

The Impact of AI Chatbots on NCE Students' Academic Writings

Chidozie Mmachi

Department of English Language and Literature
Benjamin Uwajumogu State College of Education, Ihitte-Uboma, Imo State, Nigeria
mmachichidozie@gmail.com

and

Amajuoyi Grace Ngozi

Department of Igbo, School of Languages
Benjamin Uwajumogu State College of Education, Ihitte-Uboma, Imo State, Nigeria
ngmajoy@gmail.com

Corresponding author: mmachichidozie@gmail.com; ngmajoy@gmail.com

ABSTRACT

AI chatbots such as ChatGPT have rapidly entered the academic lives of Nigerian students, raising both hopes and concerns about their effect on academic writing. This study examined the impact of AI chatbots on the academic writings of Nigeria Certificate in Education (NCE) students. Four research questions and two null hypotheses guided the study. A descriptive survey design was adopted. The population comprised 5,860 NCE III students in the two public colleges of education in Anambra State, from which a sample of 361 was drawn by proportionate stratified random sampling. Data were collected with a validated 36-item AI Chatbots and Academic Writing Questionnaire (ACAWQ), which yielded a Cronbach's alpha of .85. Mean and standard deviation answered the research questions, while the independent-samples t-test tested the hypotheses at the .05 level. Findings showed that students used AI chatbots to a high extent for academic writing ($M = 2.74$), mostly for grammar correction, idea generation and paraphrasing. Students perceived positive impacts on grammar, organisation, vocabulary and writing speed ($M = 3.06$), but also negative impacts, including over-dependence, plagiarism, fabricated references and difficulty writing without AI ($M = 2.92$). Gender did not significantly influence perceived positive impact, $t(359) = 1.00$, $p = .318$, but students in arts, languages and social sciences perceived greater negative impact than science students, $t(359) = 2.64$, $p = .009$. The study recommends institutional AI-use policies, training in responsible use, disclosure requirements and process-based writing assessment.

Keywords: AI chatbots; ChatGPT; academic writing; academic integrity; NCE students; colleges of education

Introduction

Academic writing is one of the most important skills a student teacher acquires. NCE students write assignments, term papers, examination answers, teaching practice reports and a final research project. Beyond assessment, writing is how teachers prepare lesson notes, communicate with parents and model language use for pupils. The National Policy on Education places English as the main medium of instruction from the upper primary level (Federal Republic of Nigeria, 2013), which makes teachers' own command of written English especially consequential. Colleges of education support this through General Studies courses in the Use of English and communication skills (National Commission for Colleges of Education [NCCE], 2012).

Academic writing is demanding because it is a thinking process, not only a language task. Flower and Hayes (1981) described writing as a set of cognitive processes, planning, translating ideas into text, and reviewing, all directed by the writer's goals. A student who writes an essay must generate ideas, organise them into an argument, find and acknowledge sources, and revise for clarity and accuracy. Many NCE students, most of whom write in English as a second language, find these processes difficult.

Since the public release of ChatGPT in November 2022, AI chatbots have changed how students approach writing. These tools, built on large language models, can generate outlines, draft paragraphs, paraphrase, correct grammar, summarise readings and suggest references within seconds (Kasneci et al., 2023). Free access on smartphones has made them widely available to Nigerian students. For a second-language writer, a chatbot can act as an always-available writing assistant, offering feedback that lecturers handling large classes rarely have time to provide (Su et al., 2023).

The same capabilities raise serious concerns. Students may submit chatbot-generated text as their own, undermining academic integrity (Cotton et al., 2024). Chatbots can produce fluent but inaccurate content and fabricated references (Walters & Wilder, 2023). Heavy reliance may also bypass the very cognitive processes through which writing skill develops, leaving students unable to write independently (Kosmyrna et al., 2025). UNESCO (2023) has therefore urged institutions to develop policies that harness the benefits of generative AI while protecting learning and integrity.

In Nigerian colleges of education, the extent of chatbot use in academic writing, and whether its impact is mainly positive or negative, is not well documented. Most institutions have no explicit policy on AI use, and lecturers respond inconsistently, ranging from tacit acceptance to outright

prohibition. This study was therefore undertaken to provide evidence on the impact of AI chatbots on NCE students' academic writings, using public colleges of education in Anambra State as its setting.

Statement of the Problem

Lecturers in colleges of education have noticed a sharp change in students' written work since 2023. Some assignments are unusually fluent and well structured, yet students who submitted them struggle to explain their content or to write at a similar standard under examination conditions. Lecturers also report reference lists containing sources that cannot be found. At the same time, some students appear to use chatbots constructively to improve grammar, plan essays and understand difficult readings.

These contrasting observations leave colleges without a clear picture. It is not known how widely NCE students use AI chatbots for academic writing, for which tasks, and whether students experience the impact as mainly beneficial or harmful. It is also unclear whether perceptions differ by gender or by field of study, which would affect how guidance should be targeted. Without this evidence, institutions cannot design sound policies or teaching responses. The problem of this study, therefore, is: what is the impact of AI chatbots on NCE students' academic writings?

Purpose of the Study

The main purpose of the study was to examine the impact of AI chatbots on NCE students' academic writings. Specifically, the study sought to determine:

- the extent to which NCE students use AI chatbots for academic writing tasks;
- the positive impacts of AI chatbots on NCE students' academic writings;
- the negative impacts of AI chatbots on NCE students' academic writings; and
- strategies for promoting responsible use of AI chatbots in NCE students' academic writing.

Research Questions

- To what extent do NCE students use AI chatbots for academic writing tasks?
- What are the positive impacts of AI chatbots on NCE students' academic writings?
- What are the negative impacts of AI chatbots on NCE students' academic writings?
- What strategies can promote responsible use of AI chatbots in NCE students' academic writing?

Hypotheses

The following null hypotheses were tested at the .05 level of significance:

H_{01} : There is no significant difference between the mean ratings of male and female NCE students on the positive impacts of AI chatbots on their academic writings.

H_{02} : There is no significant difference between the mean ratings of NCE students in science and those in arts, languages and social sciences on the negative impacts of AI chatbots on their academic writings.

Significance of the Study

The findings will help college managements and the NCCE formulate policies on AI use in coursework and assessment. Lecturers, particularly those teaching the Use of English and research methods, will gain insight for redesigning writing instruction and assessment. Students

will become more aware of both the benefits and the risks of chatbot use, and researchers will have baseline data for further studies on AI and writing in Nigerian teacher education.

Review of Related Literature

Conceptual Review

AI chatbots. AI chatbots are conversational applications built on large language models that generate human-like text in response to prompts. Popular examples include ChatGPT, Gemini, Claude, Copilot and Meta AI, the last of which is embedded in messaging applications widely used by Nigerian students. Kasneci et al. (2023) note that these tools can support learners with explanation, feedback and drafting, but also stress that their output can be inaccurate and that their benefits depend on critical use.

Academic writing. Academic writing is formal, evidence-based writing produced for scholarly purposes, characterised by a clear thesis, logical organisation, appropriate register, acknowledgement of sources and adherence to referencing conventions. For second-language writers, it poses challenges of grammar, vocabulary and rhetorical convention in addition to the demands of content (Hyland, 2019).

Impact of AI chatbots on academic writing. In this study, impact refers to the effects, positive and negative, that students perceive chatbot use to have on their writing process, the quality of their written products, their writing skills and their academic integrity. Positive impacts may include improved grammar, organisation and confidence; negative impacts may include over-dependence, plagiarism, factual errors and fabricated references (Cotton et al., 2024; Dwivedi et al., 2023).

Theoretical Framework

The study is anchored on the Cognitive Process Theory of Writing by Flower and Hayes (1981). The theory holds that writing involves three recursive processes: planning (generating and organising ideas and setting goals), translating (turning ideas into written language) and reviewing (evaluating and revising text). These processes are monitored by the writer and shaped by the task environment and the writer's long-term memory. The theory is relevant because chatbots can intervene in every process: they can generate ideas during planning, produce sentences during translating and correct errors during reviewing. Whether the intervention helps or harms depends on whether the student remains in control of these processes or hands them over entirely. Used as a support, a chatbot can scaffold weak processes; used as a substitute, it can prevent the student from developing them.

The study is also informed by the Technology Acceptance Model (Davis, 1989), which predicts that students adopt a technology when they perceive it as useful and easy to use. The model explains the rapid uptake of chatbots among students, independent of institutional approval.

Empirical Review

Su et al. (2023) explored the use of ChatGPT in argumentative writing classrooms and found that it could support outlining, content preparation and revision, but cautioned that students needed guidance to evaluate its suggestions and keep ownership of their arguments. Dwivedi et al. (2023), in a multidisciplinary review, identified productivity and accessibility benefits of generative AI for writing, alongside risks to originality, accuracy and the development of skills.

On integrity, Cotton et al. (2024) argued that chatbots make it difficult to detect unauthorised assistance and recommended that institutions redesign assessment rather than rely on detection alone. Liang et al. (2023) showed that AI-text detectors frequently misclassified writing by non-

native English speakers as AI-generated, a finding with direct implications for Nigerian students. On accuracy, Walters and Wilder (2023) found that a substantial proportion of references generated by ChatGPT were fabricated, and many genuine references contained errors.

On cognitive effects, Kosmyna et al. (2025) compared essay writers who used a chatbot, a search engine or no tool, and reported that chatbot users showed lower brain connectivity during writing, weaker recall of their own essays and lower sense of ownership. Although preliminary, the study supports concerns that unreflective reliance may undermine learning.

Research on chatbots and writing in Nigerian colleges of education remains scarce. Available Nigerian studies on technology and writing have mostly addressed word processing, online plagiarism and internet use, rather than generative AI.

Summary of Literature Review

The literature suggests that AI chatbots can support the planning, drafting and revising of academic writing, especially for second-language writers, but pose real risks to integrity, accuracy and skill development. The Cognitive Process Theory of Writing explains why the same tool can help or harm depending on how it is used. Evidence from Nigerian colleges of education is lacking, a gap this study fills.

Method

The study adopted a descriptive survey design, which allows the researcher to collect data from a representative sample in order to describe the opinions, practices and experiences of a population as they exist. The study was carried out in Anambra State, South-East Nigeria, in the two public colleges of education: Federal College of Education (Technical), Umunze, and Nwafor Orizu College of Education, Nsugbe. Both institutions run NCE programmes across the sciences, vocational and technical education, languages, arts and social sciences. The population comprised 5,860 NCE III students in the two colleges in the 2025/2026 academic session. NCE III students were chosen because they had completed at least two years of coursework, including written assignments, and were writing their final research projects.

A sample of 361 students was determined using the Krejcie and Morgan (1970) table. Proportionate stratified random sampling was used, with college, gender and school (sciences versus arts, languages and social sciences) as strata. The sample comprised 148 male and 213 female students, of whom 154 were in the sciences and 207 in arts, languages and social sciences.

Data were collected with a researcher-developed instrument titled the AI Chatbots and Academic Writing Questionnaire (ACAWQ). Section A captured gender and school. Section B contained 36 items in four clusters: extent of use (8 items), positive impacts (10 items), negative impacts (10 items) and strategies (8 items). Cluster A used a four-point scale of Very High Extent (4), High Extent (3), Low Extent (2) and Very Low Extent (1). Clusters B to D used Strongly Agree (4), Agree (3), Disagree (2) and Strongly Disagree (1).

The ACAWQ was face- and content-validated by three experts: one in English language education, one in educational technology and one in measurement and evaluation. Their suggestions on wording and item coverage were incorporated. The instrument was trial-tested on 30 NCE III students in a college of education in Enugu State. Cronbach's alpha (Cronbach, 1951) yielded coefficients of .82, .87, .84 and .80 for Clusters A to D, and .85 for the whole instrument.

The questionnaire was administered in person by the researchers with the help of four trained research assistants. Of the 361 copies distributed, 361 were retrieved and found usable, a 100%

return rate. Participation was voluntary, and respondents were assured that responses were anonymous and would not affect their grades, which was important given the sensitivity of questions on academic integrity.

Mean and standard deviation were used to answer the research questions, with a criterion mean of 2.50: item means of 2.50 and above were interpreted as High Extent or Agree, and those below 2.50 as Low Extent or Disagree. The independent-samples t-test was used to test the hypotheses at the .05 level of significance.

Results

Research Question 1: Extent of use of AI chatbots for academic writing tasks

Table 1. Mean ratings on the extent of use of AI chatbots for academic writing (N = 361)

S/N	Writing task	Mean	SD	Decision
1	Generating ideas and outlines for assignments	3.21	0.78	High Extent
2	Paraphrasing or rewriting text	3.08	0.82	High Extent
3	Correcting grammar and spelling	3.32	0.74	High Extent
4	Writing complete assignments or term papers	2.41	0.97	Low Extent
5	Summarising reading materials and handouts	2.97	0.85	High Extent
6	Generating references and citations	2.36	0.94	Low Extent
7	Drafting chapters of the research project	2.18	0.96	Low Extent
8	Translating ideas from mother tongue into English	2.39	0.99	Low Extent
	Grand mean	2.74	0.88	High Extent

Table 1 shows that students used AI chatbots to a high extent for academic writing (grand mean = 2.74). Grammar and spelling correction ($M = 3.32$), idea generation ($M = 3.21$), paraphrasing ($M = 3.08$) and summarising ($M = 2.97$) were the most common uses. Writing complete assignments ($M = 2.41$) and drafting project chapters ($M = 2.18$) were reported to a low extent, although their relatively large standard deviations show that a sizeable minority did so.

S/N	Positive Impact	Mean	SD	Decision
1	Improves my grammar and writing mechanics	3.38	0.69	Agree
2	Helps me organise my ideas coherently	3.14	0.76	Agree
3	Expands my academic vocabulary	3.09	0.78	Agree
4	Helps me overcome difficulty in starting to write	3.22	0.74	Agree
5	Saves time in completing assignments	3.41	0.70	Agree
6	Improves my understanding of the topics I write on	2.96	0.83	Agree

7	Improves the academic tone and style of my writing	2.94	0.81	Agree
8	Increases my confidence in writing	3.02	0.80	Agree
9	Helps me understand referencing styles such as APA	2.71	0.88	Agree
10	Improves my ability to build arguments	2.73	0.86	Agree
	Grand mean	3.06	0.79	Agree

Research Question 2: Positive impacts of AI chatbots on academic writings

Table 2. Mean ratings on positive impacts of AI chatbots on academic writings (N = 361)

S/N	Negative Impact	Mean	SD	Decision
1	Over-dependence reduces my own writing effort	3.19	0.77	Agree
2	Students submit AI text they do not fully understand	2.88	0.86	Agree
3	Chatbots produce references that do not exist	2.97	0.84	Agree
4	Chatbots give inaccurate information	2.84	0.85	Agree
5	Use of chatbots encourages plagiarism and dishonesty	3.05	0.82	Agree
6	Reliance on chatbots weakens critical thinking	2.91	0.86	Agree
7	AI writing is generic and lacks local context and personal voice	2.78	0.87	Agree
8	Students struggle to write without AI in examinations	3.12	0.79	Agree
9	Students are penalised for work that looks AI-generated	2.46	0.93	Disagree
10	Reliance on chatbots reduces reading of original sources	3.00	0.83	Agree
	Grand mean	2.92	0.84	Agree

Table 2 shows agreement with all ten positive impacts (grand mean = 3.06). Saving time ($M = 3.41$) and improving grammar ($M = 3.38$) were rated highest, while help with referencing ($M = 2.71$) and argument building ($M = 2.73$) were rated lowest.

S/N	Strategy	Mean	SD	Decision
1	Each college should issue a clear policy on AI use in coursework	3.34	0.70	Agree

2	Students should be trained in responsible use and effective prompting	3.46	0.65	Agree
3	Students should disclose how they used AI in each assignment	3.32	0.72	Agree
4	More writing tasks should be done in class or by hand	3.28	0.74	Agree
5	Lecturers should teach students to verify AI-generated references	3.41	0.67	Agree
6	Assessment should reward drafts and the writing process	3.29	0.73	Agree
7	AI literacy should be integrated into the Use of English course	3.53	0.62	Agree
8	Similarity checks should be combined with oral defence of written work	3.33	0.71	Agree
Grand mean		3.37	0.69	Agree

Research Question 3: Negative impacts of AI chatbots on academic writings

Table 3. Mean ratings on negative impacts of AI chatbots on academic writings (N = 361)

Gender	N	Mean	SD	df	t	p	Decision
Male	148	3.03	0.48	359	1.00	.318	Not rejected
Female	213	3.08	0.46				

Table 3 shows agreement with nine of the ten negative impacts (grand mean = 2.92). Over-dependence (M = 3.19), difficulty writing without AI in examinations (M = 3.12) and encouragement of plagiarism (M = 3.05) ranked highest. Respondents disagreed that students are penalised for AI-like work (M = 2.46), suggesting weak enforcement or absence of clear rules.

School	N	Mean	SD	df	t	p	Decision
Sciences	154	2.84	0.51	359	2.64	.009	Rejected
Arts, languages and social sciences	207	2.98	0.49				

Research Question 4: Strategies for promoting responsible use

Table 4. Mean ratings on strategies for promoting responsible use of AI chatbots (N = 361)

Table 4 shows strong agreement with all eight strategies (grand mean = 3.37), led by integrating AI literacy into the Use of English course (M = 3.53) and training students in responsible use (M = 3.46).

Hypothesis 1

Table 5. t-test of male and female students' mean ratings on positive impacts

Table 5 shows that $t(359) = 1.00$, $p = .318$. Since $p > .05$, H_{01} was not rejected. Male and female students did not differ significantly in their perception of the positive impacts of AI chatbots.

Hypothesis 2

Table 6. t-test of science and arts/languages/social sciences students' mean ratings on negative impacts

Table 6 shows that $t(359) = 2.64$, $p = .009$. Since $p < .05$, H_{02} was rejected. Students in arts, languages and social sciences perceived significantly greater negative impacts than science students.

Discussion of Findings

Extent of use. NCE students used AI chatbots to a high extent in academic writing, mainly as assistants for grammar correction, idea generation, paraphrasing and summarising. This rapid adoption, without formal institutional encouragement, is consistent with the Technology Acceptance Model (Davis, 1989): chatbots are free, available on phones and immediately useful. The lower but noticeable use of chatbots to write whole assignments and project chapters is a concern. Self-reports on such sensitive behaviour tend to understate it, so the true extent may be higher.

Positive impacts. Students agreed that chatbots improved their grammar, organisation, vocabulary, confidence and speed. In terms of Flower and Hayes's (1981) model, chatbots appear to support the reviewing process (correcting language) and the early planning process (overcoming the blank page and organising ideas). This mirrors Su et al. (2023), who found ChatGPT useful for outlining and revision. For second-language writers in large classes, such immediate feedback fills a real gap. The lower ratings for argument building and referencing suggest that chatbots help most with surface features and least with the deeper thinking and scholarly conventions that define academic writing.

Negative impacts. Students were candid about the downsides. They acknowledged over-dependence, difficulty writing without AI in examinations, plagiarism, fabricated references and reduced reading of original sources. These findings align with Cotton et al. (2024) on integrity, Walters and Wilder (2023) on fabricated references, and Kosmyna et al. (2025) on reduced cognitive engagement. Read through the cognitive process theory, the problem arises when students hand over the translating process, the actual composing of text, to the chatbot. They then produce writing without practising writing, which explains why they struggle in examinations. Students' disagreement that AI-like work is penalised indicates a policy vacuum: in the absence of clear rules, misuse carries little consequence.

Strategies. Students endorsed all the proposed strategies, especially building AI literacy into the Use of English course and training in responsible use. This suggests that students themselves want guidance rather than either unrestricted use or outright bans. The endorsement of disclosure, process-based assessment and oral defence is consistent with Cotton et al.'s (2024) call to redesign assessment. Given Liang et al.'s (2023) evidence that AI detectors are biased against non-native English writers, reliance on detection software alone would be both unreliable and unfair to Nigerian students.

Gender and school. Male and female students perceived the positive impacts similarly, indicating that the benefits are not gender-specific. However, students in arts, languages and

social sciences perceived greater negative impacts than science students. Writing is central to assessment in these disciplines, where essays and extended arguments dominate, so students are more exposed to the temptation of substitution and more aware of its cost to their skills. Science students, whose writing is shorter and more formulaic, may encounter these risks less directly.

Conclusion

AI chatbots have become an established part of NCE students' academic writing. Their impact is double-edged. Used as assistants, they improve grammar, organisation, vocabulary and confidence, and help students start and finish their work. Used as substitutes, they encourage over-dependence, plagiarism and fabricated references, and leave students unable to write independently. The difference lies in whether students remain the authors of their own thinking. In the absence of clear institutional policies, students are navigating this balance largely on their own. Colleges of education must therefore move from silence or prohibition to deliberate guidance.

Recommendations

Each college of education should develop and publicise a clear policy on AI use in coursework, specifying permitted uses, prohibited uses and consequences of misuse.

The NCCE should integrate AI literacy, covering responsible use, effective prompting, verification of output and academic integrity, into the General Studies Use of English curriculum.

Lecturers should require students to disclose how they used AI in each written assignment and should teach them to verify every AI-suggested reference.

Lecturers should adopt process-based assessment, including submission of outlines and drafts, in-class writing and short oral defences of written work, rather than relying on AI-detection software.

Lecturers in arts, languages and social sciences, where writing carries greater weight, should give particular attention to these measures.

College managements, with support from TETFund, should organise workshops for lecturers on adapting writing instruction and assessment to the presence of AI chatbots.

Limitations and Suggestions for Further Studies

The study relied on students' self-reports, which may understate sensitive behaviours such as submitting AI-written work, and it was limited to two colleges in one state. Further studies should analyse samples of students' actual written work, include lecturers' perspectives, extend to other states and to universities, and use experimental designs to compare guided and unguided chatbot use on writing quality.

References

- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy.

- International Journal of Information Management, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Federal Republic of Nigeria. (2013). National policy on education (6th ed.). Nigerian Educational Research and Development Council.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387. <https://doi.org/10.2307/356600>
- Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task (arXiv:2506.08872). arXiv. <https://doi.org/10.48550/arXiv.2506.08872>
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607-610. <https://doi.org/10.1177/001316447003000308>
- Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), 100779. <https://doi.org/10.1016/j.patter.2023.100779>
- National Commission for Colleges of Education. (2012). *Nigeria Certificate in Education minimum standards for general education*. NCCE.
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. <https://doi.org/10.1016/j.asw.2023.100752>
- UNESCO. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization. <https://doi.org/10.54675/EWZM9535>
- Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports*, 13, 14045. <https://doi.org/10.1038/s41598-023-41032-5>