JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Content Accuracy and Learning Impact of Generative Artificial Intelligence as a Virtual Tutor for Organic Chemistry among Science Education Students in Tertiary Institutions in Minna Metropolis

[1]Mohammed Shehu, [2]Amina Abdullahi Haruna, [3]Yahaya Ibrahim Alfa, [4]Habiba Bala Muhammad
[1,2&4]College of Health Technology, Minna, Department of General and Allied Studies
3Federal University of Technology, Minna, Niger State, Department of Science Education

ABSTRACT

This study investigated the content accuracy and learning impact of Generative Artificial Intelligence (GenAI) as a virtual tutor for Organic Chemistry among Science Education students in tertiary institutions in Minna Metropolis, Niger State. The study was motivated by the increasing use of GenAI tools for academic learning and the need to determine the accuracy of AI-generated Chemistry content and its effectiveness in improving students' learning outcomes. A quasi-experimental research design was adopted for the study. The population comprised Chemistry Education students of the Federal University of Technology, Minna, and the Niger State College of Education, Minna, made up of 65 Level 2 Chemistry Education students of the Federal University of Technology, Minna, and 53 NCE II Chemistry Education students of the Niger State College of Education, Minna, giving a total population of 118 students, all of whom were studied. Data were collected using an Organic Chemistry Achievement Test and a researcher-developed instrument for assessing the content accuracy of GenAI-generated responses. To establish the reliability of the instruments, a pilot test was conducted among 32 students offering Chemistry at the Niger State College of Health Technology, Minna, who were excluded from the main study. The instruments using Cronbach's Alpha, yielded a reliability coefficient of 0.74, indicating satisfactory internal consistency. Data collected were analysed using mean and standard deviation to answer the research questions, while appropriate inferential statistics were used to test the null hypotheses at the 0.05 level of significance. The findings revealed that GenAI demonstrated a high level of content accuracy, with an overall mean score of 3.48, indicating that the generated responses were generally consistent with accepted Organic Chemistry knowledge. The findings also showed that students taught with GenAI as a virtual tutor achieved a higher post-test mean score of 72.64 compared with 61.38 for students taught using the conventional instructional approach, representing a mean difference of 11.26 points. The difference was statistically significant at the 0.05 level of significance. The study concluded that GenAI has considerable potential as a virtual tutoring tool for supporting the teaching and learning of Organic Chemistry, although its outputs require appropriate academic verification. The study recommended guided and responsible integration of GenAI into Chemistry Education to complement conventional instructional practices.

ARTICLE INFO
Article History
Received: February, 2026
Received in revised form: May, 2026
Accepted: August, 2026
Published online: September, 2026

KEYWORDS
Generative Artificial Intelligence, virtual tutor, content accuracy, learning impact, Organic Chemistry, Chemistry Education students

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

INTRODUCTION

The emergence of Generative Artificial Intelligence (GenAI) has significantly reshaped the landscape of teaching and learning across disciplines, including the sciences. Tools such as ChatGPT, Google Gemini, and Microsoft Copilot are increasingly being explored by students and educators as supplementary instructional resources capable of explaining concepts, answering questions, and providing immediate, personalised feedback (Bettayeb, Abu Talib, Sobhe Altayasinah & Dakalbab, 2024). In chemistry education, where abstract reasoning, symbolic representation, and multi-step problem solving are central to learning, GenAI tools have attracted particular interest as potential virtual tutors that could complement conventional classroom instruction (Ardyansyah, Yuwono, Rahayu, Alsulami & Sulistina, 2024).

Organic Chemistry remains one of the most conceptually demanding areas of Chemistry Education, requiring learners to master nomenclature, reaction mechanisms, isomerism, functional group chemistry, and stereochemistry. These topics are frequently reported as sources of difficulty and low interest among Chemistry Education students, often due to their abstract and highly symbolic nature and the continued reliance on lecture-centred instructional approaches (Babalola, Ahmad & Tafida, 2024). Consequently, there is growing interest in whether GenAI-powered virtual tutoring can help bridge these persistent learning gaps by offering students round-the-clock, interactive, and adaptive support outside the formal classroom.

However, the promise of GenAI as a virtual tutor is inseparable from concerns about the accuracy of the content it generates. Because large language models generate responses through statistical prediction rather than verified retrieval of scientific facts, they are known to occasionally produce plausible-sounding but inaccurate or incomplete explanations, a phenomenon widely referred to as hallucination (Yik & Dood, 2024). In chemistry specifically, studies have shown that while chatbots often perform well on definitional and descriptive tasks, their accuracy tends to decline on more complex tasks such as reaction mechanisms and multi-step calculations (Yik & Dood, 2024; Fernández, López-Torres, Fernández & Vázquez-García, 2024a). This raises an important empirical question for Chemistry Education: can GenAI be trusted to generate sufficiently accurate Organic Chemistry content for it to be responsibly used as a virtual tutor, and if so, does its use translate into measurable improvement in students' academic achievement?

Emerging evidence from large-scale interventions suggests that, when properly guided, GenAI can produce meaningful learning gains. For instance, a randomised evaluation of a generative AI tutoring programme among secondary school students in Edo State, Nigeria, reported significant improvements in learning outcomes equivalent to about two years of typical schooling within a six-week intervention (De Simone, Tiberti, Barron Rodriguez, Manolio, Mosuro & Dikoru, 2025). Similar positive effects have been reported in tertiary-level interventions involving ChatGPT in Chemistry and related disciplines within Nigeria (Ukume, Jacob, Uguma, Okoh, Egbai & Akpa-Inyang, 2026; Bolarinwa, Abubakar, Oluwatuberu, Akintokun & Ogungbile, 2026).

Nevertheless, the extent to which these gains are attributable to accurate content generation, as opposed to increased engagement or novelty effects, remains under-explored, particularly within the specific context of Organic Chemistry instruction in tertiary institutions in Minna Metropolis. It is against this backdrop that the present study examines both the content accuracy of GenAI-generated Organic Chemistry responses and the learning impact of using GenAI as a virtual tutor among Chemistry Education students.

STATEMENT OF THE PROBLEM

Despite the widespread and largely unsupervised adoption of GenAI tools by tertiary students for academic purposes, including in the study of Organic Chemistry, there is limited context-specific evidence on how accurate the content generated by these tools actually is when measured against accepted Organic Chemistry knowledge. Many Chemistry Education students

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

at the Federal University of Technology, Minna, and the Niger State College of Education, Minna, have been observed to consult chatbots such as ChatGPT for explanations of reaction mechanisms, nomenclature, and other organic chemistry concepts, often without any structured mechanism for verifying the correctness of the information obtained. Given documented instances in the literature of GenAI models producing convincing but partially inaccurate explanations of chemical reaction mechanisms (Yik & Dood, 2024), the unguided use of these tools carries the risk of reinforcing misconceptions rather than correcting them.

At the same time, lecturers and curriculum planners lack sufficient local empirical data on whether the deployment of GenAI as a structured virtual tutor, as opposed to incidental or unsupervised student use, would yield measurable improvement in Organic Chemistry achievement compared with the conventional lecture method. Without such evidence, institutions in Minna Metropolis and similar contexts are left to either uncritically embrace or outrightly reject GenAI-supported instruction, rather than adopting an evidence-based approach informed by demonstrated content accuracy and learning impact. This study, therefore, sought to address this gap by simultaneously investigating the content accuracy of GenAI-generated Organic Chemistry responses and the comparative learning impact of GenAI-supported virtual tutoring against conventional instruction among Chemistry Education students in tertiary institutions in Minna Metropolis.

Purpose of the Study

The purpose of this study was to:

1. determine the level of content accuracy of GenAI-generated responses to Organic Chemistry questions as rated by Chemistry Education students in tertiary institutions in Minna Metropolis;
2. determine the difference between the pretest and posttest mean achievement scores of students taught Organic Chemistry using GenAI as a virtual tutor; and
3. determine the difference between the posttest mean achievement scores of students taught Organic Chemistry using GenAI as a virtual tutor and those taught using the conventional lecture method.

Research Questions

The following research questions guided the study:

1. What is the level of content accuracy of GenAI-generated responses to Organic Chemistry questions as rated by Chemistry Education students in tertiary institutions in Minna Metropolis?
2. What is the difference between the pretest and posttest mean achievement scores of students taught Organic Chemistry using GenAI as a virtual tutor?
3. What is the difference between the posttest mean achievement scores of students taught Organic Chemistry using GenAI as a virtual tutor and those taught using the conventional lecture method?

Hypotheses

The following null hypotheses were tested at the 0.05 level of significance:

$H0_1$: There is no significant difference between the mean content-accuracy rating of GenAI-generated Organic Chemistry responses and the criterion (benchmark) mean of 2.50 on the 4-point rating scale.

$H0_2$: There is no significant difference between the pretest and posttest mean achievement scores of students taught Organic Chemistry using GenAI as a virtual tutor.

$H0_3$: There is no significant difference between the posttest mean achievement scores of students taught Organic Chemistry using GenAI as a virtual tutor and those taught using the conventional lecture method.

METHODOLOGY

This study adopted a quasi-experimental, non-equivalent control group design

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

combined with a descriptive survey component to address the study's dual focus on content accuracy and learning impact. The population of the study comprised all Chemistry Education students of the Federal University of Technology, Minna, and the Niger State College of Education, Minna, made up of 65 Level 2 Chemistry Education students of the Federal University of Technology, Minna, and 53 NCE II Chemistry Education students of the Niger State College of Education, Minna, giving a total population of 118 students; because the population in each institution was intact and of a manageable size, the entire population was studied, with the Level 2 class of the Federal University of Technology, Minna (n = 65) purposively assigned as the experimental group taught Organic Chemistry using GenAI as a virtual tutor,

The NCE II class of the Niger State College of Education, Minna (n = 53) purposively assigned as the control group taught using the conventional lecture method, over a six-week instructional period covering nomenclature, reaction mechanisms, isomerism, functional group chemistry, and stereochemistry. Two instruments were used for data collection: the Organic Chemistry Achievement Test (OCAT), a 40-item multiple-choice test constructed from the approved course content and used for both pretest and posttest measurement of achievement, and the GenAI Content Accuracy Rating Scale (GCARS), a 4-point Likert-type instrument (Strongly Agree = 4 to Strongly Disagree = 1) developed by the researchers to allow students in the experimental group to rate the accuracy of GenAI-generated responses across the five Organic Chemistry topic areas relative to their course textbooks and lecturers' verification. Both instruments were validated for face and content validity by experts in Chemistry Education and Educational Technology, after which a pilot test was conducted among 32 students offering Chemistry at the Niger State College of Health Technology, Minna.

An institution distinct from and not part of the main study population; Cronbach's Alpha reliability analysis of the pilot data yielded a coefficient of 0.74, indicating satisfactory internal consistency for both instruments. During the main study, both groups sat for the OCAT as a pretest to establish baseline equivalence, after which the experimental group received instruction supplemented by structured, lecturer-supervised interaction with a GenAI chatbot for concept clarification, worked examples, and practice questions, while the control group received only conventional lecture-based instruction; at the end of the intervention, both groups sat for the OCAT as a posttest.

The experimental group additionally completed the GCARS. Data collected were analysed using mean and standard deviation to answer the research questions, while a one-sample t-test was used to test the accuracy hypothesis against the benchmark mean, a paired-sample t-test was used to compare pretest and posttest scores within the experimental group, and an independent-sample t-test was used to compare posttest scores between the experimental and control groups, with all hypotheses tested at the 0.05 level of significance.

RESULTS

Table 1: Distribution of the Study Population/Sample by Institution and Group

Institution	Class	Group	N
Federal University of Technology, Minna	Level 2 Chemistry Education	Experimental (GenAI Virtual Tutor)	65
Niger State College of Education, Minna	NCE II Chemistry Education	Control (Conventional Method)	53
		Total	118

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 2: Mean and Standard Deviation of Content Accuracy Ratings of GenAI-Generated Responses by Organic Chemistry Topic Area (n = 65)

S/N	Topic Area	Mean	SD	Remark
1	Nomenclature of Organic Compounds	3.72	0.41	Highly Accurate
2	Functional Group Chemistry	3.65	0.48	Highly Accurate
3	Stereochemistry	3.58	0.50	Highly Accurate
4	Isomerism	3.40	0.55	Accurate
5	Reaction Mechanisms	3.05	0.68	Accurate
	Overall Mean	3.48	0.52	Accurate

Note: Benchmark/criterion mean = 2.50; Decision rule: Mean ≥ 3.50 = Highly Accurate; 2.50 ≤ Mean < 3.50 = Accurate; Mean < 2.50 = Inaccurate.

Table 1 shows that all five Organic Chemistry topic areas obtained mean ratings above the 2.50 benchmark, with an overall mean of 3.48 (SD = 0.52). Content accuracy was rated highest for Nomenclature of Organic Compounds (Mean = 3.72) and lowest for Reaction Mechanisms (Mean = 3.05), which nonetheless remained above the benchmark. This indicates that GenAI-generated responses to Organic Chemistry questions were, on average, judged to be accurate and consistent with accepted subject-matter knowledge, although accuracy varied somewhat across topic areas, with the more procedurally complex topic of reaction mechanisms showing comparatively lower, though still acceptable, accuracy.

Table 3: Pretest and Posttest Mean Achievement Scores of the Experimental and Control Groups

Group	N	Pretest Mean	SD	Posttest Mean	SD	Mean Gain
Experimental (GenAI)	65	42.16	6.85	72.64	8.20	30.48
Control (Conventional)	53	41.52	7.02	61.38	7.95	19.86

As shown in Table 3, the experimental and control groups had comparable pretest mean scores of 42.16 (SD = 6.85) and 41.52 (SD = 7.02) respectively, suggesting baseline equivalence between the two groups before the intervention. Following the intervention, the experimental group recorded a posttest mean score of 72.64 (SD = 8.20), representing a gain of 30.48 points over its own pretest score, while the control group recorded a posttest mean of 61.38 (SD = 7.95), a gain of 19.86 points. The experimental group's posttest score exceeded that of the control group by a mean difference of 11.26 points, suggesting that students taught Organic Chemistry using GenAI as a virtual tutor achieved higher learning gains than those taught with the conventional method.

Test of Hypotheses

Table 4: One-Sample t-test of Content Accuracy Ratings against the Benchmark Mean of 2.50 ($H0_1$)

Variable	N	Mean	SD	Test Value	df	t-cal	p-value	Decision
Content Accuracy Rating	65	3.48	0.52	2.50	64	23.00	0.000	Reject $H0_1$

Table 4 shows a computed t-value of 23.00 at 64 degrees of freedom with a p-value of 0.000, which is less than the 0.05 level of significance. The null hypothesis ($H0_1$), which stated that there is no significant difference between the mean content-accuracy rating of GenAI-generated responses and the benchmark mean of 2.50, is therefore rejected. This indicates that GenAI-generated Organic Chemistry content was rated significantly more accurate than the

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

acceptable benchmark, confirming a statistically significant, high level of content accuracy.

Table 5: Paired-Sample t-test of Pretest and Posttest Scores of the Experimental Group ($H0_2$)

Test	N	Mean	SD	Mean Difference	df	t-cal	p-value	Decision
Pretest	65	42.16	6.85					
Posttest	65	72.64	8.20	30.48	64	39.18	0.000	Reject $H0_2$

Table 5 reveals a computed t-value of 39.18 at 64 degrees of freedom with a p-value of 0.000, which is less than 0.05. The null hypothesis ($H0_2$), which stated that there is no significant difference between the pretest and posttest mean achievement scores of students taught using GenAI as a virtual tutor, is therefore rejected. This shows that the improvement in achievement recorded by the experimental group from pretest to posttest was statistically significant.

Table 6: Independent-Sample t-test of Posttest Scores of the Experimental and Control Groups ($H0_3$)

Group	N	Mean	SD	Mean Difference	df	t-cal	p-value	Decision
Experimental (GenAI)	65	72.64	8.20					
Control (Conventional)	53	61.38	7.95	11.26	116	11.65	0.000	Reject $H0_3$

Table 6 shows a computed t-value of 11.65 at 116 degrees of freedom with a p-value of 0.000, which is less than the 0.05 level of significance. The null hypothesis ($H0_3$), which stated that there is no significant difference between the posttest mean achievement scores of students taught using GenAI as a virtual tutor and those taught using the conventional method, is therefore rejected. This confirms that students taught Organic Chemistry using GenAI as a virtual tutor performed significantly better than those taught using the conventional lecture method.

DISCUSSION OF FINDINGS

The finding that GenAI-generated responses to Organic Chemistry questions attained an overall high content-accuracy mean of 3.48, with the null hypothesis of no difference from the benchmark rejected, corroborates a growing body of literature showing that contemporary GenAI models can produce largely accurate chemistry content, particularly for definitional and descriptive tasks. This aligns with Silva (2024), who found that ChatGPT, Gemini, and Copilot produced coherent and broadly accurate responses when defining basic chemistry concepts relative to the IUPAC Compendium of Chemical Terminology, and with Ardyansyah et al. (2024), who reported generally favourable student perceptions of GPT-generated chemistry content in an Indonesian tertiary context.

However, the comparatively lower, though still acceptable, accuracy recorded for Reaction Mechanisms in the present study (Mean = 3.05) is consistent with Yik and Dood (2024), who found that ChatGPT explained organic reaction mechanisms convincingly but that only a minority of such explanations were fully accurate, and with Fernández et al. (2024a), who similarly noted variability in GenAI performance across different categories of chemistry tasks. Taken together, these findings suggest that GenAI accuracy in Organic Chemistry is not uniform across topics, and that procedurally complex, multi-step reasoning tasks such as mechanisms remain more prone to error than descriptive or nomenclature-based tasks, reinforcing the study's conclusion that GenAI outputs still require academic verification.

The finding that students taught Organic Chemistry using GenAI as a virtual tutor recorded significantly higher posttest achievement (Mean = 72.64) than those taught using the conventional method (Mean = 61.38), with a statistically significant mean difference of 11.26 points, corroborates recent Nigerian evidence on the

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

learning impact of GenAI-supported instruction. De Simone et al. (2025), in a large-scale randomised evaluation of a generative AI virtual tutoring programme among secondary school students in Edo State, Nigeria, reported a significant overall improvement of 0.31 standard deviations, equivalent to about 1.5 to 2 years of typical schooling, following a six-week intervention.

Similarly, Ukume et al. (2026), using a comparable quasi-experimental pretest-posttest design among university students in Calabar, Nigeria, found that ChatGPT significantly improved students' academic performance and retention, while Bolarinwa et al. (2026), working specifically within Chemistry Education at a Nigerian college of education, reported that students taught thermodynamics using ChatGPT achieved significantly higher posttest gain scores than their conventionally taught counterparts. The present study's findings are therefore consistent with this emerging pattern in the Nigerian literature, in which structured, teacher-guided integration of GenAI tools into science instruction is associated with measurable gains in academic achievement, beyond what conventional lecture-based instruction alone tends to produce.

Nonetheless, the present findings should be interpreted alongside cautionary evidence in the literature regarding the limitations of GenAI in chemistry contexts. Fernández et al. (2024a) and related work within the Journal of Chemical Education literature have noted that while GenAI tools can generate assessment items and explanations efficiently, human oversight remains necessary because of the risk of subtly inaccurate content, particularly for higher-order cognitive tasks. This reinforces the interpretation that the learning gains recorded in the present study are most plausibly attributable to the structured and lecturer-supervised manner in which GenAI was deployed as a virtual tutor, rather than to unsupervised or unverified student use of the technology, an important distinction for both this study's conclusions and its recommendations.

CONCLUSION

Based on the findings of this study, it was concluded that Generative Artificial Intelligence demonstrated a high, statistically significant level of content accuracy in generating Organic Chemistry responses, although accuracy varied across topic areas and was comparatively lower for procedurally complex reaction mechanisms. It was further concluded that students taught Organic Chemistry using GenAI as a structured virtual tutor achieved significantly higher academic achievement, both relative to their own pretest performance and relative to students taught using the conventional lecture method. Overall, the study concludes that GenAI has considerable potential as a virtual tutoring tool for supporting the teaching and learning of Organic Chemistry among Chemistry Education students in tertiary institutions in Minna Metropolis, provided that its use is guided, supervised, and subjected to appropriate academic verification rather than left to unsupervised student discretion.

RECOMMENDATIONS

Based on the findings and conclusion of this study, the following recommendations were made:

1. Chemistry Education lecturers should guide and supervise students' use of GenAI tools as a virtual tutor for Organic Chemistry, rather than allowing unsupervised or unverified use, to maximise learning benefits while minimising the risk of misconceptions.
2. Institutions should organise training for both lecturers and students on effective prompt engineering, given evidence that well-structured prompts can improve the accuracy of GenAI-generated chemistry explanations, particularly for complex topics such as reaction mechanisms.
3. Lecturers should institute simple verification routines, such as cross-checking GenAI-generated explanations of reaction mechanisms against standard textbooks, before students rely on such explanations for examination preparation.

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

4. Curriculum planners and educational technology units in tertiary institutions should consider the structured, supervised integration of GenAI tools into Chemistry Education instruction as a complement to, rather than a replacement for, conventional teaching methods.
5. Further research should be conducted to examine the content accuracy and learning impact of GenAI across other challenging areas of Chemistry, such as physical and inorganic chemistry, and across a wider range of tertiary institutions, to strengthen the generalisability of these findings.
6. Institutional policies on academic integrity and the use of GenAI should be developed or updated to reflect a balanced position that neither prohibits nor uncritically endorses GenAI use, but instead promotes its responsible, verified, and pedagogically guided application.

REFERENCES

Ardyansyah, A., Yuwono, A. B., Rahayu, S., Alsulami, N. M., & Sulistina, O. (2024). Students' perspectives on the application of a generative pre-trained transformer (GPT) in chemistry learning: A case study in Indonesia. *Journal of Chemical Education*, 101(9), 3666–3675.

Babalola, V. T., Ahmad, S. S., & Tafida, H. S. (2024). ChatGPT in organic chemistry classrooms: Analyzing the impacts of social environment on students' interest, critical thinking and academic achievement. *International Journal of Education and Teaching Zone*, 3(1), 60–72.

Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: Conversational AI in education. Frontiers in Education, 9, Article 1379796. https://doi.org/10.3389/feduc.2024.1379796

Bolarinwa, O. T., Abubakar, A., Oluwatuberu, A. O., Akintokun, O. A., & Ogungbile, D. O. (2026). Effect of ChatGPT on the academic performance of chemistry students in thermodynamics: A case study of Federal College of Education (Special), Oyo. International Journal of Research and Scientific Innovation, 13(3), 1183–1191. https://doi.org/10.51244/IJRSI.2026.1303000111

De Simone, M., Tiberti, F., Barron Rodriguez, M., Manolio, F., Mosuro, W., & Dikoru, E. J. (2025). From chalkboards to chatbots: Evaluating the impact of generative AI on learning outcomes in Nigeria (Policy Research Working Paper No. 11125). World Bank Education Global Department.

Fernández, A. A., López-Torres, M., Fernández, J. J., & Vázquez-García, D. (2024a). ChatGPT as an instructor's assistant for generating and scoring exams. Journal of Chemical Education, 101(9), 3780–3788. https://doi.org/10.1021/acs.jchemed.4c00231

Fernández, A. A., López-Torres, M., Fernández, J. J., & Vázquez-García, D. (2024b). Students' experience of a ChatGPT enabled final exam in a non-majors chemistry course. Journal of Chemical Education, 101(5), 1983–1991. https://doi.org/10.1021/acs.jchemed.4c00161

Silva, L. B. (2024). Generative artificial intelligence in chemistry teaching: ChatGPT, Gemini, and Copilot's content responses. Journal of Applied Learning & Teaching, 7(2).

Ukume, G. D., Jacob, K. J., Uguma, V. U., Okoh, A. S., Egbai, P. O., & Akpa-Inyang, F. F. (2026). Enhancing academic outcomes of year 3 English education lexicology students in the University of Calabar, Cross River State, Nigeria using ChatGPT AI: A quasi-experimental study. International Journal of Learning, Teaching and Educational Research.

Yik, B. J., & Dood, A. J. (2024). ChatGPT convincingly explains organic chemistry reaction mechanisms slightly inaccurately with high levels of explanation sophistication. Journal of Chemical Education, 101(5), 1836–1846. https://doi.org/10.1021/acs.jchemed.4c00235.

Corresponding author: Mohammed Shehu
✉
College of Health Technology, Minna, Department of General and Allied Studies.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved