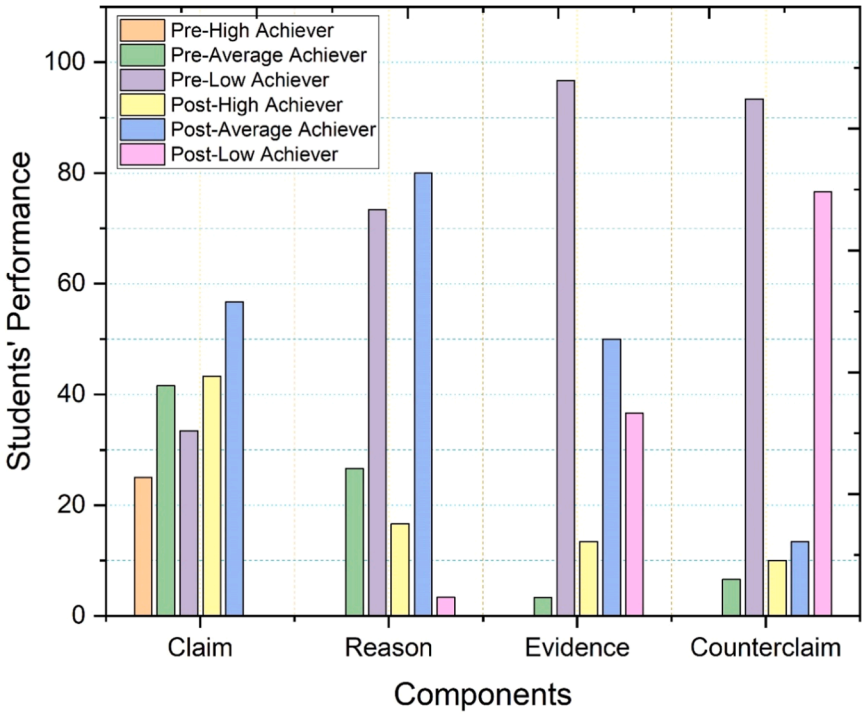


Fig. 3. Students' Performance in Key Components of Argumentative Writing.

students were marked and categorized under 3 levels: High achievers (80 %+), average (score 51–79 %), low achievers (less than 50 %). Researchers tried to align their levels with standard educational bench mark. Upon initial assessment, student performance in claim construction was largely average, with 41.60 % achieving an average score, 33.4 % falling below average, and a mere 26.70 % attaining high scores. Post-test results indicate a substantial improvement, with 43.30 % of scores categorized as high and an absence of low scores. As an example, a comparison of Student A's pre-test statement ("Social media is good") and their post-test response ("Social media promotes civic engagement and real-time activism, especially among youth, which strengthens democratic participation") reveals a significant improvement in the specificity and sophistication of their argument. These results imply that ChatGPT assisted students in developing stronger argumentative skills, specifically in the construction of clear and forceful claims, potentially through illustrative examples or highlighting the importance of precision.

The 'Reason' component demonstrates a significant improvement. In the initial assessment, no students demonstrated high-level reasoning skills in their argumentative essays. Post-intervention analysis reveals a noteworthy 16.60 % increase in student success. The results suggest a significant positive correlation between ChatGPT usage and improved reasoning skills among certain students. Also, post-test results indicated a marked improvement, with a dramatic drop in the number of students scoring below the benchmark; the rate fell from 73.40 % to 3.40 %. This suggests that the intervention proved especially beneficial for participants who initially experienced the greatest difficulty in developing argumentation within their essays. For instance, student C's reasoning, initially described as "helpful for people," evolved to a more nuanced assessment: "It helps people stay informed through rapid updates, reducing the digital divide in information access," demonstrating enhanced analytical depth. A notable shift is evident in the substantial increase of students categorized as moderate, rising from 26.60 % to 80.00 %. These results indicate that, although few students attained the highest performance level, a significant proportion progressed from low to moderate reasoning abilities in their writing.

Analysis of the 'evidence' component revealed a transformation in student performance. A pre-test revealed that an overwhelming 96.60 % of students scored low, with a complete absence of high scorers. Following the intervention, 13.40 % of students earned high scores, demonstrating that ChatGPT assisted a group of students in successfully incorporating evidence; these students' writing displayed more sophisticated argumentation and a greater depth of analysis. While in case of low scorers, a significant reduction was observed in the proportion of students achieving low scores, declining from 96.67 % in the pre-test to 36.60 % in the post-test. The marked reduction in errors suggests the intervention's considerable success in supporting students previously experiencing difficulty with appropriate evidence use. Of particular note is the substantial growth in the student population now categorized as moderate, a jump from 3.33 % to 50.00 %. This shows that though many students were not able to score high, majority improved their evidence presentation in posttest. For instance, student O posited a personal conjecture as evidence: "I believe…" Whereas in post-test, he was able to recognize his limitation and produced credible evidence to support her claim, 'According to a 2022 Pew Research study, 67 % of young adults use social media as their primary news source'. ChatGPT's training methodology seems to have focused on the importance of evidentiary support for arguments, guiding students in the proper integration and citation of relevant sources, thereby enhancing their overall argumentative proficiency.

Although the counterclaim category exhibited the least improvement, it demonstrated a notable shift from a substantial majority of low-scoring students (93.30 %) to a smaller proportion (76.60 %), coupled with a modest increase in moderate (6.67 % to 13.30 %) and high-scoring students (0.00 % to 10.00 %). The result may demonstrate the complexities inherent in managing counterarguments, necessitating the acknowledgment and effective refutation of opposing perspectives. For illustration, student C gave no counterclaim in pre-test; however, he showed an improved ability to acknowledge opposing views in post-test, 'Some argue that social media spreads misinformation, but, fact-checking features have become more strong in recent platforms.' Utilizing ChatGPT in instruction potentially provided students with diverse perspectives and models for the integration and refutation of counter-arguments—a complex skill in argumentative writing.

The findings demonstrate the efficacy of ChatGPT integration in enhancing students' argumentative writing, specifically their ability to construct claims, reason effectively, utilize evidence, and address counterarguments. The results show a notable increase in high achievers and a significant decrease in low scorers in each category. This shift reveals that interactive feedback and specifically designed illustrations provided by ChatGPT increased students' skills in developing coherent and persuasive arguments.

3.2. Impact of ChatGPT on students' errors

Error analysis shows ChatGPT substantially improved student writing by reducing linguistic and structural errors. Error analysis is used to analyze the errors committed by students in argumentative essays. Pit Corder [30] discussed the error analysis in "The Significance of Learner Errors". He wrote that important errors should be highlighted by teachers and then those errors must be eliminated. The language composition and writing performances are discussed by Ellis (2002). His discussion included a range of prevalent errors, such as omissions, misinformation, grammatical flaws, inappropriate word usage, literal translations, punctuation errors, disordered sentence structure, inaccurate use of numbers, abbreviation, prepositional misuse, differentiation, spelling mistakes, spacing issues, improper use of articles, and flawed differentiation.

Fig. 4 indicates a substantial reduction in student writing errors following the incorporation of ChatGPT into the learning practices. Pretest results indicated that a substantial majority (63 %) of students exhibited an average error frequency, while 30 % performed poorly by demonstrating a high error frequency and only 7 % performed well and displayed a low error frequency. However, posttest results showed a decrease of 10 % as compared to previous frequent errors as this time only 20 % had higher frequency of error. Concurrently, an increase in student performance corresponded with a rise in the percentage of students (43 %) demonstrating a low error rate. A comparable trend was observed within the average-frequency category, exhibiting an approximate halving of the count (from 63 % to 37 %). Present finding suggests that integration of ChatGPT reduces error rate in students' writings.

Comprehensive analysis of the research concerning ChatGPT integration and its effects on students' writing abilities shows diverse outcomes for various error types, pre- and post-intervention, as illustrated in Fig. 5. The results showed an improvement in the errors related to Grammar, punctuation, preposition use, disordering and use of articles. Though there is drastic improvement in these four types, errors were almost halved in grammar, reduced thrice in punctuation and reduced four times in case of disordering and use of preposition. A significant reduction in errors was observed in spelling, with student errors decreasing from 49 in the pretest to 3 in the posttest. Whereas, the remaining error categories—omissions, differentiation errors, and spacing issues—showed slight improvement.

There were also some error types which remained unchanged in this intervention e.g. literal translation, word-usage, abbreviation, and misinformation. The intervention may not have adequately accounted for the complexities inherent in interlingual translation. This may be attributed to instances where ChatGPT's outputs closely mirror source material structures or idioms, resulting in misinterpretations across cultural or contextual boundaries. A consistent pattern in word-usage errors suggest that either ChatGPT input was not influential or students were not able to grasp suggestion provided by AI. While the absence of progress in outcomes related to abbreviations might be attributed to ChatGPT proposing or employing them, which students subsequently integrate without a complete grasp of their proper usage or significance.

This underscores a lack of understanding of formal writing standards or subject-specific jargon. A similar pattern was observed in errors of misinformation, that shows that the use of ChatGPT did not significantly impact students' abilities to accurately interpret and communicate factual information. This may suggest that artificial intelligence may help in language mechanics but are unable to address the inaccuracy of content. The implementation of critical thinking and fact-checking skills is necessary for both teachers and students.

3.3. Statistical significance of the error reduction

A paired samples T-test was conducted to determine whether ChatGPT-based instruction effectively reduces EFL learners' writing errors. Results showed a notable reduction in error mean counts before the treatment ($M = 10.03$ SD $= 2.74$) compared to after the treatment ($M = 4.73$ SD $= 1.80$), resulting in a mean difference of -5.30 errors. This improvement was statistically highly significant, $t(29) = 9.44$, $p < .001$, and statistically significant at the 95 % confidence interval of 4.15 to 6.45. This difference means that this confidence interval does not include zero, indicating a reliable and consistent improvement across

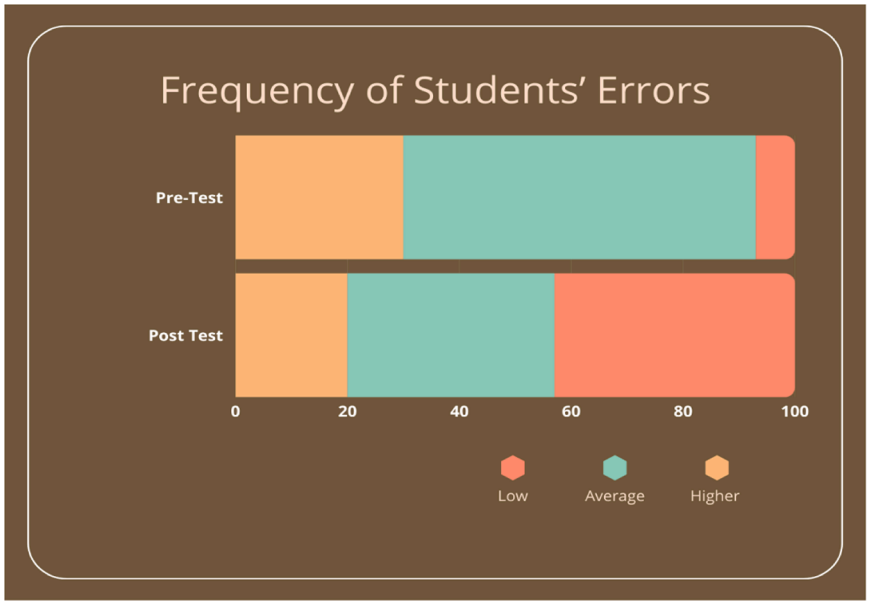


Fig. 4. Impact of ChatGPT on Students' errors.

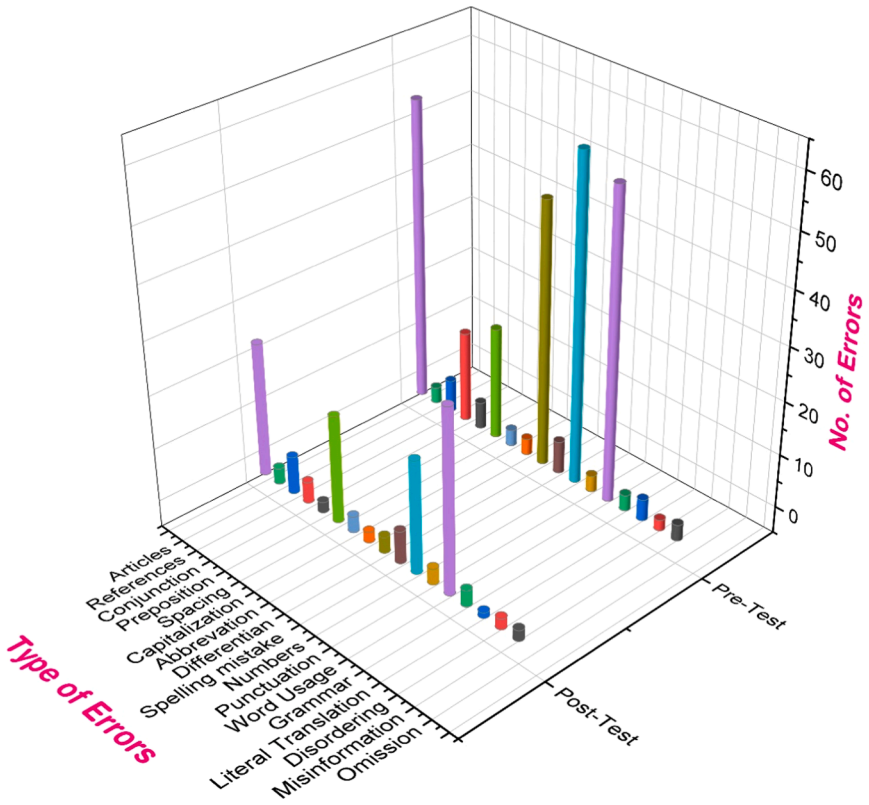


Fig. 5. Impact of ChatGPT Integration on Students' Error Types.

participants.

The magnitude of change resulting from the intervention was calculated to be very large on the effect size (Cohen's d) measure (1.72). Thus, ChatGPT was helpful for students in decreasing writing errors, probably because the feedback is more precise, the suggestions are more structured, and they can view the corrections. Overall, the statistics view favors using machine-based instructional models for language improvements (Table 1).

3.4. Discussion

The findings show a marked improvement in students' argumentative writing skills following the integration of ChatGPT into their learning. The students exhibited a remarkable progress in learning key components of argumentative writing: claim, reason, evidence, and counterclaim. The results also showed a reduction in linguistic and structural errors. This paper contextualizes the presented results within the existing body of research on AI-assisted language learning, further examining their pedagogical ramifications for English as a Foreign Language (EFL) instruction in Pakistan. The analysis of the above data shows that AI tools can be used for educational purposes. It encourages the teachers to use AI for teaching and involves the students in language learning. The focus of the current study is to know the impact of ChatGPT in the improvements of key components and reducing the error types. These are observed from the pre-test and the post-test. The outcomes are discussed in the light of different research. The students gain knowledge through the utilization of ChatGPT. The findings showing students' improved writing skills are consistent with Xiao [31] who concluded that technology has different challenges in language learning but it has benefits for interactive materials. Our results corroborate Kim et al. [32] findings that AI-based writing feedback significantly improves the structural aspects of student writing, leading to better organization and more cohesive essays. It is used to enhance the language learning skills. The employment of ChatGPT has influences in teaching methods and also it is a great benefit for students for language learning. The findings correspond with another research by Athanassopoulos [33] whose findings indicate that ChatGPT is a language learning and teaching for the improvement of language learning skills [33]. The key components of writing argumentative essays are also assisted by the employment of ChatGPT. The outcomes are aligned with the findings of Smith (2020) who concluded that Artificial Intelligence has much importance in the field of education. It was related that is used for the embellishment of writing skills. He (2020) found that the key components of argumentative writing are refined by the utilization of ChatGPT. The current study also found that the utilization of ChatGPT has enabled the students to write and maintain the key components. The exploration of the pre-test, before the utilization of ChatGPT, and the post-test, after the employment of ChatGPT, show improvements in writing. It has developed the abilities for the formulation claims, reasons, evidence, and counterclaim. These are improved by 15 % to 20 % from before utilizing ChatGPT to after its utilization. This is also supported by the work of Deane (2023) who suggested that writing with the help of AI is better than other writings. It has substantial gain in the structure of writing and improves the arguments. This is very close to the research of Herbold (2023). He performed his research on the quality of essays, written by ChatGPT, and reported that these essays are better than the essays which were written by human beings. The current research is

Table 1
Paired-Samples T-Test Results.

Measure	Pre-Test (V1)	Post-Test (V2)	Difference (V1 – V2)
Mean	10	4.73	5.3
Standard Deviation	2.74	1.8	3.08
Standard Error	0.5	0.33	0.56
t-value	–	–	**9.44**
Df	–	–	**29**
p-value	–	–	**< 0.001**
Cohen's *d*	–	–	**1.72**

established on the utilization of ChatGPT for the writing of argumentative essays.

The current study identified the errors committed by Pakistani SSC students in pre-tests and found a significant reduction after integration of ChatGPT as learning tool. A survey of existing literature reveals that grammatical inaccuracy is a persistent obstacle for ESL learners, resulting in frequent errors in their writing [34]. Our findings echo the results of Green and Chen (2019) who demonstrated that AI assisted tools mitigate errors by providing real time feedback. This study showed a significant decrease in grammatical and punctuation errors, indicating that students improved their writing precision with the help of ChatGPT's feedback mechanisms. Students' consistency in some errors like misinformation and errors due to interference of first language highlight necessitates the supplimental instruction in critical thinking and contextual understanding, a finding corroborated by Liu and Liu and Liu [35]. This study expands upon existing literature, incorporating findings from Mahboob's [36] research, which attributes errors to first-language interference.

Besides observed advantages, certain limitations of ChatGPT also surfaced. Some errors including literal translation, word-usage, abbreviation, and misinformation remained unchanged after intervention that submits that AI tools may not work well in addressing culture specific areas and other comprehension issues. Present study validates the warning of Marzuki et al. [16] the potential negative impact of over-dependence on AI on students' development of independent critical thinking skills. A hybrid model should be the focus of future academic inquiry, combining AI feedback with teacher-led instruction in argumentation and critical thinking.

The current findings align with and expand prior studies [16,21] by demonstrating how AI tools not only enhance writing mechanics but also support deeper argumentative structures. By applying Toulmin's framework alongside Ellis's error analysis, this study uniquely shows how ChatGPT intervention intersects with both form-focused and meaning-focused instruction, a key concern in EFL pedagogy [37]. Moreover, this study contributes theoretically by situating ChatGPT within sociocultural learning paradigms, illustrating how AI tools can simulate scaffolding traditionally provided by teachers.

4. Conclusion

This study primarily investigated the effects of the AI tool ChatGPT on the argumentative writing abilities of Pakistani secondary school students. The pre-test and the post-test are conducted for data collection and analysis by following the key components of Toulmin (1958) and steps of error analysis by Ellis (2002). The results predict that argumentative writings are positively affected by the utilization of ChatGPT. It is disclosed that the key components such as claim, reason, evidence, and counterclaim are improved and students overcame several types of errors. The results of this study have important implications for education, notably in the advancement of English language pedagogy in environments comparable to Pakistan. This work offers a theoretical contribution by integrating and applying the frameworks of Toulmin and Ellis, demonstrating their combined utility for enhancing student writing skills. It also shows a shift from traditional teaching methods to new technological-based methods of teaching. The effective enhancement of argumentative writing skills by ChatGPT suggests that educators should explore broader integration of AI tools into curricula, encouraging more interactive and student-centered teaching methodologies. It can assist the teachers in preparing their material. Also, the demonstrable advantages of ChatGPT for EFL learners might encourage educational authorities to prioritize investment in technologically advanced teaching materials, thus fostering a transition away from traditional rote learning towards more stimulating and productive learning environments. Furthermore, this research may encourage additional studies investigating the application of artificial intelligence within diverse language learning contexts and educational levels, thereby fostering advancements in educational technology. It is also suggested that the limitation of AI tools in developing critical thinking should be focused in future research. Finally, this study emphasizes the necessity of incorporating modules focused on effective technology integration into teacher training programs to better equip educators with the skills to improve student learning.

CRediT authorship contribution statement

Musarat Yasmin: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Formal analysis, Conceptualization. **Walees Fatima:** Writing – original draft, Data curation, Conceptualization. **Isra Irshad:** Software, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] P. Sharma, B. Dash, AI and VR enabled modern LMS for students with special needs, J. Foreign Lang. Educ. Technol. 8 (1) (2023) 2023–2040.
[2] J. Bryant, F. Child, E. Dorn, S. Hall, New global data reveal education technology's impact on learning. McKinsey & Company blog. (2020) Available at https://www.mckinsey.com/industries/public-and-social-sector/our-insights/new-global-data-reveal-education-technologys-impact-on-learning.
[3] M. Yasmin, Online chemical engineering education during COVID-19 pandemic: lessons learned from Pakistan, Educ. Chem. Eng. 39 (2022) 19–30, https://doi.org/10.1016/j.ece.2022.02.002.
[4] R.F. Kizilcec, To Advance AI Use in education, focus on understanding educators, Int. J. Artif. Intell. Educ. 34 (2023) 12–19, https://doi.org/10.1007/s40593-023-00351-4.
[5] L. Yan, L. Sha, L. Zhao, Y. Li, R. Martinez-Maldonado, G. Chen, X. Li, Y. Jin, D. Gašević, Practical and ethical challenges of large language models in education: a systematic literature review, Br. J. Educ. Technol. 55 (1) (2023) 90–112, https://doi.org/10.1111/bjet.13370.
[6] M. Soori, F.K.G. Jough, R. Dastres, B. Arezoo, AI-Based Decision Support Systems in Industry 4.0, A Review, J. Econ. Technol. (Pre-proof) (2024), https://doi.org/10.1016/j.ject.2024.08.005.
[7] S. Akgun, C. Greenhow, Artificial intelligence in education: Addressing ethical challenges in K-12 settings, AI Ethics 2 (2022) 431–440, https://doi.org/10.1007/s43681-021-00096-7.
[8] C. Greenhow, S. Galvin, D. Brandon, E. Askari, A decade of research on K-12 teaching and teacher learning with social media: insights on the state-of-the-field, Teachers College Record 122 (6) (2020) 1–7.
[9] M.Z. Haider, T.A. Chowdhury, Promoting CLT within a computer-assisted learning environment: a survey of the communicative English course of FLTC, Engl. Lang. Teach. 5 (3) (2012) 74–86, https://doi.org/10.5539/elt.v5n3p74.
[10] R.K. Dwivedi, Seema, Approaches to computer-assisted language learning (CALL): a brief survey, Int. J. Res. Publ. Rev. 3 (8) (2022) 10–14. https://www.researchgate.net/publication/371486668.
[11] F. Jamaldeen, N.A. Razak, H. Mohamed, A theoretical approach to initiate mobile-assisted language learning among school leavers and university students of Sri Lanka, *arXiv preprint*, https://arxiv.org/abs/1606.01365, 2016.
[12] N. Yeasmin, T.A. Chowdhury, Teachers' perception of using technology in a blended learning environment to facilitate collaborative learning in Bangladesh, in: Proceedings of the 19th International Conference of the Asia Association of Computer-Assisted Language Learning (AsiaCALL 2022), Atlantis Press, 2022, pp. 201–210, in: https://www.atlantis-press.com/proceedings/asiacall-22/125972706.
[13] T. Başar, L. Şahin, Technology integration in teaching English as a foreign language: a content analysis study, J. Educ. Technol. Online Learn. 5 (1) (2022) 204–222.
[14] M. Ratnam, B. Sharma, A. Tomer, ChatGPT: Educational Artificial Intelligence. Int. J. Adv. Trends Comput. Sci. Eng. 12 (2) (2023) 84-91. Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse091222023.pdf.
[15] B. Green, Y. Chen. The Principles and Limits of Algorithm-in-the-Loop Decision Making.Proc. ACM Human-Comput. Interac. 3 (CSCW), 2019, pp. 1–24, https://doi.org/10.1145/3359152.

[16] Widiati Marzuki, U. Rusdin, D. Darwin, I. Indrawati, The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective, Cogent Educ. 10 (2) (2023) 2236469.
[17] T.T.H. Nguyen, EFL teachers' perspectives toward the use of ChatGPT in Writing classes: a case study at van lang university, Int. J. Lang. Instr. 2 (3) (2023) 1–47.
[18] S. Sheikh, S.Y. Yayilgan, B. Klimova, M. Pikhart, Assessing the usability of ChatGPT for formal English language learning, Eur. J. Investig. Health Psychol. Educ. 13 (9) (2023) 1937–1965.
[19] Cao, S., & Zhong, L. (2023). Exploring the effectiveness of ChatGPT-based feedback compared with teacher feedback and self-feedback: evidence from Chinese to English translation. *arXiv preprint arXiv:2309.01645*.
[20] H. Smith, Algorithmic bias: should students pay the price? AI Soc. 35 (2020) 1077–1078.
[21] S. Herbold, A. Hautli-Janisz, U. Heuer, Z. Kikteva, A. Trautsch, A large-scale comparison of human-written versus ChatGPT-generated essays, Sci. Rep. 13 (1) (2023) 18617.
[22] M. Yasmin, F. Naseem, N. Abas, Constraints to developing learner autonomy in Pakistan: university lecturers' perspectives, Educ. Res. Policy Pract. 19 (2) (2020) 125–142, https://doi.org/10.1007/s10671-019-09252-7.
[23] M. Yasmin, A. Sohail, Socio-cultural barriers in promoting learner autonomy in pakistani universities: english teachers' beliefs, Cogent Educ. 5 (1) (2018) 1–12, https://doi.org/10.1080/2331186X.2018.1501888.
[24] M. Yasmin, A. Sohail, M. Sarkar, R. Hafeez, Creative methods in transforming education using human resources, Creat. Stud. 10 (2) (2017) 145–158, https://doi.org/10.3846/23450479.2017.1365778.
[25] M. Javed, W.X. Juan, S. Nazli, A study of students' assessment in writing skills of the english language, Int. J. Instr. 6 (2) (2012).
[26] M.M. Albuhairy, J. Al-Garaady, DeepSeek vs. ChatGPT: Comparative Efficacy in Reasoning for Adults' Second Language Acquisition Analysis. (2025) Available at SSRN: https://ssrn.com/abstract=5137935.
[27] M. Aziz, R. Rawian, Modeling higher order thinking skills and metacognitive awareness in English reading comprehension among university learners. In Frontiers in Education, 7, (2022) 991015. Frontiers Media SA.
[28] S.E. Toulmin. *The Uses of Argument*, Cambridge University Press, 1958.
[29] R. Ellis. The Study of Second Language Acquisition, 2nd ed., Oxford University Press, 2002.
[30] S.P. Corder, The significance of learners' errors, Int. Rev. Appl. Linguistics Lang. Teaching 5 (4) (1967) 161–170.
[31] Y. Xiao, Y. Zhi, An exploratory study of EFL learners' use of ChatGPT for language learning tasks: experience and perceptions, Languages 8 (3) (2023) 212–230.
[32] J. Kim, S. Yu, R. Detrick, et al., Exploring students' perspectives on Generative AI-assisted academic writing, Educ. Info. Tech. 30 (2025) 1265–1300, https://doi.org/10.1007/s10639-024-12878-7.
[33] S. Athanassopoulos, P. Manoli, M. Gouvi, K. Lavidas, V. Komis, The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom, Adv. Mob. Learn. Educ. Res. 3 (2) (2023) 818–824, https://doi.org/10.25082/AMLER.2023.02.009.
[34] D. Ferris, Treatment of Error in Second Language Student Writing, Second Edition, University of Michigan Press/ELT, 2011, https://doi.org/10.3998/mpub.2173290.
[35] Y. Liu, D. Liu, Morphological awareness and orthographic awareness link Chinese writing to reading comprehension, Read Writ. 33 (2020) 1701–1720, https://doi.org/10.1007/s11145-019-10009-0.
[36] A. Mahboob. English medium instruction in higher education in Pakistan: Policies, perceptions, problems, and possibilities. English medium instruction in higher education in Asia-Pacific: From policy to pedagogy, 2017, pp. 71–91.
[37] M. Celce-Murcia, Discourse analysis and grammar instruction, Annu. Rev. Appl. Linguist. 11 (1990) 135–151.

Dr Musarat Yasmin earned her MA in Applied Linguistics from University of Reading, UK. She availed a pre-doc scholarship for San Jose State University, USA, and completed her PhD. She is currently working as Chairperson/Associate Professor at the Department of English, University of Gujrat, Pakistan. She has authored more than 35 peer-reviewed international journal papers with an h-index 17. Her research interests include ESP, TESOL, Education and technology and Discourse Analysis. She serves as editor for Hayatian Journal of Linguistics and Literature and as reviewer of Taylor and Francis, Springer, Elsevier, IEEE, Wiley and De Gruyter journals. Orcid.org/0000–0002/0571–7575