

AN EMPIRICAL INVESTIGATION OF THE USE OF AI CHATBOTS AS PREDICTORS OF ENGLISH AS A SECOND LANGUAGE PERFORMANCE IN AZAD JAMMU AND KASHMIR

Summera Sarwer¹, Rukhsana Ameer², Muhammad Faisal Ishaque^{*3}

³Lecturer, Department of Education University of Poonch Rawalakot Azad Kashmir

¹summera.sarwer@gmail.com, ²rukhsana.usmanakb@gmail.com, ³faisalishaque7@gmail.com

Corresponding Author: *

Muhammad Faisal Ishaque

DOI: <http://doi.org/10.5281/zenodo.19591728>

Received
15 March 2025

Accepted
19 April 2025

Published
30 April 2025

ABSTRACT

The research was conducted to determine how AI chatbot-supported instruction can improve the English as a Second Language (ESL) proficiency of Class 10 students of Government Pilot High School Bagh, Azad Jammu and Kashmir. It was a quasi-experimental non-equivalent control group design that included 60 students in an experimental group and control group. The experimental group was supplied with AI chatbot-mediated instruction with the aid of ChatGPT and Google Gemini, and the control group was instructed in the traditional English. The intervention took place during the eight to twelve weeks period and was in accordance to Class 10 English syllabus by Azad Jammu and Kashmir Textbook Board. The data were gathered through the use of a curriculum-based ESL proficiency test and speaking and writing assessment through the use of analytic rubrics. The data was analyzed by descriptive statistics, independent samples *t*-tests, and Analysis of Covariance (ANCOVA). The results showed no significant difference of the groups during the baseline. Nonetheless, the post-intervention outcomes showed significant statistical gain in the overall ESL proficiency of students having AI chatbot-based instruction with a big effect size.

There was also a great improvement in speaking, writing, vocabulary, and grammar skills. It was concluded that AI chatbot integration in a structured manner could be of great help in improving ESL learning at the secondary level. The results imply that AI-mediated instructional strategies may be used as useful pedagogical strategies to enhance communicative competence in any English language classroom within Azad Jammu and Kashmir.

Keywords: AI chatbot-assisted instruction, English as a Second Language, ESL proficiency, quasi-experimental design, secondary education, ChatGPT, Google Gemini, Azad Jammu and Kashmir.

INTRODUCTION

The artificial intelligence (AI) chatbots are quickly exiting the customer service sector and entering the mainstream educational sector, particularly where the large language models have more natural and interactive conversation. Chatbots have been introduced in language learning as the so-called always-available partners, which may remind learners, reply to their

messages, and give them examples as soon as they need them. Systematic review of AI in language education indicate increased adoption of conversational agents and writing/Speech support software, with much focus on real-time feedback and less anxiety of learners practicing a second language. Such trends indicate that the variety of computer-assisted tasks will take place less often and be replaced by dialogue-focused

practice and embedded in the daily learning activities (Albedah, 2025).

In English as a Foreign Language (EFL) setting, the area of research focused on speaking practice has gained more and more popularity due to the fact that tradition classrooms generally offer little time to speaking solo. A systematic review of AI-based chatbots as a means of EFL speaking practice offered evidence that the activities supported with chatbots can be useful in achieving learning outcomes, including confidence, engagement, motivation, and pronunciation practice, and reduced speaking anxiety, but also the field remains a developing area and requires stronger and context-sensitive research designs. This body of international research contextualizes chatbots as potential useful supplements and not substitutes to teachers, particularly in situations with limited practice opportunities (Du and Daniel, 2024).

Simultaneously, global policy and research community emphasize the benefits of education to be based on safe and pedagogical incorporation. According to the UNESCO instructions, the risks in chatbot use in the educational process can be listed as the potential issues of erroneous results and prejudicial stereotyping of information, paired with the necessity of questionable inclusion by the learner and instructor, and regard AI-generated information unquestioningly. Such warnings are central to any evaluation since language learners can perceive errors that are in fluent responses as the normative ones so that the learning and assessment practices can be distorted (UNESCO, 2023).

In education, the evaluation of AI chatbots worldwide is increasingly being perceived through the prism of effectiveness, equity and governance. Advantages like instant feedback and alleviated anxiety are highlighted by reviews, whereas deficiencies in longitudinal evidence and atypical access to infrastructure being uneven and increasing inequalities have been highlighted. In English learning outcomes, rigorous assessment is usually specified to provide answers to what, on whom, and under which conditions in the classroom and connectivity are some outcomes of

English learning maintained (Albedah, 2025; UNESCO, 2023).

The improvement of the outcomes of the study of English language in Azad Jammu and Kashmir (AJK) is directly correlated with the increased education enrolment rates, quality, and resource realities. According to the official statistics, the literacy rate in AJK (age 10 +) in 2017 was 74.79 percent, and the presence of considerable gender discrepancy proves the same official table. Those figures testify to a rather good baseline in the area of literacy but, also, reveals that literacy support, in both instances, the girls, as well as students in isolated learning conditions is still a key issue to make subject to any learning technology new (Bureau of Statistics, P&DD, Government of AJK, 2023).

This means the school system of AJK is functioning on a large scale which further justifies the necessity to conduct a systematic assessment of the system and only after it is conducted, start using Linux-style chatbot-based learning of English in large scale. The official portal of Elementary and Secondary Education Department represents that more than six thousand schools exist and the numbers of registered students are hundreds of thousands, and the number of teachers tens of thousands. In such system, AI chatbots may be viewed as interesting in terms of extending the practice beyond the classroom, however, the quality of it will be conditioned by the correspondence with the goals of the curriculum, the desire of the teachers, and the fact of access to the devices and connectivity within the districts (Elementary & Secondary Education Department, Government of AJK, n.d.).

The policy course is also pertinent because it predetermines not only what the notion of better results should consist of but also the manner in which the education technology is being controlled in the schools. The National Education Policy (2017 2025) of Pakistan not only states the importance of instructional technologies and the introduction of ICT to facilitate student learning, it also acknowledges the challenge of the meaningful integration not only as the provision of hardware. Regarding the case of AJK, one should therefore assess AI

chatbots not only by the attributes of gained benefits in the English language among learners, but also the possibility to implement, professionally develop, and monitor the quality and responsible use of AI chatbots (Ministry of Federal Education and Professional Training, 2017-2025).

In AJK, recent activities have depicted proactive policy making with the support of the UNICEF through announcement of an Education Policy 2024-2040. This scenario supports the thesis in explaining why AI chatbots can be viewed as part of a more extensive modernization agenda, in which equity and protection of children are and can be central to it. The education materials created by UNICEF to Pakistan are also dedicated to the continuous work concerning the access, learning outcomes, equity, and support that subsequently assesses the technical interventions as possibly shrinking or expanding the existing gaps (Dawn, 2024; UNICEF, n.d.).

That being the case, the research problem of the present study, which is how AI chatbots can be used to improve the performance of English language learning of AJK, is timely and valid. It sets global evidence touching chatbot-assisted language learning and worldwide precautions of dependability, biases, and critical use - in the documented education level of AJK and literacy profile. The plausible assessment in AJK should examine quantifiable English outcomes (speaking, writing, vocabulary, understanding), it should also examine student involvement and anxiety, facilitation activities of teachers, and the appropriateness and practicality of the introduction of using AI tools in real-life classes and homes (Du and Daniel, 2024; UNESCO, 2023; Government of AJK, Bureau of Statistics, P&DD, 2023).

Problem Statement

Although AI chatbots have become a point of interest in English as a Second Language (ESL) education, the connection between the use of chatbots and the quantifiable ESL learning effectiveness is not outlined in an empirical clarity. The major independent factor here is the incorporation and a systematic application of AI chatbots to practice English language, whereas

the most vital dependent variables consist of speaking fluency and pronunciation accuracy of students, vocabulary acquisition, grammatical competence, writing quality, and communicative competence in general. The current literature tends to concentrate on the engagements and satisfaction of the learners, instead of subjectively tested proficiency gains, launching a knowledge gap as to whether one of the potential benefits of the chatbots interaction is the long term sustaining effect of ESL performance.

The contextual variables that contribute to the problem in the Azad Jammu and Kashmir (AJK) include a lack of exposure to the English language outside the classroom, teaching and learning approaches that focus on teachers, class size, and insufficient availability of digital resources. The impact of such moderating and control variables on the ESL learning effectiveness of AI chatbots intervention can be overwhelming. Thus, the underlying research gap is the lack of a variable contextually relevant study, which studies what the use of AI chatbots can do or not do in teaching ESL proficiency, and also takes into consideration institutional, pedagogical and technological conditions in the school setting of AJK.

Purpose of the Study

This research aims to determine the influence of the introduction of AI chatbots on the English as a Second Language (ESL) learning outcomes among Azad Jammu and Kashmir students. Particularly, the proposed study is expected to analyze the hypothesis on whether structured AI chatbots use can boost measurable ESL proficiency as assessed by speaking fluency, pronunciation accuracy, vocabulary acquisition, grammatical competence, and writing performance. The aim of the study is also to establish the correlation between the use of AI chatbots (independent variable) and ESL academic performance (dependent variables) and also takes into account the contextual characteristics of instructional practices, access to online resources, and exposure of English to learners. Through the production of empirical data fitting into the local educational context, the study aims to assess the potential of AI chatbot-

mediated learning being used as a viable supplement tool to overcome the long-term ESL learning issues in Azad Jammu and Kashmir.

Objectives

- To identify the impact of instruction guided by artificial intelligence (AI) chatbots on English as a Second Language (ESL) proficiency in students in Azad Jammu and Kashmir.
- To compare the ESL learning performance of students who experienced the programs based on the AI chatbots with that of the students that were taught the subject through traditional means.

Research Hypothesis

H01: English as a Second Language (ESL) proficiency is a significant difference and does not exist between students getting AI chatbot-assisted instruction and traditional instruction.

H02: English as a Second Language (ESL) proficiency does not differ significantly between students who are given AI chatbots-assisted and traditional instructions.

Significance of the Study

The presented research is contingent because it indicates empirical data about the utility of AI chatbot-mediated teaching to enhance the level of English as a Second Language (ESL) proficiency when the population is the students of Azad Jammu and Kashmir. The study provides information to a paucity of context-bound literature that investigates technology-mediated ESL education under-resourced conditions by providing a systematic comparison of AI-assisted learning and traditional instructions. The results will also contribute to the understanding of whether AI chatbots can be meaningful to adopting the unresolved difficulties in speaking, vocabulary growth, grammar, and writing abilities.

The research also holds significance to teachers and curriculum developers since it provides empirical information on how to introduce AI tools into pre-established ESL teaching processes. Educators can be interested in knowing that chatbot-mediated learning can improve student engagement levels, communicative skills, and self-

practice outside the classroom restrictions. Since the findings are applicable to their professional development, professional development initiatives should be based on efficient and pedagogically well-grounded technology integration.

As a research on evidence-based policy and practices in language education, this study can serve policy makers and education administrators in Azad Jammu and Kashmir to make informed decisions to invest in digital learning and innovate in language education. In case AI chatbot-assisted instruction proves any measurable positive change in ESL achievement, it can help in the creation of scalable, cost-effective mechanisms to improve the English command. Finally, the study works towards the enrichment of the academic performance, employment opportunities, and the availability of higher education opportunities that require sufficient functional English language skills.

Delimitations of the Study

- The research was restricted to Class 10 learners of Government Pilot High School Bagh, Azad Jammu and Kashmir, with the involvement of only two sections (Computer Science and Biology) only.
- The intervention was limited to the Class 10 English curriculum as recommended by the Azad Jammu and Kashmir Textbook Board and was done within an eight to twelve-week teaching term.
- The study devoted attention to only some of the aspects of ESL proficiency, i.e. speaking, writing, vocabulary, and grammar and did not independently match the skill of listening and the acquisition in the long run.
- In the exclusive use of two designated chatbots, ChatGPT and Google Gemini, which help with verbal (conversational practice) and non-verbal communication, the intervention was based on AI.

REVIEW LITERATURE

Chatbots involving artificial intelligence (AI) have become increasingly popular in the English as a Second Language (ESL) education as they can deliver an interactive exercise, quick feedback,

and personalized support that cannot be achieved within the constraints of classroom learning. This is in line with the principles of second language acquisition (SLA) that have highlighted the role of interaction and output in language acquisition, especially on enhancement of fluency, accuracy, and communicative competence (Long, 1996; Swain, 1995). The limited chances of speaking, limited feedback, and anxiety in interactions are harming learning issues experienced by students in most ESL environments, which limits their capability to gain functional levels of proficiency and, therefore, chatbot-based practice is an appropriate instructional aid. Conversational agents of language learning were already studied before the advent of large language models, though research had indicated imbalanced value in teaching since previous systems would generate limited interactions that were unnaturally guided. One major initial conclusion was that chatbots could be used to learn languages, but more pedagogical alignment, better design of conversations, and evidence-based integration into the classroom than new uses of novelty (Fryer et al., 2020). Such an issue is still topical due to a lack of differentiation between the efficiency of working with chatbots to teach ESL and the organization of tasks, mediation of the teacher, and the connection between learner communication and specific, measurable goals in language acquisition.

Empirical research and reviews of chatbots in language learning continue to increase in the period between 2020 and 2022. One large-scale systematic review study found that chatbot-mediated language learning often increased the engagement, motivation, and desire to practice but found less evidence that chatbot-based interventions led to a systematic improvement in the outcome of language proficiency as a result of the methodological constraints, the relatively short duration of intervention, and the lack of continued testing measures (Huang et al., 2022). The results are significant in the context of the study in quasi-experimental research since they indicate a discrepancy between the attitude of learners positively and the proved enhancement of core ESL outcomes including speaking

fluency, grammatical competence, vocabulary learning, and writing skills. The operational scope of generative artificial intelligence (GenAI) has grown rapidly since late 2022, and currently, the research literature on the topic is highly varied with DXplain, ChatGPT, and AI-based chatbots as the most common forms of learning. The UNESCO guidance added that GenAI could be used to help in learning due to personalization, feedback, still, there were risks like misinformation and bias, privacy and over-reliance, which can easily destroy authentic learning (UNESCO, 2023). In case of ESL learners, such risks are particularly serious, as, due to the fluent chatbot responses, the learners may assume that the correct language input is made and may thus strengthen their mistakes in case they do not have instructions and verification abilities.

GenAI has been used in a body of literature concerning second language writing. The academic reporting of second language writing assessment has pointed to ChatGPT as a tool to assist in drafting, lexical and grammatical editing, and revision, although it brings up the issues of dependence, shallow learning, and unidentified authorship, unless teachers explicitly scaffold and delineate signing standards (Barrot, 2023). Empirical studies also propose that in case AI support is provided in the form of formative guidance, but not writing replacement, it is possible that the learners will show the increase of writing outcomes and motivation (Fathi et al., 2023). Findings however are not always similar across outcomes and situations, which supports the argument that writing gains should be assessed through standardized rubrics and systematic instructional comparisons and not isolated based on impressions. By the year 2024, research syntheses had become more organized and comparative. A systematic review applying to ChatGPT in ESL/EFL teaching assessed the existing empirical evidence that is available and published soon after the release date and revealed that most studies were centered on the writing and not on speaking, listening, or reading. It was also stated in the review that a significant number of studies have used self-report instead of objective proficiency measures, and they required

more controlled research to investigate measurable learning outcomes and conditions of implementation (Lo et al., 2024). This justifies the reason as to why one conducts quasi-experimental work comparing chatbot-assisted instruction to traditional one based on validated outcome measures.

Speaking development has emerged to be one of the key concerns since productive skills are often underdeveloped in the situation of ESL where classroom dialogue is not a significant issue. A meta-analysis of AI-based chatbots in the context of EFL speaking practice identified positive tendencies in the area of engagement, self-confidence, and intensity of practice, but emphasized that the evidence of long-term speaking proficiency benefits is yet to be established and well-designed studies and more efficient operationalization of chatbots use are needed (Du and Daniel, 2024). It means that one of the main research gaps is to go beyond the results of motivation to the quantifiable enhancement of oral proficiency, including fluency, accent, and communicative competence. Contact characteristics Material that is stronger 2025 has started being put into practice using quasi-experimental and comparative designs. System 2025 study examined the effects of AI chatbot interaction on the oral proficiency and motivation to communicate of EFL learners, with statistically significant positive effects of chatbot-supported communication in the structured conditions and with current measurements methodologies (Lu et al., 2025). This is meaningful since it proves that one can get some measurable speaking benefits in case chatbot interaction is structured as intentional communicative exercise, but not as a not-premeditated conversation. The other 2025 donation in Computer Assisted Language Learning followed the attitude of EFL learners towards AI communication and associated readiness to communicate with AI to the real output of the chat. The article reinforces the body of evidence by relating the perceptions of the learners to the behavioral participation, which is applicable to the ESL process since the greater the output and time engagement the more the behavioral utterance is strongly linked to

language development opportunities (Peng and Liang, 2025). These results indicate that the effectiveness of a chatbot can be conditional on the acceptability and engagement patterns of learners that are to be considered as the relevant variable in intervention-based researches.

A 2025 study based on a quasi-experimental research design, where the authors indicated that AI-assisted learning produced strong effects on several dimensions of writing proficiency (e.g., content, organization, vocabulary, language use, mechanics) found that a quasi-experimental design and further analyses enabled to investigate patterns in the performance of learners (Alangari, 2025). Such evidence contributes to the idea that chatbot-supported instruction it can be effective to improve the outcome of writing in case the intervention can be structured and evaluated in a systematic way. It also validates the usefulness of skill specific rubrics and objective scoring as a method of grading ESL results. Assessment conditions and effect of learners were also touched upon in research in 2025. The *Frontiers in Psychology* study on AI-facilitated speaking tests (2025) included the information of whether AI-mediated settings yield a lower anxiety level than human-mediated settings and whether a lower level of anxiety is associated with a better performance in speaking (*Frontiers in Psychology*, 2025). This line of research is the most relevant in regard to ESL learners, where anxiety usually stifles the participation and oral output, implying that chatbot can indirectly affect performance through the reduction of affective barriers in the rightful setting of instructions.

Evidence was also further enhanced in 2025 by massive syntheses about AI integration in EFL teaching. A systematic review of empirical research on AI in school EFL teaching in the *International Journal of Artificial Intelligence in Education* synthesized the results of empirical studies on the effectiveness, challenges, and pedagogical implications of AI use in schools, noting that the influence of various types of tools, instructional design, and constraints relevant to the context differed (Kundu and Bej, 2025). This confirms the opinion that AI chatbots cannot be discussed as universal effective; it is rather the

effects of their usage that are tied to their integration into teaching plans and their corresponding with learning goals and assessment criteria.

Notably, literature on system-level issues (as of 2025-2026) is also aware of meaningful learning and cognitive engagement of concern. Digital Education Outlook 2026 is evidence-based and expert-analysed on how the GenAI can enable learning, but the authors cautions that education value relies on risk management processes, including over-reliance and the use of GenAI as a learning partner and not a substitute to competently developing all underlying skills (OECD, 2026). In the case of ESL research, it would then be imperative to not only identify whether the students are doing better when engaging in language activities, but also whether they achieve transferable competence as opposed to delegating language production to AI. Quasi-experimental evidence is still found in 2026 such as tests of assisted ChatGPT methods to enhance particular language abilities in controlled classroom conditions. As an example, the Arab World English Journal conducted a quasi-experimental study in 2026 that assessed a situation involving a type of supported ChatGPT intervention to enhance grammatical accuracy during the writing process and measured of pre-intervention and post-intervention results and the structured implementation of such intervention (Vo, 2026). This research is of particular significance to your design since it reveals that quasi-experimental methods are rapidly being utilized to examine the progressively testable classroom-based language outcomes as opposed to solely depending on the attitudinal results.

All in all, the present state of AI chatbots discussed in the literature of 2020-2026 offers avenues of AI-assisted ESL application as road to more effective ESL practice in productive skills where learners are prone to limited speaking time and lack of one-to-one feedback. Nonetheless, it also proves the fact that the results strongly rely on the instructional design, teacher scaffolding, learner acceptance, and the validity of chatbot feedback (Huang et al., 2022; Lo et al., 2024; UNESCO, 2023). Evidence built to date is

progressing more towards the structured, result-oriented analyses especially quasi-experimental comparisons that indicate proficiency-related variables such as oral proficiency, vocabulary growth, grammatical proficiency, and quality of writing, and address contextual moderators such as access, teacher preparedness, and constraints on learning functions (Kundu and Bej, 2025; Lu et al., 2025; OECD, 2026).

RESEARCH METHODOLOGY

Research Design

The study was based on the quasi-experimental, non-equivalent control group design that investigated the impact of an AI chatbot-assisted intervention on English as a Second Language (ESL) in Class 10 students. The design fitted the objectives and hypotheses of the study as it intended to find out whether a significant difference existed between students who were provided with the instructions assisted with AI chatbot and students who were provided with the traditional instructions. Natural schooling where intact room sections were utilized to ensure the natural school setting and at the same time structured comparison was maintained.

Population and Sample

Students in public secondary schools located in Azar Jammu and Kashmir who are Class 10 were used as the population. Pursuant to the presence of digital facilities that are to be used to integrate the chatbot, purposive sampling was used to select the sample since it is located in Government Pilot High School Bagh, AJK. Class 10 of the two sections i.e. the Computer Science section and the Biology section were chosen. One of them was an experimental group and the other one was a control group. A total of 60 students were used as the sample, that is, 30 students in each group.

Research Instruments

An ESL Achievement Test referred to as ESL EST was designed with respect to AJK Class 10 ESL English syllabus to test the vocabulary, grammar and writing competencies. Proficiency in speech was tested with the help of structured oral activities by using analytic rubric. The

instruments were subjected to review by experts who were subject to the content validity and pilot tested to determine the reliability of the instruments.

Intervention Procedure

The intervention took place in eight to twelve weeks. Structured AI chatbot-based instruction built into routine English classes was offered to the experimental group in the course of which the typical instructional techniques were applied to the control group. The same content of syllabus was taught in both groups and they were provided with the same amount of time of instruction.

Data Collection and Analysis.

Before the intervention baseline measurement carried out to estimate the initial level of proficiency. The same assessment tools were given to the two groups after the intervention had been done. The analysis of the data was conducted through descriptive statistics, independent samples t-tests, and Analysis of

Covariance (ANCOVA) test based on whether there were any significant differences between the groups. The magnitude of the instructional impact was also measured by calculating effect size.

Ethical Considerations

The administration of Government Pilot High School Bagh, AJK, was approached and allowed to participate in the study. The participation of the participants and their guardians was obtained by informed consent. All information that was collected remained confidential and the engagement was voluntary throughout the research.

RESULTS

These results give the statistical analysis that was performed to identify the disparities between ESL proficiency in students undergoing AI chatbot-assisted instruction and conventional instruction. Descriptive statistics were then calculated and independent samples t-tests and ANCOVA.

Table 1
Reliability Analysis of Research Instruments

Instrument	Cronbach's α
ESL Achievement Test	.74
Vocabulary Subscale	.72
Grammar Subscale	.71
Writing Rubric (5 Criteria)	.76
Speaking Rubric (5 Criteria)	.73

The results of the reliability analysis of the research instruments conducted in Table 1 indicate that all the research instruments had acceptable internal consistency. The total ESL Achievement Test produced a Cronbach's alpha of .74, which was found to be acceptable. The grammar and vocabulary subscales yielded a

reliability coefficient of .72 and .71 respectively. On the same note, both the analytic writing and speaking rubrics were revealed to be acceptable having coefficients of .76 and .73. Such results affirmed that the instruments were reliable enough to be used to measure ESL proficiency amongst Class 10 students.

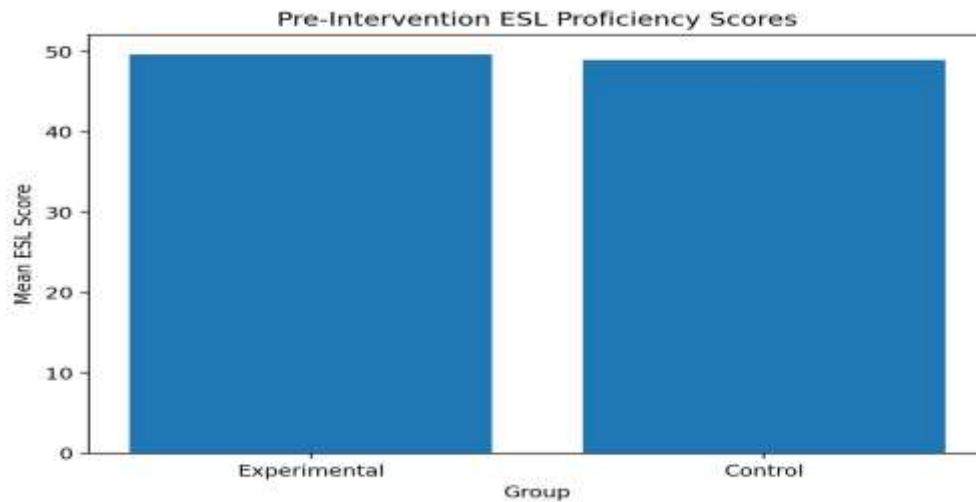
Table 2
Pre-Intervention ESL Proficiency Scores of Experimental and Control Groups

Group	N	M	SD	t	p	Cohen's d
Experimental	30	49.63	6.82	0.39	.698	0.09
Control	30	48.97	7.14			

Table 2 shows pre intervention ESL proficiency scores of the experimental and control groups. The experimental group received a mean of 49.63 (SD = 6.82), and the control group received a mean of 48.97 (SD = 7.14). Independent samples t-test found that the difference between two

groups was not significant $t(58) = 0.39, p = .698$. This was very small (0.09 Cohen d) and suggests that there was no practical difference at baseline. Such findings indicate that the two groups were initially equal in terms of statistics.

Figure 1
 Pre-intervention comparison of mean ESL scores.



The visual representation of this comparison in Figure 1 includes almost the same mean scores of both groups before the instructional treatment. The low statistical result of the difference in bar

heights helps to prove the statistical result that there was no significant difference between the experimental and the control group at the pre-intervention stage.

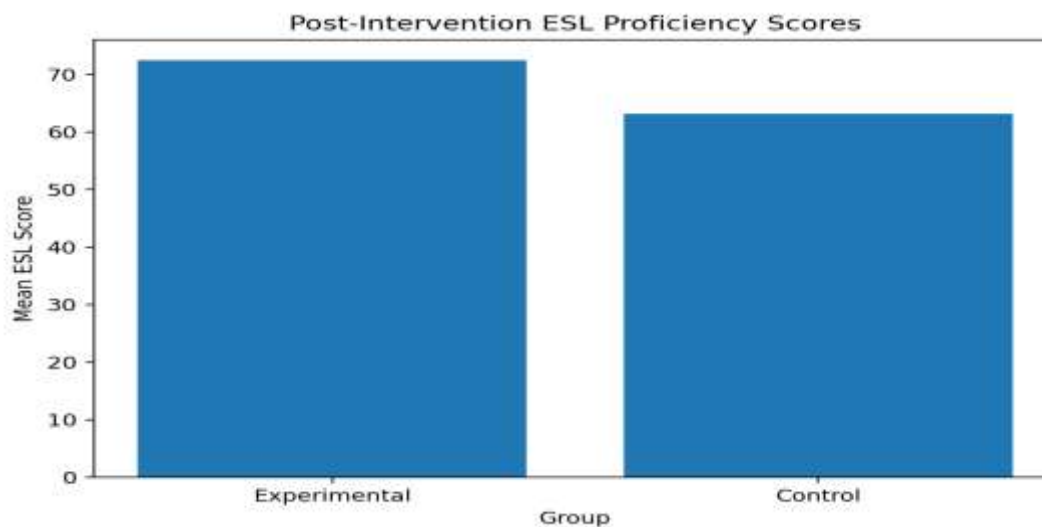
Table 3
 Post-Intervention ESL Proficiency Scores of Experimental and Control Groups

Group	N	M	SD	t	p	Cohen's d
Experimental	30	72.41	8.15	4.52	< .001	0.98
Control	30	63.18	7.96			

Table 3 shows the (after intervention) ESL proficiency scores of the experimental and control groups. The experimental condition recorded a mean score of 72.41 (SD = 8.15), and that of the control group was 63.18 (SD = 7.96). A t-test independent samples test found that this difference was statistically significant $t(58) = 4.52, p < .001$. The size of the effect (Cohen d = 0.98)

was large, which means that there is a significant practical influence of AI chatbot-assisted instruction on the ESL proficiency. These findings prove that students who were taught through chatbots scored tremendously higher in comparison to those who were taught traditionally.

Figure 2
 Post-intervention comparison of mean ESL scores.



The post-intervention comparison presented in figure 2 is in a graphic format, and one can easily see that there is a difference in the mean scores representing the two groups. The visual representation of the improvement, which is

statistically significant in Table 2, is further validated with the help of the higher bar of the experimental group, which confirms the effectiveness of AI chatbot-assisted teaching.

Table 4
 Skill-wise Post-Intervention Comparison between Groups

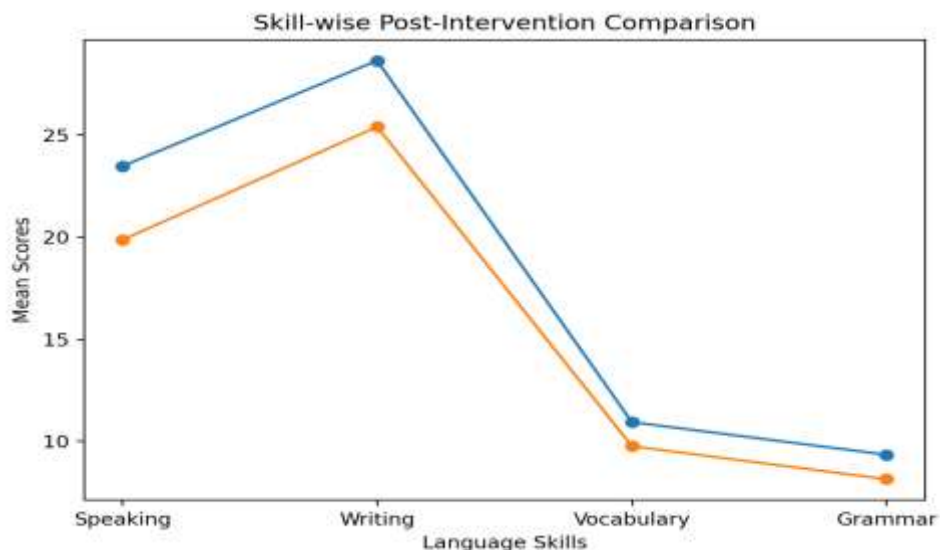
Skill	Group	M	SD	t	p
Speaking	Experimental	23.48	3.12	3.96	< .001
	Control	19.87	3.45		
Writing	Experimental	28.65	4.21	2.84	.006
	Control	25.41	4.37		
Vocabulary	Experimental	10.94	1.86	2.12	.038
	Control	9.76	2.04		
Grammar	Experimental	9.34	1.72	2.27	.027
	Control	8.14	1.89		

Table 4 indicates the comparison of the experimental and control groups in terms of speaking, writing, vocabulary, and grammar in terms of skills after the interventions. During speech, the experimental group got a mean of 23.48 and (SD) = 3.12 and the control group got 19.87 and (SD) = 3.45, which was significantly different and thus, $t(58) = 3.96$ and $p = 0.001$. In the written version the means were 28.65 (SD = 4.21) and 25.41 (SD = 4.37) in the experimental

and control group respectively, ($t = 2.84$) $p = .006$. There was also significant difference in vocabulary, $t(58) = 2.12$, $p = .038$, grammar, $t(58) = 2.27$, $p = .027$, and higher mean scores has always been achieved in the experimental group. These results demonstrate that the use of AI chatbot-mediated instruction has a positive impact on all the quantified elements of language.

Figure 3

Skill-wise comparison of mean scores (line chart).



The comparison of the skills is done using a line graph as shown in figure 3. The line of the experimental group is always higher in all four abilities of language, which is a perfect indication of high-quality performance in speaking, writing,

vocabulary, and grammar. In Table 3, the statistical data is corroborated by the graphical trend, that is, the AI chatbot-assessed instruction provided proves to be effective in various aspects of ESL proficiency.

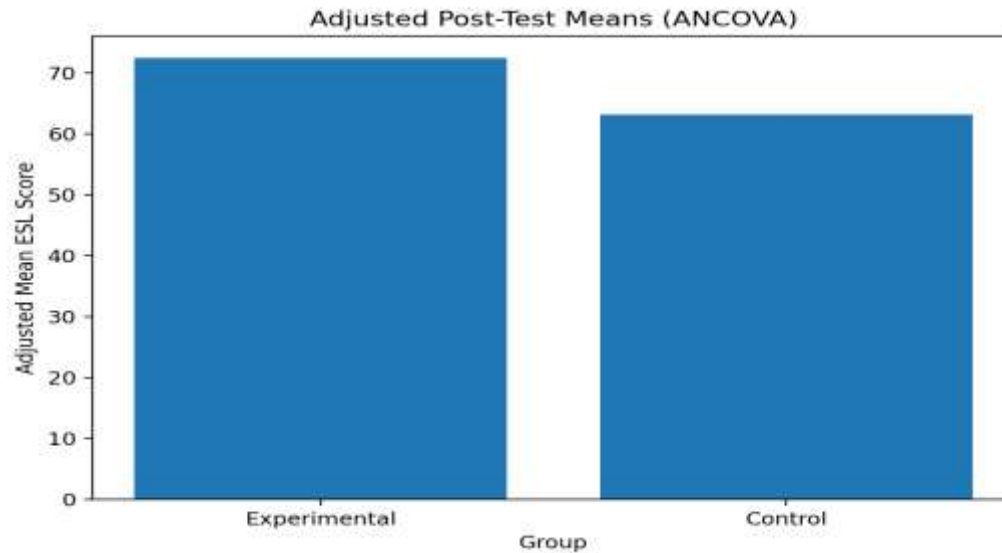
Table 5
 ANCOVA Results for Post-Intervention ESL Proficiency

Source	SS	df	MS	F	p
Pre-test (Covariate)	214.56	1	214.56	6.47	.013
Instructional Method	621.74	1	621.74	18.74	< .001
Error	1867.29	57	32.76		

Table 5 gives the findings of the Analysis of Covariance (ANCOVA) that was done to establish the impact of instruction approach on post intervention ESL proficiency with baseline scores held constant. The covariate (pre-test scores) was also statistically significant, $f(1, 57) = 6.47$ the $p = .013$ to show that initial proficiency was significantly related to post-test performance.

More to the point, the impact of the instructional method was found to be statistically significant, $f(1, 57) = 18.74$, $p < .001$. This finding indicates that AI chatbot guided instruction significantly influenced the ESL proficiency even when differences on the pre-existing differences were controlled.

Figure 4
 Adjusted post-test means estimated via ANCOVA.



The ANCOVA analysis provided in figure 4 shows the adjusted post-test means. The adjusted mean is much greater in the experimental compared to the control group which supports statistically, but visually, the fact that there is a significant effect of instructional method on ESL outcomes. The value puts to rest the fact that the progress that was seen in the experimental group was not because of the differences in the basis but could be credited to the AI chatbot-based instructional solution.

Findings of the Study

- The study findings showed that no statistically significant difference could be seen between experimental and control group at the pre-intervention stage which ensured baseline equivalence in ESL proficiency.
- It was found that students in the AI chatbot-assisted instruction performed far better in terms of post-intervention ESL proficiency test scores than those who were in the traditional instruction.
- A statistically significant difference in total ESL performance was also proved with a large effect size, meaning that the effect of AI chatbot-assisted instruction is not only statistically but also considerably significant.

- Findings indicated that, the experimental group had better talking competence as compared to the control group and there was a significant improvement in fluency, pronunciation and communicative effectiveness.

- Findings also provided a substantial improvement in writing in students who received AI chatbot-guided learning, including writing organization, vocabulary use, and grammatical correctness.

- It was found that statistically significant changes favoring the experimental group were found to be favorable in improvement in vocabulary and grammatical competence.

- Results of the ANCOVA showed that the instructional approach was a strong predictor of the post intervention ESL competence despite matching the with the baseline scores.

- All in all, the results attested to the fact that ESL learning outcomes with the use of AI chatbot-assisted instructions were manifestly superior to conventional ways of teaching Class 10 students.

Conclusions of the Study

- According to the study findings, the AI chatbot-led instruction was more productive than the traditional instruction in the enhancement of the overall ESL proficiency among Class 10 students.

- The researchers have found that AI chatbot technology applied in the classroom setting of the English language learning helped to significantly improve speaking, writing, vocabulary, and grammatical competencies of students.
- The results justify the statement that structured chatbot or human interaction created significant practice possibilities, real-time feedback, and the enhancement of the engagement among learners and resulted in a better language outcome.
- It was also discovered that the statistically and practically significant effect of AI chatbot-assisted instruction was observed among students of communicative competence, especially speaking performance.
- The study also found that the progress made in the ESL proficiency was due to the instruction procedure and not the difference that existed before as ascertained by covariance analysis.
- In general, the research has found that the use of AI chatbots can be offered as a useful method of pedagogy to improve the performance of English language learning in the secondary school system in Azad Jammu and Kashmir.

Implications of the Study

- The research results suggest that AI chatbot-guided learning could be easily deployed in the secondary-level English classrooms to improve ESL proficiency. AZJAM institutions can embrace use of AI based tools as additional educational materials to aid language instruction.
- This paper has proposed that to improve communicative competence and learner engagement, policy-makers and curriculum planners need to consider the potential of incorporating technology-based learning practices in the AJK framework of secondary English curriculum.
- The implication of the study is that English language instructors can be helped by professional development courses related to how to use AI chatbots in their pedagogical practice to enable interactive learning, formative feedback, and student-centered teaching.

- The fact that the speaking and writing skills were positively affected shows that the AI chatbot-based learning may be used to tackle the lingering issues related to productive language skills as it is often identified in ESL learners at the secondary level.
- The results also suggest that schools with elementary digital infrastructure have an opportunity to utilize AI technologies to generate blended learning environments and facilitate independent learning and extensive training outside the classroom.
- At a greater scale, the present research also indicates that responsible and organized introduction of artificial intelligence into learning and teaching can serve to enhance the overall language learning results and update the approach to teaching practices in developing education settings, like Azar Jammu and Kashmir.

Recommendations of the Study

- It is suggested that AI chatbot facilitated learning should be adopted in secondary schools of Azar Jammu and Kashmir as an additional means to increase ESL proficiency.
- It is suggested that the AJK education authorities should take into consideration implementing technology-based language learning strategies in the curriculum guidelines and teaching patterns in the secondary level.
- English language instructors are to receive professional training on effective use of AI chatbots in the teaching process, which will lead to a consistent implementation approach and significant integration with classrooms.
- Schools in house should dedicate efforts to making basic digital infrastructure, such as access to the internet and supervised use of the equipment, available to ensure effective use of AI-based learning equipment.
- It is suggested that AI chatbot activities have to be matched to the recommended Class 10 English syllabus to ensure the relevance of the curriculum and academic consistency.
- Future researchers are advised to repeat the same studies by grade levels, in other schools, and during the intervention period to further

support the results and make known implications.

- More studies can also be conducted to determine the long-term consequences of AI chatbot-based instruction on learner motivation, independent learning, and test performance in ESL.

REFERENCES

- Alangari, T. S. (2025). The effect of AI-assisted learning on EFL writing proficiency: Quasi-experimental and cluster analysis. *Educational Process: International Journal*, 17, e2025345.
- Albedah, F. (2025). Artificial intelligence in language education: A systematic review of multilingual applications, large language models, and emerging challenges. *Language Teaching Research Quarterly*, 49, 247–268.
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. Bureau of Statistics, Planning & Development Department, Government of Azad Jammu & Kashmir. (2023). *Azad Jammu & Kashmir Statistical Year Book 2023*.
- Dawn. (2024). AJK to launch new 'Education Policy' in collaboration with Unicef.
- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 5, 100230.
- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 5, 100230.
- Elementary & Secondary Education Department, Government of Azad Jammu & Kashmir. (n.d.). *E&SE Department AJK(official portal)*.
- Fathi, J., Ebadi, S., & Amini, M. (2023). Enhancing academic writing skills and motivation: Assessing the effectiveness of ChatGPT in EFL writing. *Frontiers in Psychology*, 14, 1260843.
- Frontiers in Psychology. (2025). Reducing anxiety and enhancing performance: The impact of AI chatbots on speaking assessment and performance. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2025.1745942>
- Fryer, L. K., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology*, 24(2), 8–22.
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257.
- Koc, F. S., & Savas, P. (2025). The use of artificially intelligent chatbots in English language learning: A systematic meta synthesis study of articles published between 2010 and 2024. *ReCALL*, 37(1), 4–21. <https://doi.org/10.1017/S0958344024000168>
- Kundu, A., & Bej, T. (2025). Transforming EFL teaching with AI: A systematic review of empirical studies. *International Journal of Artificial Intelligence in Education*, 35, 2281–2314. <https://doi.org/10.1007/s40593-025-00470-0>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S.-Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, 11, Article 342. <https://doi.org/10.1186/s40561-024-00342-5>

- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). Academic Press.
- Lu, X., [Author initials as published], & [Co-authors as published]. (2025). The impact of AI chatbots on EFL learners' oral proficiency and willingness to communicate. *System*, 103919. <https://doi.org/10.1016/j.system.2025.103919>
- Ministry of Federal Education and Professional Training. (2017-2025). *National Education Policy 2017-2025 (Pakistan)*. OECD. (2026). *OECD Digital Education Outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. <https://doi.org/10.1787/062a7394-en>
- Peng, J.-E., & Liang, W. (2025). Willingness to communicate with artificial intelligence (AI)? Insights from tracking EFL learners' perceived acceptance and chat output. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2025.2486147>
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford University Press.
- UNESCO. (2023). *Guidance for generative AI in education and research*.
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO.
- UNICEF. (n.d.). *Education | UNICEF Pakistan*.
- Vo, T. L. (2026). A quasi-experimental study on enhancing EFL learners' grammar precision in writing skills under supported ChatGPT approach at a university in a remote region of Vietnam. *Arab World English Journal*.

