

COMPARATIVE ANALYSIS OF ENGLISH LANGUAGE TEACHERS' AND CHAT GPT-GENERATED ASSESSMENT AND FEEDBACK ON ARGUMENTATIVE ESSAYS BY PAKISTANI UNDERGRADUATES

Uswa Aslam

Department Linguistics and Communications (DLC), University of Management and Technology, (UMT), Sialkot, Pakistan

uswaaslam47@gmail.com

Maryam Dilawar

Department Linguistics and Communications (DLC), University of Management and Technology, (UMT), Sialkot, Pakistan

maryamawan300@gmail.com

Dr. Muhammad Nadeem Anwar

Associate Professor of English, Higher Education Department, Punjab, Pakistan

mnadeemanwar@gcslahore.edu.pk

Muhammad Zahid

Associate Professor of English, Higher Education Department, Punjab, Pakistan

zahidmian89@gmail.com

Corresponding Author: *Muhammad Zahid zahidmian89@gmail.com

ABSTRACT

English language teachers' assessment and feedback on writing productions of novices has been a source of evaluation over the years. However, this feedback could be limited in certain ways. Contrastively, the use of Chat GPT as an assessment and feedback tool has also gained a special area of interest. The current study compared the assessment and feedback of teachers and Chat GPT on argumentative essay writings of Pakistani BS English undergraduates. The study utilized qualitative research design. A purposive sampling technique was used to select participants (n=60) from a public-sector graduate college. The tool used in the study was writing productions of argumentative essays. The data were analyzed by comparing traditional teachers' assessments and Chat GPT assessments, and the findings revealed that teachers' assessments and feedback were limited while those of Chat GPT were comprehensive. Moreover, this study concluded that Chat GPT as an assessment and feedback tool would improve students' learning experience. This study opened opportunities for future researchers to conduct studies at a broader level. They may explore the inclusion of Chat GPT in teacher education or the development of personalized strategies for assessment with the use of Chat GPT in Pakistan.

Key Words: Argumentative essay writing, assessment, feedback, Chat GPT, Pakistani undergraduates

INTRODUCTION

Writing is a fundamental form of communication, and the ability to communicate oneself clearly and effectively in personal, professional, and academic settings is essential. As students and practitioners, we must be able to effectively communicate in writing. This is required to meet professional practice obligations such as assessing, planning, implementing, and evaluating care, as well as writing nursing and midwifery notes and supporting and justifying changes in practice to managers and other interdisciplinary colleagues. In addition, one could choose to write for publication to share best practices and new breakthroughs with a larger audience. Remember that there is probably no such thing as a natural writer. The effective writer is one who sees writing as a skill that can be studied, perfected, and

continually improved. Successful writing is about using ideas and facts to express something in a clear, unambiguous manner that others can understand (Lloyd-Jones & Masterton 2010). There are many types of writing. One of them is argumentative. It is probably the most common writing genre that L2 students encounter in colleges. According to Ferretti, Andrews-Weckerly, and Lewis (2007), this type of writing is particularly challenging for L2 learners due to its complex cognitive and linguistic requirement.

Widyastuti (2018) argues that critical thinking (CT) is a measurable skill that can be integrated into language teaching. CT is a difficult topic to define, but it has become an important skill in the 20th century. A writing course that motivates students can be built on the concept that, first and foremost, CT is an essential component of argumentative writing, and when combined with other important language skills, it is more effective than traditional writing methods. Integrating CT into argumentative writing promotes critical thinking and other areas of writing. Writing means arguing, and arguing means thinking. Argumentative writing is a complex skill that requires background knowledge, writing skills, and the ability to analyze facts. It can help students generate and organize ideas, defend opinions, and develop cognitive and social skills, including making comparisons, drawing inferences, and evaluating. Writing is the verbal representation of CT. Critical thinking requires mastery of multiple arguments and problem-solving skills (Chance, 1986).

Therefore, improving writing skills is necessary, but this improvement requires timely assessment and effective feedback without negatively affecting student self-image and motivation. It is considered that feedback always produces favorable results, but in some cultures, it is observed that this is not always the case. Feedback intervenes in one's task performance. Kulger and DeNisi (1996) argued that feedback usually enhances performance, but it actually decreases it in one third of the cases. For instance, in Pakistan we have often observed that a teacher does not have enough time in the classroom to tell students their marks individually because the class strength is usually large. Therefore, they traditionally announce marks in the class, which some students consider embarrassing. Moreover, there is a tendency to use harsh words like *very poor*, *very bad*, *you always perform badly*, *you cannot succeed*, etc., in some cultural settings that may hurt students' self-esteem and demotivate them.

In recent years, the advancement of artificial intelligence (AI) has emerged as a critical factor in redefining educational paradigms, notably in the field of language learning. This surge of AI in education heralds a transition towards more responsive and adaptive learning environments, where language acquisition and instruction are optimized to meet the specific needs of individual students and teachers (Malik et al. 2023). Artificial intelligence (AI) is transforming how things are taught and learned in educational institutions.

Chat GPT is one of the most easily operated and accessible AI tools. Chat GPT plays a very significant role in assessing the feedback. It helps in analyzing sentiments by determining whether it is negative or positive. It also helps in identifying key themes by assessing what the strengths and weaknesses of certain writings are. Moreover, it provides personalized responses and gives suggestions and instructions for improvement that help the students and teachers to encourage and take an active role in learning. It also offers suggestions for improvement and automated feedback analysis that enhance student engagement and support feedback. Chat GPT also prioritizes feedback based on different topics and relevance and also concludes large amounts of data into concise insights. Additionally, it can improve and provide multilingual support and analyze feedback to give refined feedback strategies.

Statement of Problem

It is considered that feedback always has a positive impact on improving students' performance. However, teachers' traditional way of giving feedback is not very helpful. A

number of researches have been done on the use of AI tools for assessment and feedback. (Allen, Likens & McNamara, 2019; Wang, 2024). Similarly, many researchers consider the use of Chat GPT for providing feedback (Cao & Zhong, 2023; 2023; Kerman, Noroozi, et al., 2024; Polakova & Ivenz, 2024). However, they neglect the context of Pakistani students. Therefore, this study will focus on the assessment of argumentative essay writing of Pakistani BS English first-semester students in comparison to traditional teacher feedback.

Research Objectives

- To compare and contrast English language teachers' and Chat GPT-generated assessment and feedback on argumentative essays of Pakistani L2 undergraduates
- To propose recommendations for improving assessment and feedback practices on argumentative essays

Research Questions

1. What are the similarities and differences in the assessment and feedback of English language teachers and Chat GPT on argumentative essays of Pakistani L2 learners?
2. What are the recommendations that can help improve assessment and feedback practices on argumentative writing essays?

RELATED STUDIES

The evaluation of argumentative writing typically includes evaluations of the specific linguistic and rhetorical aspects found in the individual essays written by students. However, the characteristics of the task, learner, and educational situation that determine this adaptability are largely unknown. Allen, Likens, and McNamara (2019) conducted a study which involved editing of six argumentative essays using an automated writing evaluation system, and the researchers obtained both summative and formative comments on their work. The findings support the hypothesis that talented writers display linguistic flexibility throughout the argumentative essays they generate. However, research shows that lower-level feedback (i.e., spelling and grammatical input) has little to no effect on the characteristics of students' essays, nor on their variability among prompts or drafts. Overall, the current study sheds light on the function of flexibility in argumentative writing competence and establishes a solid framework for future research and educational interventions.

Wang (2024) compared human teacher feedback with an AI-driven program named Poe to investigate how corrective feedback affected language learners' writing anxiety, complexity, fluency, and accuracy. Three intact classes of twenty-five language learners each participated in the quasi-experimental design; one class received feedback from the teacher, one class received feedback from Poe, and one class received no feedback at all. Both teacher and AI-generated feedback had a substantial impact on writing anxiety, accuracy, and fluency, according to the results. The AI-generated feedback group performed better than the others and saw a larger decrease in writing anxiety. The study emphasizes the advantages of incorporating AI applications into language learning settings and stresses how crucial it is to take learners' psychological health into account when designing language learning interventions.

Sharadgah, Sa'di, and Ahmad (2020) conducted a study that is heightened by the fact that critical thinking (CT) is widely regarded as a major goal of higher education and the foundation for developing learning outcomes. Thus, this quasi-experiment seeks to promote and measure students' critical thinking skills (CTSs) through argumentative essay writing. The study included 98 male English major students enrolled in an essay-writing course at Prince Sattam bin Abdulaziz University (PSAU), Saudi Arabia. The participants were randomly assigned to either the intervention (n = 49) or control (n = 49) groups. Quantitative and qualitative methodologies were used. Both groups were subjected to pre- and post-tests. The Facione and Facione (1994) CT score rubric was used to assess CTSs. The findings

demonstrated a substantial positive correlation between CT and essay writing ability. Assessment of students' essays revealed that the intervention group considerably outperformed the control group in the five CTSs: "interpretation, analysis, evaluation, inference, and explanation."

Polakova and Ivenz's (2024) study looks into the efficiency of Chat GPT feedback in improving the writing skills of Gen Z pupils, who make up a sizable proportion of today's global population. The study takes a mixed-methods approach, investigating both quantitative and qualitative features of students' interactions with Chat GPT in the context of foreign language instruction. A quasi-experimental design was used, with 110 university students learning English as a foreign language (EFL). Pre- and post-tests, questionnaires, and focus group interviews were used to assess the influence of Chat GPT feedback on students' writing proficiency. Following Chat GPT comments, quantitative analysis demonstrated considerable improvements in different aspects of writing, such as conciseness, grammar, inclusion of crucial information, and usage of passive voice. Qualitative insights from focus group interviews confirmed similar findings among students. Overall, the study shows that ChatGPT has the potential to be an innovative instructional tool for developing writing skills among Generation Z students. The findings highlight the necessity of incorporating AI-driven technologies into language learning processes to satisfy the changing needs of today's learners.

Peer feedback is an effective learning approach, particularly in big courses with high instructor workloads. Peers may struggle to provide high-quality criticism for complicated tasks like writing an argumentative essay, as they demand advanced cognitive processing, critical thinking, and topic knowledge. The debut of Chat GPT has sparked debate over whether AI technologies can provide feedback for hard activities on this topic, making it unclear. Banihashem et al. (2024) researched using Chat GPT to provide feedback on students' argumentative essay writing assignments and compared the quality of Chat GPT-generated comments to peer feedback. Participants included 74 graduate students from a Dutch university. The study was conducted in two phases: first, students' essay data were acquired as they wrote essays on one of the assigned subjects; second, peer feedback and Chat GPT-generated feedback data were obtained by engaging peers in a feedback process and using Chat GPT as a feedback source. Two coding schemes, one for essay analysis and another for feedback analysis, were employed to measure the quality of feedback. Then, a MANOVA analysis was used to assess whether there were any differences between peer feedback and Chat GPT. Spearman's correlation was used to investigate the relationship between essay quality, peer feedback, and Chat GPT. The results revealed a considerable difference in feedback generated by Chat GPT versus peers. While Chat GPT gave more descriptive comments, including information about how the essay was written, peers supplied feedback that included information on the essay's problem identification.

Cao and Zhong (2023) state Chat GPT, a cutting-edge, AI-powered chat bot, can respond swiftly to provided commands. While Chat GPT has been shown to be capable of providing helpful feedback, its usefulness in comparison to traditional feedback systems such as teacher feedback (TF) and self-feedback (SF) remains unknown. To address this issue, this study compared revised Chinese to English translation texts produced by Chinese Master of Translation and Interpretation (MTI) students who learned English as a Second/Foreign Language (ESL/EFL) using three types of feedback (Chat GPT-based feedback, TF, and SF). The data were examined using BLEU score to gauge the overall translation quality and Coh-Metrix to examine linguistic aspects across three dimensions: lexicon, syntax, and coherence. The results showed that TF- and SF-guided translation texts outperformed those with Chat GPT-based input, as demonstrated by their BLEU score. In terms of linguistic aspects, Chat

GPT-based feedback outperformed, particularly in improving lexical competence and referential consistency in translation texts. However, TF and SF were more helpful in teaching syntax-related skills since they addressed inappropriate use of the passive voice. These different results demonstrate Chat GPT's potential as a supplementary resource to standard teacher-led techniques in translation practice.

“Feedback intervention is defined as actions taken by an external agent to provide information regarding some aspect of one's task performance.” Kluger and DeNisi's (1996) work fundamentally altered the field of feedback study. Until this publication (and even now), many people felt that feedback always produced favorable results. There are numerous reasons for this, including the fact that much of the previous feedback research assumed that feedback givers always provided both excellent and effective feedback. After doing a thorough examination, the feedback intervention article discovered that, while feedback enhanced performance on average, it actually decreased performance in one-third of situations. Feedback can have a profound impact, but its effectiveness is contingent upon the manner of delivery, as poorly conveyed feedback can precipitate defensiveness, diminished motivation, and a penchant for self-protection, whereas timely, specific, balanced, and actionable feedback, delivered with empathy, can catalyze growth, improvement, and positive transformation, underscoring the imperative of a thoughtful and nuanced approach. So, students should be aware of the use of technology to improve their skills. This highlights that technology should be introduced and accepted by institutes to promote high-level education.

The technology acceptance model (TAM) serves this purpose. It is an information systems theory that describes how to persuade consumers to accept and use new technology (Davis, 1989). It has been frequently used by information systems scholars to solve the organizational challenge of increasing adoption of new information systems (Liu, Dedehayir, and Katzy, 2015). TAM identifies two essential aspects influencing user acceptance: perceived utility and perceived ease of use (Davis, 1989). The underlying idea is that the more users believe a given program would improve their performance, and the less effort required to utilize the application, the higher the adoption rate. TAM is used to predict technology adoption rates, assess user behavior, and guide technology design in order to produce user-friendly and appealing solutions. TAM also assesses the success of technology implementations, compares various technologies, and drives the creation of training and support programs. TAM's simplicity, flexibility, and empirical support have made it a popular choice for researchers and practitioners looking to understand and promote technology adoption, resulting in more informed decision-making and better technical results.

In conclusion, a good number of studies have been done on the use of Chat GPT as a feedback provider. Nevertheless, the use of Chat GPT in certain cultural settings where traditional teacher feedback is limited and time-taking has not been studied much. Moreover, these studies neglected Pakistani students. Therefore, the current study mainly focuses on the assessment of Pakistani students through Chat GPT in comparison with teacher feedback and highlights the ease and usefulness of Chat GPT for improvement. In addition, feedback intervention and TAM provide a theoretical framework for the recommendation of using Chat GPT for giving effective and instant feedback.

METHODOLOGY

This research utilized a qualitative research design. The purposive sampling technique was used to select samples (n=60) from BS English first-semester students from the public sector institute: Govt. Graduate College, Hajipura, Sialkot, Punjab, Pakistan. The participants of the study were between 18 to 25 years of age. The tool used in the research was argumentative essays. The researchers personally visited the college. Taking ethics into account, the researchers asked for permission from the head as well as the class teacher, and the following

surprise test activity was conducted. Write an argumentative essay on the topic of "Co-education: Beneficial or Harmful."

The researchers collected the writing manuscripts of the samples. Subsequently, the collected data were evenly distributed between two different English language teachers of the same public sector institute. The teachers assessed the essays, graded them traditionally and gave their feedback on them. The researchers did not ask the students to submit it online because the researchers wanted to make sure of the authenticity of the material and the knowledge of students. Those manually assessed essays were collected from teachers one week later and were typed by the researchers. After that, a prompt was given to Chat GPT along with typed essays of students one by one. The prompt given to the Chat GPT to assess the essays and get feedback was the following:

Prompt: "Assess the following essay and grade it out of 10; also provide possible suggestions for improvement."

Chat GPT assessed the essays and provided feedback. In the next step, both feedbacks were compared, and content analysis was done to check the effective methods of feedback and recommend Chat GPT for improvement of argumentative essay writing skills.

DATA ANALYSIS

The assessment provided by teachers and Chat GPT was thoroughly examined and categorized into emerging themes. These themes were then analyzed to identify patterns of similarities and differences in the assessments of Chat GPT and teachers. Identified themes and the analysis have been provided below:

Content and Argumentation

Teachers analyzed the content in a general way without highlighting the strengths or weaknesses of the essays. They did not highlight the quality of the content. They did not provide any details regarding arguments represented in the essays. They did not even write general statements like content is poor, weak arguments, some ideas are vague, justify your arguments, some ideas are vague, etc. Adding to it, they did not give any specific suggestions to improve content and argumentative skills.

On the other hand, Chat GPT deeply analyzed the content of each essay and provided strengths of the essay, like the essay *acknowledges societal stereotypes, it emphasizes the importance of education for girls, confidence issues are discussed, distractions caused by co-education, and personal reasons like some parents allow some don't for opposing co-education*. It also highlighted weaknesses of each essay, e.g., *content is repetitive, vague claims like "our religion is not allowed co-education, lacks depth, arguments are not developed in a good way, generalizability of claims like "students are not concentrating, and ideas are underdeveloped*. Moreover, it also gave suggestions like how the content could be written in a more comprehensive way. For example, a few suggestions from the Chat GPT assessment of different essays are as follows: *this can be improved by adding more detailed arguments and real-world examples, avoiding repetition and vague ideas, expanding your arguments like how co-education benefits students and addresses both genders in an equal manner, justifying your arguments, and saying the co-education system has more accidents. Give evidence*. This could be an encouraging step for students that is not catered to by teachers.

Clarity and Organization

Analysis of this theme in teachers' assessments highlighted that their assessment was limited. They just wrote on the essay that "There is no clarity" but did not mention where the clarity is required or in which specific paragraph they were lacking this point. Moreover, not all teachers commented about the clarity and organization of the essay.

Whereas Chat GPT assessed and provided feedback related to the theme of clarity and organization in a more comprehensive way. It stated that ideas were repetitive and lacked logical flow, e.g., *people thinking in this style, she is enough for her* is vague and unclear. It also stated that the learners mixed the unrelated ideas in one sentence, e.g., *It gives them the belief that she is also sitting with boys, and she understands the mentality of boys*. According to Chat GPT-generated feedback ideas were disorganized and lacked clear transition. For instance, *I am not good at co-education*. Even some sentences were contradictory. For example, *I am not good at co-education, not to say bad*. Furthermore, Chat GPT talked about the organization of the essays specifically, like the essays were not thoroughly organized the way in which they had to be organized. It mentioned that *the essay lacks introductory lines, a topic sentence, a thesis sentence, an outline, and a proper ending*. Also gave suggestions like *using separate paragraphs for introduction, advantages, disadvantages, and conclusion*.

Tone and Language

Teachers provided feedback by using statements like *very poor, you cannot succeed in life, and you always perform like this*. They used such words that could hurt them and demotivate them. And perhaps due to this reason, most of the students don't take interest in the class, which eventually become the reason of their failure.

In contrast, Chat GPT provided feedback, which was free from harsh remarks and highlighted students' mistakes in a friendly manner that might motivate them to be the better version of themselves. It avoided judgmental and overly negative language. It guided them neutrally. For example, you could write *some individuals believe co-education leads to distractions* instead of *people's thoughts are not clean*, and it also suggested that students should have avoided biased and emotional statements such as *our society and religion, it is not liked by the institutions*; rather, it should be presented more objectively.

Grammar and Syntax

Teachers highlighted this theme only by writing on the test that they should *improve grammar, poor sentence structure, and a lot of grammatical errors* but did not specify the sentences and paragraphs in which they made such errors. Even teachers didn't correct the errors.

On the other hand, Chat GPT presented a detailed version of this theme by showing frequent grammar issues that might make the essay difficult to understand. For example, *girls and boys together studying should be like this: girls and boys study together*. Chat GPT not only mentioned the grammatical and syntactical errors but also provided their corrections. For correcting the subject-verb agreement error it suggested that the writer should have written *education has a greater impact on students* instead of *Education has more impact on students*; *co-education helps students not (helps)*. It also suggested that the writer should have used correct possessives and apostrophes: e.g., *girls, not girl's; boys, not boy's*. It also suggested that the writer should have avoided grammatical errors. Instead of writing *some students do not get admitted to co-educate universities* they should have written *some students do not enroll in co-educational universities*.

Quality

Analysis of teachers' assessment style highlighted that teachers provided low-quality feedback. It did not help much in improving students writing because there was no explanation and correction of mistakes in the assessment of the essay. It consisted of only grades with a few remarks written on it.

On the other hand, Chat GPT's perspective on the quality of essays was highlighted in a better way. For instance, this essay exhibited a satisfactory comprehension of the issue of co-education, but it suffered from a number of constraints. The language utilized was simple, and the vocabulary was fairly limited, which reduced the essay's clarity and persuasive

power. Furthermore, the essay's structure and organization were not ideal, with abrupt paragraph transitions and a lack of clear linkages between topics.

Specificity and Generality

The teachers provided assessment in a very limited way. Their feedback was generalized. They marked red circles or underlined the test overall but did not mention what were grammar mistakes, where they lacked in syntax, where they used poor sentence structure, what were grammatical errors, what would be the correct form of verb, or the overall sentence.

On the other hand, Chat GPT provided comprehensive and detailed feedback on the basis of content. It provided an explanation along with strengths and weaknesses. It highlighted spelling errors. *e.g. (experience), their (their), managed (managed)*. It also pointed out grammatical mistakes and syntactical errors with corrections such as *many students are highly confident, but many are very low confidence. It should have been like this: Some students are highly confident, while others have very low confidence*. The essay contained many poor, fragmented, and unclear sentences, such as, *because some girls and boys do not explain what they thought*.

Improvement Suggestions

Teachers did not suggest any improvement in the essay except marking a few correct spellings. Teacher feedback was overall and generalized. They marked the test but did not specify where students had made mistakes. Like they simply mentioned at the top of the test *poor sentence structure, a lot of spelling mistakes, grammatical errors, vague ideas, and weak arguments*.

On the other hand, Chat GPT provided a comprehensive feedback, like essays lacked clarity and organization, ideas were repetitive, vague, and unclear. It held this sentence *People thinking in this style* as an unclear structure. It also pointed out that essays had lacked organization. It stated that the introductory paragraph is faulty. It should have explained what co-education is, what are the arguments related to this topic, and then give concluding lines.

DISCUSSION

After the distribution of tests between two teachers for assessment and feedback, they return the tests. The marked scripts did not manifest quality feedback. The feedback only consisted of grades and a few words or sentences as a remark. For example, spelling mistakes, a lot of grammatical errors, poor sentence structure, sentences that are not properly constructed, and there were a few suggestions like correct vocabulary suggestions, weak arguments, vague ideas, and a lot of red circles on the test. In comparison, Chat GPT gave more descriptive feedback. It gave a detailed explanation of mistakes related to grammar, structure, and cohesion. Furthermore, Chat GPT graded essays in different sections like clarity and cohesion, grammar and language use, content and depth, and organization. It also provided suggestions for further improvement, like improving grammar and language, organizing the essay, expanding on ideas, and focusing on clarity. It was clear through the analysis of Chat GPT feedback that it provided feedback solely on the input text without any biases and emotional factors, as seen in the teachers' feedback that they use phrases in the classroom like very poor, you always perform badly, and you have topped from bottom. In addition, Chat GPT provides consistent feedback, whereas human feedback varies depending on teachers' moods and teaching styles.

Moreover, large class sizes contribute significantly to delayed feedback in educational contexts. When teachers are responsible for managing a large number of students, grading and delivering individualized comments become more time-consuming. With a heavy workload and limited time, teachers may struggle to keep up with grading, resulting in delayed feedback. This can result in students obtaining feedback that is too late to be useful,

limiting their capacity to change their learning practices and enhance their knowledge of the content. Furthermore, large class sizes might reduce the quality of feedback since professors are obliged to deliver brief, general comments rather than thorough, personalized instruction. Moreover, various preparations can result in delayed feedback. Teaching numerous subjects or courses might lead to delayed feedback since it increases the teacher's workload and divides their focus. When instructors must prepare lectures, grade assignments, and provide comments for many classes, it can be difficult to organize their time properly. The extra complexity of managing numerous curricula, tests, and student needs might result in a backlog of grading and feedback, causing delays in providing students with appropriate assistance. As a result, students may receive feedback too late to be useful, limiting their capacity to modify their learning practices and enhance their comprehension of the topic. Due to the need to deliver marks and reports on time, teachers may speed through the grading process, reducing the quality and detail of their feedback. (Cao & Zhong, 2023; Gul et al., 2016)

Furthermore, the time and effort required to satisfy administrative requirements can limit the time available for offering thoughtful and constructive feedback to students. As a result, students may receive brief, superficial comments rather than thorough, actionable feedback that can aid in their learning. All these factors highlight that chat GPT provides thorough and detailed assessments and feedback as compared to teachers' assessment and feedback. Chat GPT provides more concerns and opportunities for effective language learning (Kim et al., 2023). This shows that both have some slight similarities, but many factors that contribute to the limited, low-quality, and delayed assessment offered by teachers create much more difference than the chat GPT. It answers the first research question. These factors eventually impede student advancement and weaken the efficacy of the feedback process. Keeping this in view The use of Chat GPT should be encouraged at a personal and academic level.

CONCLUSION

Analysis shows that including Chat GPT assessment into English argumentative essays is a pivotal step towards improving students' learning experiences in Pakistan. Using Chat GPT fast evaluation, real-time comments, and language support capabilities, educators may deliver timely, tailored, and effective feedback to students. This not only increases the quality of student writing but also promotes a more autonomous and self-directed learning environment. Furthermore, Chat GPT assessments can help Pakistani students overcome problems such as restricted access to quality feedback and linguistic barriers. Chat GPT can help overcome the linguistic and academic gap by offering multilingual support and automating the grading process. However, it is critical to recognize the limitations and inherent biases of Chat GPT assessments. Educators must ensure that Chat GPT technologies are built and educated to handle the complexities of Pakistani English and cultural contexts. Finally, the strategic use of Chat GPT assessment in English argumentative essays has the potential to transform how Pakistani students learn and engage with the topic. Educators may use Chat GPT to build a more inclusive, efficient, and effective learning environment that equips students to succeed in an increasingly complicated and interconnected world. The current study may open opportunities for future researchers to conduct studies at a broader level, including various perspectives. They may explore the inclusion of Chat GPT in teacher education or the development of personalized strategies for assessment with the use of Chat GPT to enhance writing skills and promote learning outcomes.

REFERENCES

- A Sharadgah, T., A Sa'di, R., & H Ahmad, H. (2020). Promoting and assessing EFL college students' critical thinking skills through argumentative essay writing. *Arab World English Journal (AWEJ) Volume, 10*.

- Al-Amin, M., Saqui, F. Z., & Khan, M. R. (2024). Enhancing Assessment Systems in Higher Education: A Review on Artificial Intelligence Usage. *Utilizing AI for Assessment, Grading, and Feedback in Higher Education*, 28-56.
- Allen, L. K., Likens, A. D., & McNamara, D. S. (2019). Writing flexibility in argumentative essays: A multidimensional analysis. *Reading and Writing*, 32, 1607-1634.
- Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education*, 21(1), 23.
- Bellhäuser, H., Dignath, C., & Theobald, M. (2023). Daily automated feedback enhances self-regulated learning: a longitudinal randomized field experiment. *Frontiers in Psychology*, 14, 1125873.
- Cao, S., & Zhong, L. (2023). Exploring the effectiveness of Chat GPT-based feedback compared with teacher feedback and self-feedback: *Evidence from Chinese to English translation*. *arXiv preprint arXiv:2309.01645*.
- Chance, P. (1986). Teaching Thinking. *Curriculum Report*, 15(5), n5.
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1), 57.
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of Chat GPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474. <https://doi.org/10.1080/14703297.2023.2195846>
- Ferretti, R. P., Andrews-Weckerly, S., & Lewis, W. E. (2007). Improving the argumentative writing of students with learning disabilities: Descriptive and normative considerations. *Reading & Writing Quarterly*, 23(3), 267-285.
- Fitria, T. N. (2021a). Artificial Intelligence (AI) In Education: Using AI Tools for Teaching and Learning Process. *Prosiding Seminar Nasional & Call for Paper STIE AAS*, 4(1), 134–147. <https://prosiding.stie-aas.ac.id/index.php/prosenas/article/view/106>
- Gul, R. B., Tharani, A., Lakhani, A., Rizvi, N. F., & Ali, S. K. (2016). Teachers' perceptions and practices of written feedback in higher education. *World Journal of Education*, 6(3), 10.
- Kim, S., Shim, J., & Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System*, 10(1), 79-88.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, 119(2), 254.
- Lloyd-Jones, N., & Masterton, A. (2010). *Writing skills and developing an argument*. In *Bailliere's study skills for nurses and midwives*. Edinburgh Bailliere Tindall.
- Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., & Darwis, A. (2023). Exploring artificial intelligence in academic essay: higher education student's perspective. *International Journal of Educational Research Open*, 5, 100296.
- Mustafa, G., Urooj, T., & Aslam, M. (2024). Role of artificial intelligence for adaptive learning

- environments in higher education by 2030. *Journal of Social Research Development*, 5(3), 12-22.
- Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), 2410101.
- Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), 2410101.
- Tajik, E. (2024). A comprehensive Examination of the potential application of Chat GPT in Higher Education Institutions.
- Valero Haro, A., Noroozi, O., Biemans, H. J., Mulder, M., & Banihashem, S. K. (2024). How does the type of online peer feedback influence feedback quality, argumentative essay writing quality, and domain-specific learning? *Interactive Learning Environments*, 32(9), 5459-5478.
- Wang, D. (2024). Teacher-versus AI-generated (Poe application) corrective feedback and language learners' writing anxiety, complexity, fluency, and accuracy. *International Review of Research in Open and Distributed Learning*, 25(3), 37-56.
- Widyastuti, S. (2018). *Fostering critical thinking skills through argumentative writing*. Yogyakarta State University.
- Zhang, Z. V., & Hyland, K. (2022). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, 51, 100586.