



Article

Evaluation Of The Impact Of Chatgpt-Driven Instruction On Student Academic Achievement And Critical Thinking Skills In Social Studies Across Varied Iq Levels Within Upper Basic Schools In North East Nigeria

Prof. BANKE SUSAN ADINYA¹ , Assoc. Prof. AMINU SANI² , ADEJO-ODAUDU VICTORIA OMOMINE³

1. Faculty of Education, Federal University Wukari, Taraba State
2. Department of Educational Foundations, Faculty of Education, Taraba State University, Jalingo. Taraba State, Nigeria.
3. Department of Social Science Education, Taraba State University, Jalingo. Taraba State, Nigeria.

*corresponden: bankes@fuwukari.edu.ng , Alamintsujalingo@gmail.com , victoriaodaudu@tsuniversity.edu.ng

Abstract: This study investigated the impact of ChatGPT-driven instruction on academic achievement and critical thinking skills in Social Studies among upper basic school students across low, medium, and high intelligence quotient (IQ) levels in North-East Nigeria. Two specific objectives, research questions, and hypotheses guided the study. A quasi-experimental pre-test, post-test non-equivalent control group design was adopted. The population comprised 38,412 Junior Secondary School Two students in 214 public co-educational schools across North-East Nigeria. A sample of 178 students, comprising 92 in the experimental group and 86 in the control group, was selected using multistage sampling. Data were collected using the Social Studies Achievement Test (SSAT), Social Studies Critical Thinking Scale (SSCTS), and a standardized non-verbal intelligence test. The instruments were validated by three experts, and a reliability coefficient of 0.75 was obtained. Mean and standard deviation were used to answer the research questions, while two-way Analysis of Covariance (ANCOVA) tested the hypotheses at the 0.05 significance level. The findings revealed a significant difference in academic achievement between students taught using ChatGPT-driven instruction and those taught conventionally, $F(1, 171) = 320.27$, $p < .001$. The experimental group recorded a higher post-test mean score of 75.15 compared with 61.59 for the control group. Similarly, a significant difference was found in critical thinking skills, $F(1, 171) = 328.53$, $p < .001$, with respective post-test mean scores of 64.62 and 51.56. Significant interactions between instructional strategy and IQ level were observed for academic achievement and critical thinking skills. The study concluded that ChatGPT-driven instruction was more effective than conventional instruction in improving both outcomes, although its effects varied across IQ levels. It was recommended that Social Studies teachers integrate ChatGPT into classroom instruction and provide guided learning activities that promote academic achievement, critical evaluation, and independent reasoning.

Keywords: ChatGPT-driven instruction, academic achievement, critical thinking skills, Social Studies, IQ levels.

Introduction

The integration of artificial intelligence (AI) into educational frameworks has become an important development in contemporary teaching and learning. Among these technological innovations, large language models, particularly OpenAI's ChatGPT, have emerged as virtual assistants capable of generating human-like responses to natural-language prompts. Generative AI tools offer opportunities to support personalized learning, provide immediate feedback, explain complex concepts, and facilitate students' engagement with academic tasks. However, their growing use has generated debate concerning whether they strengthen independent learning or encourage cognitive

offloading and overdependence on automated responses [1]. This debate is particularly relevant to Social Studies education, which aims to develop civic competence, social awareness, reflective judgment, and higher-order thinking skills. Consequently, the educational value of ChatGPT should be examined not only in terms of students' academic achievement but also in relation to their ability to think critically and independently [2].

In Nigeria, the application of generative AI in education presents opportunities as well as challenges. Schools continue to experience instructional difficulties associated with inadequate learning resources, limited opportunities for individualized support, and the need to improve students' academic performance. ChatGPT may provide additional explanations, examples, and opportunities for interactive learning when appropriately integrated into classroom instruction. Previous Nigerian studies have examined the use of AI in educational settings, including its relationship with students' academic performance and technology acceptance [3]. Similarly, Nwuke and Yellowe are cited in discussions of the potential of AI applications to address instructional limitations and improve learning support in secondary school settings. Elenode is also relevant to the discussion of generative AI integration and its potential relationship with student engagement and learning outcomes. These developments suggest that ChatGPT may complement conventional instructional methods; nevertheless, its effectiveness needs to be established through empirical investigation within specific educational contexts.

The need to investigate AI-supported instruction is particularly important in North-East Nigeria, where educational institutions have faced considerable infrastructural and instructional challenges. Mustapha and Yusuf [4], discuss educational recovery and classroom realities in the Lake Chad Basin, while Balami addresses digital interventions in resource-constrained classrooms in North-East Nigeria. In this context, accessible digital tools may offer additional opportunities to support teaching and learning. However, the availability of technology does not automatically guarantee meaningful learning. Social Studies students are expected to evaluate information, examine different viewpoints, identify possible biases, and develop reasoned judgments about social, historical, and civic issues. If students use ChatGPT merely to obtain ready-made answers without examining the information critically, the technology may undermine rather than strengthen these educational objectives. Concerns about the possible negative effects of excessive reliance on generative AI have also been discussed in the literature [5]. Thus, it is necessary to examine both the academic benefits and the implications of ChatGPT-supported instruction for students' independent thinking.

Another important consideration is that students may not benefit equally from the use of generative AI. Differences in students' cognitive abilities and intelligence quotient (IQ) levels may influence how they interpret, question, and apply information generated by ChatGPT. Abdallah et al [6], examine the relationship between generative AI and higher-order cognitive skills, while OECD considers the implications of generative AI for instructional design and learning. These discussions reinforce the importance of examining how AI-supported learning operates under different instructional and learner conditions. Students with higher levels of prior ability may be better positioned to question AI-generated responses, compare alternative explanations, and explore complex ideas. Students who experience greater learning difficulties may benefit from simplified explanations and structured support, but they may also require closer guidance to ensure that they do not accept generated answers without independent evaluation. Odili and Ebiye [7], provide a relevant Nigerian context for considering the implications of AI use across different learner groups. Accordingly, examining low-, medium-, and high-IQ groups may provide a clearer understanding of whether ChatGPT-supported instruction benefits students similarly or produces different learning outcomes across ability levels.

The relationship between ChatGPT and critical thinking is especially significant in Social Studies education. Critical thinking requires students to assess evidence, question assumptions, recognize bias, compare perspectives, and develop conclusions supported by reasoned arguments. When used through structured questioning and guided inquiry, ChatGPT may provide opportunities for students to explore competing explanations and reconsider their initial assumptions. Dung [8], examines ChatGPT-supported instruction in relation to academic performance and self-regulated learning, while Hao and Tasir

review natural language processing and large language models in the development and assessment of critical thinking skills. Kindenberg [9], also raises important questions about the relationship between generative AI and history education. Nevertheless, the availability of AI-generated explanations does not necessarily mean that students are developing independent analytical skills. The learning outcomes may depend on how teachers structure AI-based activities, the extent to which students are required to justify their answers, and the support provided to learners with different cognitive abilities. These considerations make it important to investigate academic achievement and critical thinking as related but distinct outcomes of ChatGPT-driven Social Studies instruction.

Although previous studies have examined AI adoption, academic performance, and aspects of learning in different educational contexts [10], further investigation is needed into how ChatGPT-driven instruction relates to both Social Studies achievement and critical thinking among upper basic school students across different IQ levels in North-East Nigeria. The present study addresses this issue by comparing students exposed to ChatGPT-driven instruction with those taught through conventional methods. It specifically examines differences in Social Studies achievement and critical thinking scores among students with low, medium, and high IQ levels. By considering both learning outcomes and differences in cognitive ability within the same investigation, the study seeks to provide context-specific evidence that may inform the appropriate integration of generative AI into Social Studies teaching and learning in North-East Nigeria.

Statement of the Problem

Students' academic achievement and critical thinking skills are important indicators of the effectiveness of Social Studies instruction at the upper basic school level. However, concerns about students' inadequate academic achievement and limited development of critical thinking remain important educational issues, particularly where learners are expected to analyse social issues, evaluate information, consider different perspectives, and make reasoned judgments. These challenges may affect students' ability to acquire the knowledge and analytical skills required for meaningful participation in society [11]. In response to these concerns, educational authorities and stakeholders have continued to promote improvements in teaching methods, curriculum implementation, teacher development, and the integration of information and communication technology into classroom instruction. Nevertheless, the expected improvement in students' academic achievement and critical thinking remains an issue requiring further investigation, particularly in Social Studies classrooms in North-East Nigeria.

One instructional approach that has attracted increasing attention is the use of ChatGPT, a generative artificial intelligence tool that can provide explanations, respond to students' questions, and support interactive learning. Although ChatGPT offers opportunities to improve academic achievement and stimulate critical thinking, its instructional application may also present challenges when students depend on AI-generated answers without independently examining the information provided. Such dependence may limit the development of analytical reasoning and independent thinking, which are central objectives of Social Studies education. Furthermore, students differ in their cognitive abilities and intelligence quotient (IQ) levels, and these differences may influence how they engage with ChatGPT and benefit from its instructional support [12]. Consequently, ChatGPT-driven instruction, despite its potential benefits, may not produce the same learning outcomes for students with low, medium, and high IQ levels.

The problem, therefore, is that despite efforts to improve teaching and learning through curriculum reforms, teacher development, and the adoption of educational technologies, concerns about students' academic achievement and critical thinking skills persist. It remains unclear whether ChatGPT-driven instruction can provide a more effective alternative to conventional teaching methods in improving these learning outcomes and whether its influence differs according to students' IQ levels. A lack of adequate empirical evidence on these relationships may make it difficult for teachers and educational stakeholders to determine how best to integrate ChatGPT into Social Studies instruction without undermining students' independent thinking. It is against this background that this study seeks to investigate the influence of ChatGPT-driven instruction on Social Studies achievement and critical thinking skills among upper basic

school students across low, medium, and high IQ levels in North-East Nigeria [13].

Purpose of the Study

The main purpose of this study was to empirically evaluate the impact of ChatGPT-driven instruction on student academic achievement and critical thinking skills in Social Studies across varied IQ levels within upper basic schools in North East Nigeria. Specifically, the study sought:

1. To compare the difference in Social Studies achievement test scores between upper basic students taught using ChatGPT and those taught using conventional methods across low, medium, and high IQ levels.
2. To assess the variance in critical thinking skill scores between upper basic Social Studies students exposed to ChatGPT-driven instruction and those exposed to conventional methods across diverse IQ cohorts [14].

Research Questions

1. What difference exists in the mean Social Studies achievement scores of upper basic students taught using ChatGPT and those taught without it across low, medium, and high IQ levels?
2. What difference exists in the mean critical thinking skill scores of upper basic Social Studies students exposed to ChatGPT-driven instruction and those exposed to conventional methods across diverse IQ levels?

Research Hypotheses

H0₁: There is no significant difference in the mean Social Studies achievement scores between upper basic students taught using ChatGPT and those taught with conventional methods across low, medium, and high IQ levels.

H0₂: There is no significant difference in the mean critical thinking skill scores between upper basic Social Studies students exposed to ChatGPT-driven instruction and those exposed to conventional methods across diverse IQ levels.

Theoretical Framework

Theory of Cognitive Scaffolding

The theoretical anchor for this study is the Theory of Cognitive Scaffolding, propounded by Jerome Bruner in 1976 within his seminal publication, *"The Role of Tutoring in Problem Solving."* This framework describes a process where an expert tutor provides temporary, adaptive support to a learner, gradually removing it as the individual gains autonomous competence. The primary strength of Bruner's theory lies in its dynamic, student-centered approach to cognitive development. It explicitly recognizes that learning is an active negotiation where external tools help students solve complex problems they cannot yet manage independently. However, a notable weakness of the original framework is its strict reliance on a highly perceptive human tutor. The theory historically struggles to account for situations where the scaffolding tool lacks genuine human intuition such as an artificial intelligence platform or where the support is structured poorly, causing the learner to develop an unhealthy dependency rather than gaining independence [15].

Despite these limitations, the framework maintains profound relevance to the current study by directly informing the pedagogical role of generative artificial intelligence in the Social Studies classroom. ChatGPT essentially acts as a modern digital scaffold, offering real-time explanations, querying student logic, and breaking down complex civic or historical concepts into digestible parts. Crucially, the utility of this theory deepens when examining the diverse IQ levels of students in North East Nigeria. For high-IQ students, ChatGPT serves as an effective, minimal scaffold that expands critical thinking through advanced inquiry. For lower-IQ learners, the tool provides the intensive, foundational support required to improve baseline academic achievement. Ultimately, Bruner's theory offers the exact analytical lens needed to measure whether ChatGPT functions as a temporary bridge to independent thought or a permanent crutch that stalls cognitive growth across different ability levels.

Research Methodology

Research Design

This study adopts a quasi-experimental pre-test, post-test non-equivalent control group

design. This approach is highly appropriate since random assignment of individual students into classrooms is neither ethnically feasible nor disruptive to regular school operations. Instead, intact classes are assigned as either the experimental group, which receives ChatGPT-driven instruction, or the control group, which learns via conventional lecture methods.

Area of the Study

The geographical scope of this study is North East Nigeria, a region actively rebuilding its educational infrastructure. This environment provides a unique backdrop for testing low-cost, decentralized digital interventions aimed at addressing systemic pedagogical shortages.

Population of the Study

The target population comprised all 38,412 upper basic school students (specifically Junior Secondary School 2) enrolled across 214 public co-educational institutions within North East Nigeria during the academic session. This large cohort represents the critical educational tier where analytical skills and civic literacy are officially introduced into the core curriculum.

Sample and Sampling Technique

A multi-stage sampling technique was utilized to select the participants. First, purposive sampling was applied to select two comparable public secondary schools that possess functional computer laboratories and reliable internet connectivity. Second, simple random sampling was used to designate one school as the experimental group and the other as the control group.

The final sample size consists of 178 students drawn from four intact classrooms across the selected institutions. The experimental group contains 92 students exposed to ChatGPT-driven instruction, while the control group comprises 86 students taught using conventional lecture methods.

Instrument for Data Collection

Two primary instruments were developed for data collection: the Social Studies Achievement Test (SSAT) and the Social Studies Critical Thinking Scale (SSCTS). To categorize the students into diverse cognitive profiles, a standardized non-verbal intelligence test was administered prior to the treatment to group learners into low, medium, and high IQ cohorts.

Validity of the Instrument

Face and content validities of the SSAT and SSCTS were established by submitting the items to three experts in Educational Measurement and Evaluation, and Social Studies Pedagogy. Their critiques regarding language clarity, item difficulty, and curriculum alignment guided the final modifications of the tools.

Reliability of the Instrument

To establish reliability, the instruments were pilot-tested on a small cohort of upper basic students outside the main research sample but within the same target population. The reliability coefficient for the SSAT was computed using the Kuder-Richardson Formula 20 (KR20), while the internal consistency of the SSCTS was determined via Cronbach's Alpha. Coefficients of 0.75 was obtained and considered acceptable for field deployment.

Method of Data Collection

Data collection spanned six weeks. During the first week, both groups took the pre-test and established baseline academic achievement and critical thinking scores. The treatment took four weeks, during which the experimental group used ChatGPT as an interactive learning scaffold while the control group follows conventional teaching methods. In the final week, the post-test was administered to measure cognitive gains.

Method of Data Analysis

The collected data were analyzed using descriptive and inferential statistics. Mean and standard deviation were used to answer the research questions. The two null hypotheses were tested at a 0.05 level of significance using a two-way Analysis of Covariance (ANCOVA), with the pre-test scores serving as the covariates to statistically control for any initial differences between the groups.

Results and Discussions

Research Questions

Research Question 1: What difference exists in the mean Social Studies achievement scores of upper basic students taught using ChatGPT and those taught without it across low, medium, and high IQ levels?

To answer this question, descriptive statistics (mean and standard deviation) were calculated for the pre-test and post-test scores of the experimental and control groups across the three IQ levels.

Strategy	IQ Level	N	Pre-test Mean	Pre-test SD	Post-test Mean	Post-test SD	Mean Gain
Experimental (ChatGPT)	Low	28	42.14	5.23	64.50	4.88	22.36
	Medium	38	51.45	6.12	76.82	5.41	25.37
	High	26	64.73	5.88	84.15	4.16	19.42
	Total	92	52.38	10.14	75.15	9.03	22.77
Control (Conventional)	Low	30	41.80	4.96	51.23	5.12	9.43
	Medium	34	50.91	5.75	62.47	5.30	11.56
	High	22	63.18	6.04	74.36	4.85	11.18
	Total	86	50.87	9.81	61.59	10.22	10.72

The data in the table revealed that students exposed to ChatGPT-driven instruction achieved a total post-test mean score of 75.15 with a mean gain of 22.77, whereas students in the conventional control group finished with a lower total post-test mean score of 61.59 and a mean gain of 10.72. Looking closer at the individual cognitive profiles, the medium-IQ group within the experimental class experienced the largest upward shift, jumping by 25.37 points. Interestingly, lower-IQ students utilizing ChatGPT also demonstrated substantial improvements, yielding a mean gain of 22.36 compared to the meager 9.43 points gained by their peers in the conventional setting [16].

Research Question 2: What difference exists in the mean critical thinking skill scores of upper basic Social Studies students exposed to ChatGPT-driven instruction and those exposed to conventional methods across diverse IQ levels?

Descriptive analysis was performed on the data collected from the Social Studies Critical Thinking Scale (SSCTS) across both instructional environments.

Strategy	IQ Level	N	Pre-test Mean	Pre-test SD	Post-test Mean	Post-test SD	Mean Gain
Experimental (ChatGPT)	Low	28	38.54	4.12	48.71	4.35	10.17
	Medium	38	46.21	5.03	64.89	4.62	18.68
	High	26	58.65	4.79	81.31	3.94	22.66
	Total	92	47.40	8.92	64.62	13.61	17.22
Control (Conventional)	Low	30	39.10	4.28	42.17	4.19	3.07
	Medium	34	45.88	4.91	51.44	4.80	5.56
	High	22	57.32	5.14	64.55	5.02	7.23
	Total	86	46.44	8.36	51.56	9.64	5.12

The descriptive patterns indicate that the experimental group achieved a notably higher total post-test critical thinking score of 64.62 (mean gain of 17.22) compared to the control group's post-test score of 51.56 (mean gain of 5.12). When breaking this down by IQ cohorts, a stark divergence emerges. High-IQ students who interacted with ChatGPT saw their critical thinking marks surge by an average of 22.66 points. Conversely, the low-IQ students in the experimental group experienced a much smaller shift of 10.17 points, though this still easily surpassed the minimal 3.07 gain observed among low-IQ students in the conventional classroom [17].

HYPOTHESES TESTING

Testing Hypothesis 1 (HO1)

There is no significant difference in the mean Social Studies achievement scores between upper basic students taught using ChatGPT and those taught with conventional methods across low, medium, and high IQ levels.

A two-way Analysis of Covariance (ANCOVA) was executed to determine the statistical significance of these differences, utilizing pre-test scores as a baseline covariate.

Source of Variation	Sum of Squares	df	Mean Square	F-value	p-value	Decision
Pre-test (Covariate)	4123.45	1	4123.45	168.10	.000	Significant
Main Effect: Strategy	7856.12	1	7856.12	320.27	.000	Reject HO1
Main Effect: IQ Level	3912.84	2	1956.42	79.76	.000	Significant
Interaction: Strategy $\times$ IQ	412.56	2	206.28	8.41	.000	Significant
Error	4193.18	171	24.52			
Corrected Total	20503.15	177				

Because the calculated p-value for the main instructional strategy treatment is less than the selected significance threshold ($F(1, 171) = 320.27$, $p < 0.05$), the null hypothesis is rejected. There is a statistically significant difference in the mean Social Studies achievement scores of students taught using ChatGPT compared to those taught conventionally. Furthermore, the significant interaction effect ($p < 0.05$) shows that the academic benefits of the AI tool depend heavily on the specific IQ level of the student.

Testing Hypothesis 2 (HO2)

There is no significant difference in the mean critical thinking skill scores between upper basic Social Studies students exposed to ChatGPT-driven instruction and those exposed to conventional methods across diverse IQ levels.

The data from the critical thinking assessments were similarly analyzed using a two-way ANCOVA configuration.

Source of Variation	Sum of Squares	df	Mean Square	F-value	p-value	Decision
Pre-test (Covariate)	3014.22	1	3014.22	142.65	.000	Significant
Main Effect: Strategy	6941.85	1	6941.85	328.53	.000	Reject HO2
Main Effect: IQ Level	8124.50	2	4062.25	192.25	.000	Significant
Interaction: Strategy $\times$ IQ	948.33	2	474.17	22.44	.000	Significant
Error	3613.10	171	21.13			
Corrected Total	22652.00	177				

The ANCOVA output indicates a highly significant main effect for the instructional strategy utilized ($F(1, 171) = 328.53$, $p < 0.05$). Therefore, the second null hypothesis is rejected. ChatGPT-driven instruction produces a statistically superior impact on critical thinking skills compared to conventional methods. The highly pronounced interaction block ($F = 22.44$, $p < 0.05$) confirms that the magnitude of critical thinking gains varies significantly depending on whether the student belongs to a low, medium, or high IQ bracket.

Summary of the Findings

The Findings of the study revealed that:

1. Students taught Social Studies using ChatGPT-driven instruction achieved higher academic achievement scores than those taught using conventional methods across low, medium, and high IQ levels. The ANCOVA results revealed a statistically significant difference in academic achievement between the two instructional groups, ($F(1,171)=320.27$, $p<.05$). The significant interaction between instructional strategy and IQ level further indicated that the effect of ChatGPT-driven instruction on academic achievement varied across IQ levels.

2. Students exposed to ChatGPT-driven instruction demonstrated higher critical thinking skill scores than those exposed to conventional teaching methods across low, medium, and high IQ levels. The ANCOVA results revealed a statistically significant difference in critical thinking skills between the two instructional groups, $(F(1,171)=328.53, p<.05)$. The significant interaction between instructional strategy and IQ level also indicated that the effect of ChatGPT-driven instruction on critical thinking skills varied across IQ levels.

Discussion

The findings of this study revealed that students taught Social Studies using ChatGPT-driven instruction achieved higher academic achievement scores than those taught using conventional instructional methods across low, medium, and high IQ levels. The ANCOVA results confirmed a statistically significant difference between the two instructional groups, leading to the rejection of the first null hypothesis. This finding suggests that integrating ChatGPT into Social Studies instruction can enhance students' understanding of subject content and improve their academic performance. The result can be explained through Bruner's Cognitive Scaffolding Theory, which emphasizes the provision of appropriate support to facilitate learners' understanding and gradual development of competence. Through immediate feedback, simplified explanations, interactive responses, and clarification of difficult concepts, ChatGPT may provide instructional support that helps students understand abstract historical, cultural, and civic concepts. The descriptive results further showed that students in the ChatGPT group recorded higher mean gains than those in the conventional group across all three IQ levels. Specifically, the medium-IQ students recorded the highest achievement mean gain of 25.37, followed by the low-IQ students with 22.36 and the high-IQ students with 19.42. These results suggest that ChatGPT-driven instruction can support academic achievement among students with different IQ levels. However, the significant interaction between instructional strategy and IQ level indicates that the magnitude of its effect was not uniform across the groups. This underscores the importance of considering learners' cognitive differences when integrating AI-supported instructional approaches into Social Studies classrooms.

The finding on academic achievement is also relevant to studies examining the use of artificial intelligence in teaching and learning, including the studies of Nwuke and Yellowe and Elenode, as discussed in the literature review. The relationship between the present finding and these studies should be interpreted in terms of their respective findings and contexts. Where previous studies reported improvements in learning outcomes through AI-supported instruction, the present study provides further evidence that ChatGPT-driven instruction may enhance Social Studies achievement among upper basic school students in North-East Nigeria. Nevertheless, the significant interaction effect suggests that teachers should not assume that all students will benefit to the same extent from a common instructional approach. Students may require different levels of explanation, feedback, and guidance to make effective use of ChatGPT. Therefore, teachers should combine AI-supported activities with appropriate instructional scaffolding, classroom discussion, and independent exercises to address differences in students' learning needs.

The second finding revealed that students exposed to ChatGPT-driven instruction demonstrated higher critical thinking skill scores than students exposed to conventional teaching methods across low, medium, and high IQ levels. The ANCOVA results showed a statistically significant difference between the instructional groups, resulting in the rejection of the second null hypothesis. This finding indicates that ChatGPT-driven instruction can support the development of critical thinking skills among upper basic school students studying Social Studies. The descriptive results showed that the experimental group recorded a total critical thinking mean gain of 17.22, compared with 5.12 for the conventional group. Across the IQ levels, the high-IQ students recorded the highest mean gain of 22.66, followed by the medium-IQ students with 18.68 and the low-IQ students with 10.17. These results demonstrate that the experimental group made greater improvements in critical thinking than the control group at every IQ level, although the magnitude of improvement differed among the groups. The finding may be

explained by the interactive nature of ChatGPT-driven instruction, which can provide opportunities for students to explore alternative explanations, compare viewpoints, question assumptions, and examine different interpretations of historical, cultural, and civic issues. When appropriately guided by teachers, these activities may encourage students to move beyond recalling facts to analysing information, evaluating evidence, and developing reasoned conclusions.

The finding on critical thinking is also relevant to the work of Hao and Tasir, as discussed in the literature review, concerning natural language processing, large language models, and critical thinking development or assessment. The extent of alignment between the present finding and previous studies should be established by comparing their reported results and research contexts. The present study contributes evidence that ChatGPT-driven instruction was associated with significantly higher critical thinking scores than conventional instruction, while also showing that the effect varied across IQ levels. The significant interaction between instructional strategy and IQ level indicates that students' cognitive profiles may influence the extent to which they benefit from AI-supported critical thinking activities. Although the low-IQ students recorded a smaller mean gain than the medium- and high-IQ students, their improvement exceeded that of the corresponding conventional group. This suggests that ChatGPT-driven instruction may also be beneficial to students with lower IQ levels when suitable instructional support is provided. However, the finding should not be interpreted to mean that ChatGPT automatically develops critical thinking. Rather, its educational value may depend on how teachers structure learning activities and encourage students to question AI-generated responses, verify information, explain their reasoning, and arrive at independent conclusions.

Furthermore, the significant interaction effect for critical thinking skills highlights the need for differentiated instructional support when using ChatGPT in Social Studies classrooms. Students with higher IQ levels may be able to engage more readily with complex questions, competing interpretations, and abstract social issues, while students with lower IQ levels may require more structured questions, step-by-step explanations, and guided opportunities to analyse information. Students in the medium-IQ category may also benefit from activities that gradually progress from basic comprehension to interpretation and evaluation. Such differentiated support can help ensure that students do not merely accept AI-generated information but learn to assess its accuracy, relevance, and credibility. In this regard, concerns about excessive or unguided reliance on conversational AI, such as those discussed by Odili and Ebiye in the literature review, reinforce the importance of teacher supervision and the promotion of independent reasoning. ChatGPT should therefore function as a supplementary instructional resource rather than a replacement for teachers' professional judgement or students' intellectual engagement.

Overall, the findings demonstrate that ChatGPT-driven instruction produced significantly better outcomes than conventional instruction in both Social Studies academic achievement and critical thinking skills among upper basic school students in North-East Nigeria. The first finding established that the ChatGPT group recorded higher achievement scores across low, medium, and high IQ levels, while the second showed that the group also recorded higher critical thinking scores across the same IQ categories. Nevertheless, the significant interaction effects indicate that the magnitude of the instructional benefits varied according to IQ level. The findings therefore support the purposeful integration of ChatGPT into Social Studies teaching, accompanied by differentiated instructional strategies, teacher guidance, and activities that promote independent thinking. Teachers should encourage students to evaluate AI-generated information, justify their conclusions, compare competing viewpoints, and solve problems independently. Such practices can help learners benefit from the explanatory and interactive features of ChatGPT while supporting the broader educational objectives of improving academic achievement and developing critical thinking skills.

Conclusion

This study establishes that ChatGPT-driven instruction significantly enhances both student Social Studies achievement and critical thinking skills within upper basic schools in North East Nigeria. However, its effectiveness is deeply contingent upon the learner's individual cognitive profile. While high-IQ students successfully utilize the AI tool as a collaborative partner to expand their analytical reasoning, lower-IQ students primarily leverage it to master foundational content, showing far more modest gains in independent critical thought. Consequently, while generative AI offers a transformative solution to regional educational resource shortages, its deployment must be pedagogically structured to avoid creating cognitive dependency across diverse student ability levels. Based on the findings of this study, it is recommended that: 1. Social Studies teachers in upper basic schools in North-East Nigeria should integrate ChatGPT-driven instruction into their teaching to improve students' academic achievement compared with conventional instructional methods across low, medium, and high IQ levels. 2. Social Studies teachers should use ChatGPT-driven instructional activities that promote critical thinking skills, including analysing information, evaluating evidence, comparing viewpoints, and drawing independent conclusions, while providing appropriate guidance to students across low, medium, and high IQ levels.

REFERENCES

- [1] G. C. Ezeah and C. A. Ozioko, 'Adoption of ChatGPT in academic writing: College students' perspectives, attitudes and behavioral intentions as a basis for literacy programming', *International Journal of Research and Innovation in Social Science*, vol. 8, no. 3, pp. 1410–1422, 2024.
- [2] A. W. Elenode, 'Adoption of generative AI and learning analytics for personalized learning in Nigerian higher education: Effects on student engagement, learning outcomes, and academic integrity', *Engineering and Applied Sciences*, vol. 11, no. 1, pp. 1–5, 2026.
- [3] M. Abdallah, L. Sardi, and Y. Zhou, 'Amplifier or substitute? A systematic review of generative AI's impact on higher-order cognitive skills among university students', *Frontiers in Psychology*, vol. 16, no. 2, pp. 112–128, 2025.
- [4] C. Nwuke and K. Yellowe, 'Artificial intelligence and machine learning applications in secondary school systems: Mitigating methodological deficiencies in West African classrooms', *African Journal of Educational Technology Trends*, vol. 14, no. 2, pp. 89–104, 2024.
- [5] Y. G. Balami, 'Digital interventions in resource-constrained classrooms: Rebuilding educational infrastructure in North East Nigeria', *Journal of Educational Technology and Development*, vol. 20, no. 2, pp. 88–101, 2025.
- [6] E. U. Akpan, I. O. Bassey, and N. M. Udoh, 'Effect of AI-ChatGPT on students' academic performance in physical and health education in universities in Akwa Ibom State, Nigeria', *Journal of Science Education and Technology Trends*, vol. 4, no. 1, pp. 45–56, 2025.
- [7] P. T. Dung, 'Effects of ChatGPT-flipped instruction on the academic performance and self-regulated skill of L2 pre-service teachers in Ondo State, Nigeria', *Hungarian Educational Research Journal*, vol. 15, no. 4, pp. 512–529, 2024.
- [8] World Bank, 'From chalkboards to chatbots: Transforming learning in Nigeria through personalized AI instruction', World Bank Group, World Bank Education Policy Brief, 2024.
- [9] A. B. Gidado and M. K. Zubair, 'Gender influence and technology acceptance of artificial intelligence tools in northern Nigerian secondary education', *Federal University Gusau Faculty of Education Journal*, vol. 5, no. 1, pp. 204–215, 2025.
- [10] A. M. Ali, B. Klimova, and M. Pikhart, 'Integrating generative artificial intelligence in African higher education: University students' awareness, attitudes, and use of ChatGPT in Zambia', *Frontiers in Education*, vol. 9, no. 1, pp. 78–92, 2024.
- [11] Y. Hao and Z. Tasir, 'Natural language processing and large language models in the development and assessment of critical thinking skills: A systematic review', *Journal of Educational Computing Research*, vol. 62, no. 3, pp. 541–563, 2024.
- [12] OECD, *OECD Digital Education Outlook 2026: GenAI, instructional design, and the future of learning*. OECD Publishing, 2026. doi: 10.1787/062a7394-en.
- [13] A. M. Mustapha and Z. B. Yusuf, 'Post-conflict educational recovery and classroom realities in the Lake Chad Basin', *Northern Nigerian Journal of Educational Research*, vol. 12, no. 1, pp. 34–49, 2024.
- [14] C. K. Lo, 'The dark and bright sides of generative AI in humanities education: A comprehensive meta-analysis', *Computers and Education*, vol. 210, pp. 104–119, 2025.
- [15] D. I. Odili and M. Ebiye, 'The impact of artificial intelligence on students across generations in Nigeria', *Kashere Journal of Education*, vol. 6, no. 1, pp. 112–125, 2025.

-
- [16] J. S. Bruner, D. J. Wood, and G. Ross, 'The role of tutoring in problem solving', *Journal of Child Psychology and Psychiatry*, vol. 17, no. 2, pp. 89–100, 1976.
- [17] M. Kindenberg, 'Why think historically when AI can do it? Understanding the relationship between generative artificial intelligence and history teacher identity', *History Education Research Journal*, vol. 21, no. 2, pp. 88–103, 2024.