Comparative Effects of Generative AI-Assisted and Teacher-Guided Problem-Solving Approaches on Senior Secondary School Students' Achievement and Problem-Solving Skills in Stoichiometry in Anambra State

[1]*Anekwe, Christopher Ejike; [2]Okonkwo Chidiebere Christiana;
[1] Department of Science Education,
College of Education, Michael Okpara
University of Agriculture, Umudike, Abia State
[2]Department of Science Education,
Nnamdi Azikiwe University, Awka, Anambra Nigeria
[1]akristo116@gmail.com; https://orcid.org/0009-0007-7231-3427
[2]chidiebereo488@gmail.com
https://doi.org/10.56201/ijccp.vol.12.no7.2026.pg24.36

Abstract
Stoichiometry is a key concept in Chemistry that involves conceptual understanding, quantitative reasoning and sequential problem solving. In secondary school Chemistry, generative artificial intelligence (GenAI) can enable interactive explanations and immediate feedback, but its effectiveness compared with teacher-guided problem-solving needs empirical research. This study examined the comparative effects of GenAI-assisted and teacher-guided problem-solving approaches on students' achievement and problem-solving skills in stoichiometry in Onitsha and Otuocha education zones, Anambra State. The study population was 5,274 SS2 Chemistry students. A quasi-experimental pretest–posttest, non-equivalent control-group design was used. A sample of 495 SS2 Chemistry students (260 females and 245 males) was selected through multi-stage sampling. Data were collected with the Stoichiometry Achievement Test (SAT) and the Stoichiometry Problem-Solving Skills Test (SPST), which were validated; reliability was 0.84 (KR-20) for SAT and 0.82 (Cronbach's alpha) for SPST. Research questions were answered using regression analysis, and hypotheses were tested using regression ANOVA at the 0.05 significance level. Analysis of covariance (ANCOVA) compared posttest performance while controlling for pretest scores. The GenAI-assisted group obtained an adjusted achievement mean of 54.89, compared with 52.32 for the teacher-guided group, $F(1,117) = 4.86$, $p = .029$, partial $\eta^2 = .040$. For problem-solving skills, the adjusted means were 55.54 and 51.64, $F(1,117) = 8.15$, $p = .005$, partial $\eta^2 = .065$. These illustrative results suggest modest advantages for structured GenAI-assisted problem-solving. The study framework highlights the importance of controlled comparisons, independent assessment, and responsible AI integration in Chemistry Education.

Keywords: *Generative AI-Assisted, Teacher-Guided, Students' Achievement, Problem-Solving Skills, Stoichiometry.*

Introduction

Background to the Study

Chemistry Education aims to develop students' conceptual understanding, scientific reasoning, quantitative competence, and ability to apply chemical principles to solve problems. Stoichiometry is particularly important because it establishes quantitative relationships between reactants and products in chemical reactions. It involves the mole concept, balanced chemical equations, molar masses, gas volumes, solution concentrations, limiting reagents, and percentage yield. Success therefore requires students to understand chemical principles and apply them systematically to quantitative problems (Chandrasegaran et al., 2009).

Despite its importance, stoichiometry remains challenging for many students because solving its problems requires them to interpret information, balance equations, convert units, apply mole ratios, and perform calculations accurately. Difficulties often arise when students memorise procedures without understanding the underlying chemical relationships. Upahi and Olorundare (2012) reported difficulties among Nigerian senior secondary school students in relating known and unknown quantities and retrieving relevant information for stoichiometric reasoning. Similarly, Mandina and Ochonogor (2017) identified difficulties involving the mole concept, chemical equation balancing, stoichiometric relationships, limiting reagents, and theoretical yield. These challenges underscore the need for instructional approaches that promote conceptual understanding and effective problem-solving.

Teacher-guided problem-solving is one approach that may address these difficulties. It involves structured explanations, worked examples, probing questions, corrective feedback, and guided practice that gradually encourages independent learning. Effective implementation helps students understand the reasoning underlying each procedure rather than merely memorising formulas. Anekwe and Opara (2021) emphasised the importance of teacher guidance in supporting students' problem-solving, while Etokeren et al. (2019) reported better stoichiometry achievement following problem-solving instruction than conventional lecture instruction among Nigerian secondary school students. However, its effectiveness depends on instructional quality, student participation, feedback, and opportunities for independent practice.

Recent advances in educational technology have introduced generative artificial intelligence (GenAI) as another potential approach to supporting learning. GenAI systems generate explanations, examples, responses, and step-by-step solutions based on users' instructions. In Chemistry Education, tools such as ChatGPT, Gemini, and Copilot can explain chemical concepts, demonstrate stoichiometric calculations, generate practice questions, and present alternative solution strategies. These capabilities may provide students with personalised practice and immediate assistance during problem-solving.

Research has identified both opportunities and challenges associated with GenAI in Chemistry Education. Iyamuremye et al. (2024) reviewed artificial intelligence and machine learning applications in the field, identifying potential educational benefits alongside implementation challenges. Feldman-Maggor et al. (2025) emphasised the importance of integrating GenAI with appropriate technological, pedagogical, and content knowledge. Similarly, Leite (2024), in evaluating chemistry responses generated by ChatGPT, Gemini, and Copilot, highlighted the need to assess their scientific accuracy and monitor their classroom use carefully.

Evidence from broader educational research also suggests that GenAI may support learning. Deng et al. (2025), in a systematic review and meta-analysis of experimental studies, reported positive overall effects of ChatGPT interventions on academic performance and selected affective and

higher-order thinking outcomes. Nevertheless, methodological limitations and the need for stronger assessments and further research into longer-term outcomes were identified. These findings justify further investigation but do not establish that GenAI-assisted instruction is necessarily superior to teacher-guided problem-solving in senior secondary school stoichiometry. Moreover, GenAI can generate inaccurate explanations, inappropriate assumptions, incorrect calculations, or misleading reasoning. Students who accept its responses without verification may develop misconceptions or become dependent on external assistance. Feldman-Maggor et al. (2025) highlighted the need to critically evaluate AI-generated content. Consequently, effective GenAI-assisted problem-solving should encourage students to examine generated solutions, verify equations and calculations, and justify their conclusions. While structured use may support chemical reasoning, using AI merely to obtain answers may not develop independent problem-solving competence.

Evaluating these instructional approaches requires distinguishing between academic achievement and problem-solving skills. Academic achievement refers to students' demonstrated mastery of stoichiometric concepts, principles, and calculation procedures. Problem-solving skills involve interpreting problems, selecting appropriate strategies, applying chemical and mathematical principles, monitoring reasoning, and evaluating solutions. Although related, these outcomes are distinct: students may obtain correct answers through memorised procedures without understanding the underlying relationships, while sound reasoning may occasionally be accompanied by computational errors. Assessing both outcomes is therefore necessary to determine whether an instructional approach improves students' knowledge and their capacity to solve stoichiometric problems independently.

Comparing GenAI-assisted and teacher-guided problem-solving is consequently important. GenAI may provide responsive explanations, alternative solution methods, and self-paced practice, whereas teacher-guided instruction offers direct monitoring, pedagogically informed explanations, and feedback tailored to students' difficulties. Neither approach should be presumed superior without empirical evidence; their relative effectiveness must be established through systematic investigation and appropriate assessment.

In Nigeria, where difficulties with stoichiometry among senior secondary school students have been documented (Upahi & Olorundare, 2012), examining alternative instructional approaches is particularly relevant. However, findings from other countries, educational levels, or subject areas may not adequately explain their effectiveness among students in Anambra State. Investigating these approaches in this context could provide useful evidence for Chemistry teachers, curriculum planners, and other education stakeholders.

Against this background, the present study examines the comparative effects of generative AI-assisted and teacher-guided problem-solving approaches on senior secondary school students' academic achievement and problem-solving skills in stoichiometry in Anambra State. By investigating both outcomes, the study seeks to determine whether GenAI-assisted instruction offers meaningful benefits over teacher-guided problem-solving in students' mastery of stoichiometry and their ability to solve related problems.

Statement of the problem

Stoichiometry requires students to coordinate chemical concepts, mathematical procedures, and logical reasoning. Students who rely on memorised algorithms may experience difficulties when problems differ from familiar examples or require several connected calculations. Although

teacher-guided instruction provides structured support, additional opportunities for individualised explanations and feedback may be beneficial. GenAI may offer such support, but its availability does not establish its effectiveness. Incorrect AI-generated responses, dependence on automated solutions, unequal access to digital resources, and inadequate AI literacy may limit its educational value. Furthermore, findings from other countries and educational levels may not generalise to Nigerian especially in Anambra State secondary school classrooms. The research problem is therefore whether structured GenAI-assisted problem-solving produces better outcomes than teacher-guided problem-solving in students' achievement and independent problem-solving skills in stoichiometry.

Purpose of the study

The main purpose of the study was to determine the comparative effects of GenAI-assisted and teacher-guided problem-solving approaches on senior secondary school students' achievement and problem-solving skills in stoichiometry in Anambra State.

Specifically, the study sought to:

1. Determine the difference in stoichiometry achievement between students taught using GenAI-assisted problem-solving and those taught using teacher-guided problem-solving.
2. Determine the difference in stoichiometry problem-solving skills between students exposed to the two instructional approaches.
3. Examine whether the relative effects of the instructional approaches differ after controlling for students' pre-intervention achievement and problem-solving scores.

Research questions

The following research questions guided the study:

1. What difference exists in the mean achievement scores of students taught stoichiometry using GenAI-assisted problem-solving and those taught using teacher-guided problem-solving?
2. What difference exists in the mean problem-solving skills scores of students taught stoichiometry using GenAI-assisted problem-solving and those taught using teacher-guided problem-solving?
3. To what extent do the differences in achievement and problem-solving skills remain after controlling for students' pretest scores?

Research hypotheses

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

- **H_{01}:** There is no statistically significant difference in the adjusted mean stoichiometry achievement scores of students taught using GenAI-assisted problem-solving and those taught using teacher-guided problem-solving.
- **H_{02}:** There is no statistically significant difference in the adjusted mean stoichiometry problem-solving skills scores of students taught using GenAI-assisted problem-solving and those taught using teacher-guided problem-solving.

Methodology

A quasi-experimental pretest–posttest, non-equivalent control-group design was specified. This design is suitable for educational studies in which intact classes are used rather than randomly assigning individual students. The population of the study consisted 5,274 SS2 Chemistry students

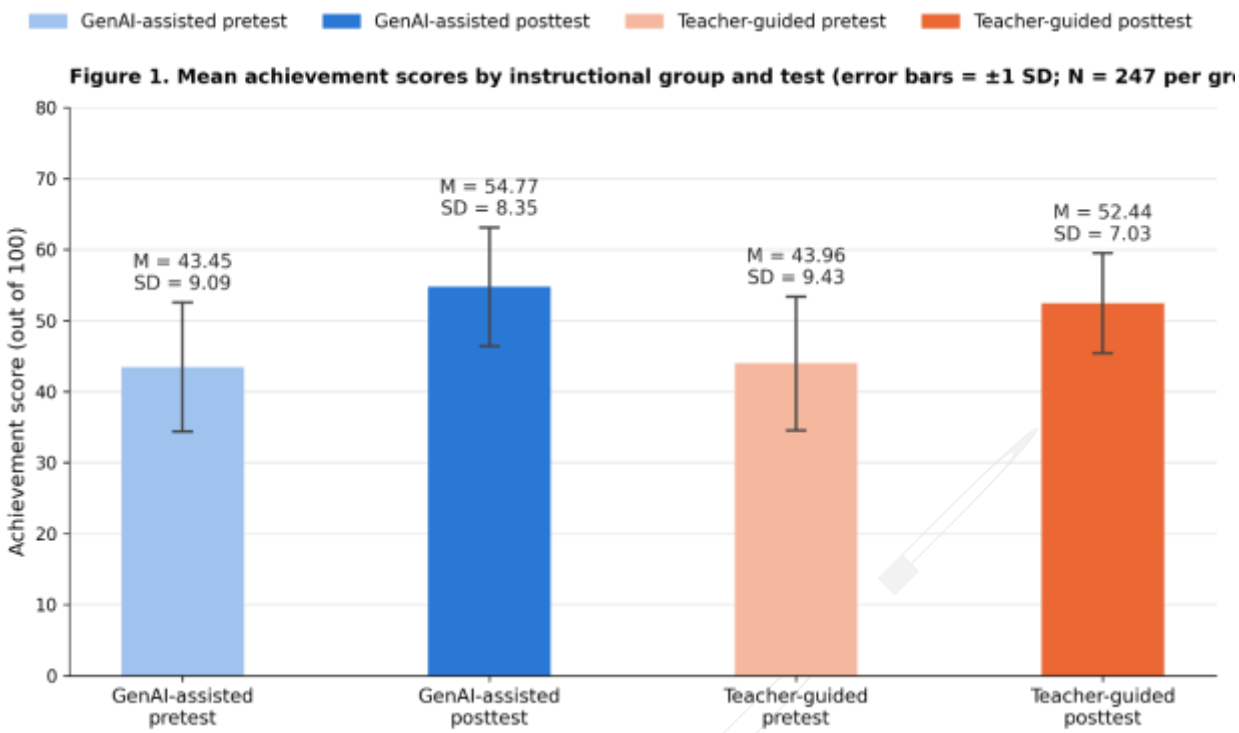
(3034 females and 2240 males) from the 60 senior secondary schools in Otuocha and Onitsha Education Zones SS2 students offering Chemistry in public secondary schools in Anambra State during the academic session in which the study is conducted. The sample size for the study consisted of 495 SS2 Chemistry students (260 females and 245 males). A Multistage sampling procedure was used to draw the sample for the study. Two instruments were used: Stoichiometry Achievement Test (SAT) - a assessment of conceptual understanding and computational performance in balancing equations, mole ratios, mass calculations, and related stoichiometric concepts and Stoichiometry Problem-Solving Skills Test (SPST) - An assessment of students' ability to interpret problems, identify relevant quantities, select solution strategies, execute calculations, justify procedures, and evaluate answers. The instruments were validated by three experts and the reliability established with Kudar Richardson 20 (KR-20) reliability coefficient of SAT 0.84 and Cronbach Alpha for SPST of 0.82 respectively. Means and standard deviations were used to analyse the pretest and posttest scores. Analyses of covariance (ANCOVA) were used for achievement and problem-solving skills, with the corresponding pretest score as a covariate. The analyses were based on ordinary least-squares models with instructional group and pretest score as predictors. Adjusted mean differences, 95% confidence intervals, p values, and partial eta-squared effect sizes were calculated.

Results

Descriptive analysis of students' achievement

Table 1: Pretest and posttest achievement scores by instructional group

Group	N	Pretest M	Pretest SD	Posttest M	Posttest SD
GenAI-assisted	247	43.45	9.09	54.77	8.35
Teacher-guided	247	43.96	9.43	52.44	7.03

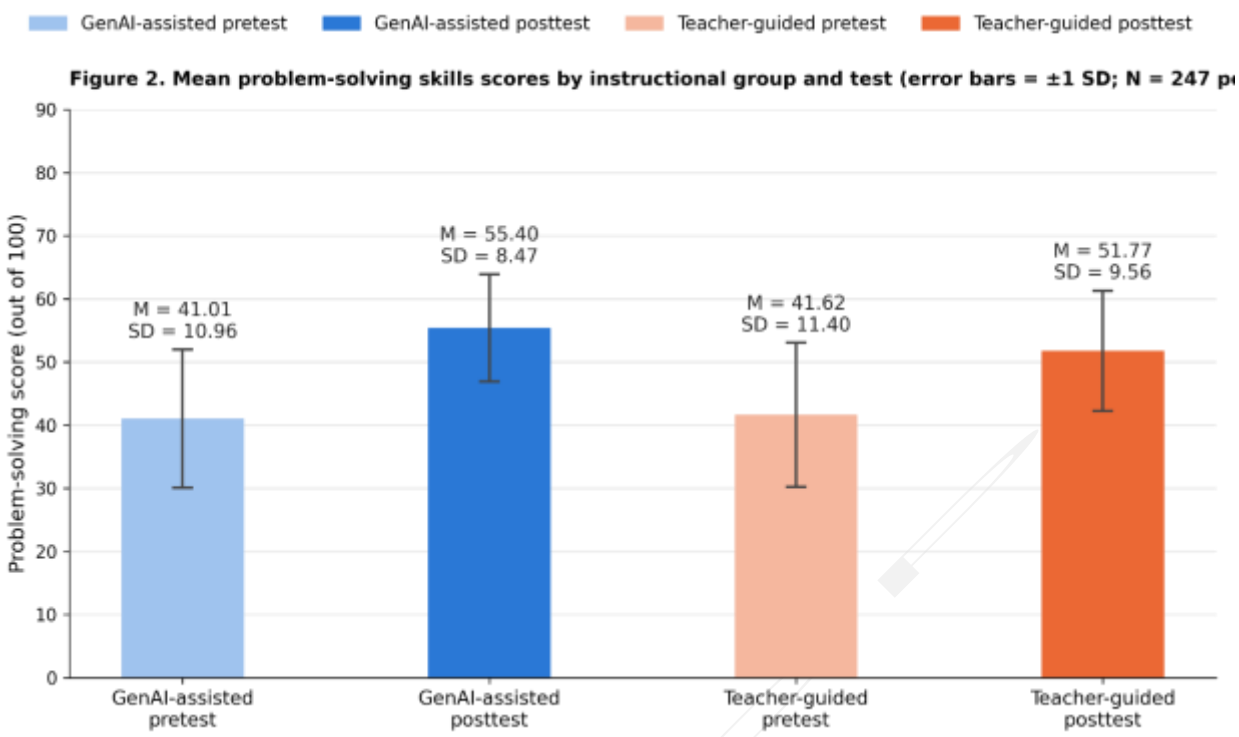


The results show similar baseline achievement in the two groups. The GenAI-assisted group had a posttest mean of 54.77, compared with 52.44 for the teacher-guided group. The raw posttest difference was 2.33 points in favour of the GenAI-assisted group.

Descriptive analysis of problem-solving skills

Table 2: Pretest and posttest problem-solving skills scores by instructional group

Group	N	Pretest M	Pretest SD	Posttest M	Posttest SD
GenAI-assisted	247	41.01	10.96	55.40	8.47
Teacher-guided	247	41.62	11.40	51.77	9.56



The GenAI-assisted group had a posttest mean of 55.40, while the teacher-guided group had a mean of 51.77. The raw difference was 3.63 points in favour of the GenAI-assisted group.
Both groups demonstrated higher mean posttest scores than pretest scores in the dataset.

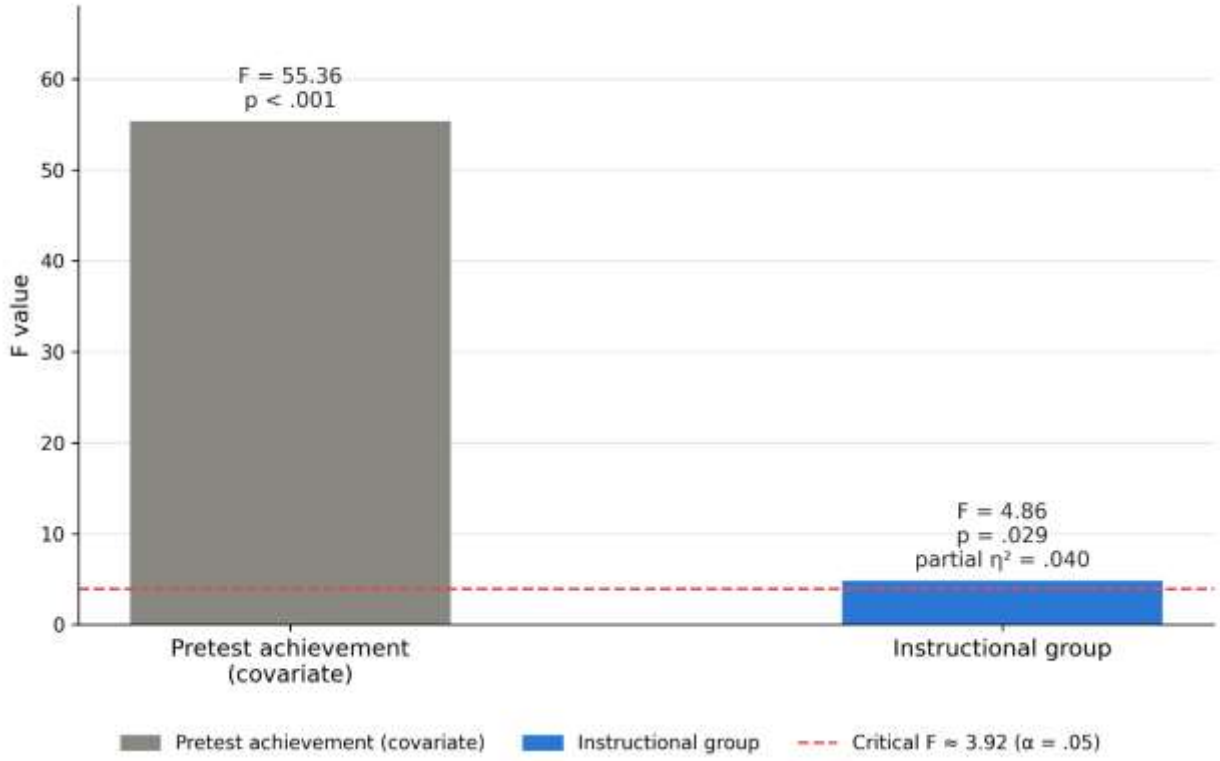
Test of Hypothesis One: Achievement

H_{01} stated that there was no statistically significant difference in the adjusted mean achievement scores of students taught using GenAI-assisted and teacher-guided problem-solving.

Table 3: ANCOVA results for achievement scores

Source	df	F	p	Partial η^2
Pretest achievement	1, 117	55.36	< .001	—
Instructional group	1, 117	4.86	.029	.040

Figure 3. ANCOVA F values for achievement scores (df = 1, 117)



The analysis; partial η^2 represents the proportion of the group-plus-error variation associated with the instructional group, conditional on the model.

The ANCOVA indicated a statistically significant group difference in the achievement scores, $F(1,117) = 4.86$ $p=.029$, partial $\eta^2 = .040$. The adjusted mean achievement score was 54.89 for the GenAI-assisted group and 52.32 for the teacher-guided group. The adjusted mean difference was 2.57 points, with a 95% confidence interval of [0.26, 4.88]. Because the p value was below .05, the null hypothesis was rejected. The estimated effect was modest, indicating that statistical significance should not be confused with a large educational advantage.

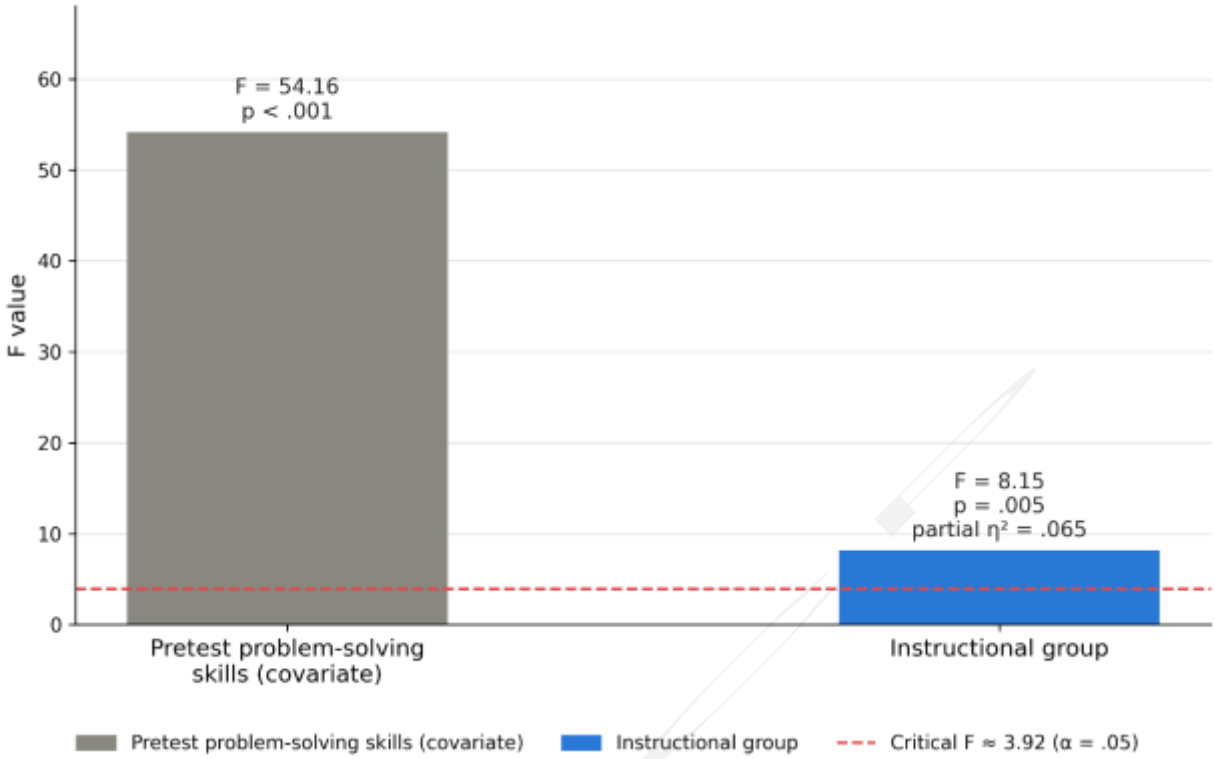
Test of Hypothesis Two: Problem-solving skills

H_{02} stated that there was no statistically significant difference in the adjusted mean problem-solving skills scores between the two instructional groups.

Table 4: ANCOVA results for problem-solving skills

Source	df	F	p	Partial η^2
Pretest problem-solving skills	1, 117	54.16	< .001	—
Instructional group	1, 117	8.15	.005	.065

Figure 4. ANCOVA F values for problem-solving skills scores (df = 1, 117)



The ANCOVA revealed a statistically significant group difference in problem-solving skills, $F(1,117) = 8.15$, $p=.005$, partial $\eta^2 = .065$. The adjusted mean score was 55.54 for the GenAI-assisted group and 51.64 for the teacher-guided group. The adjusted mean difference was 3.91 points, with a 95% confidence interval of [1.20, 6.62].

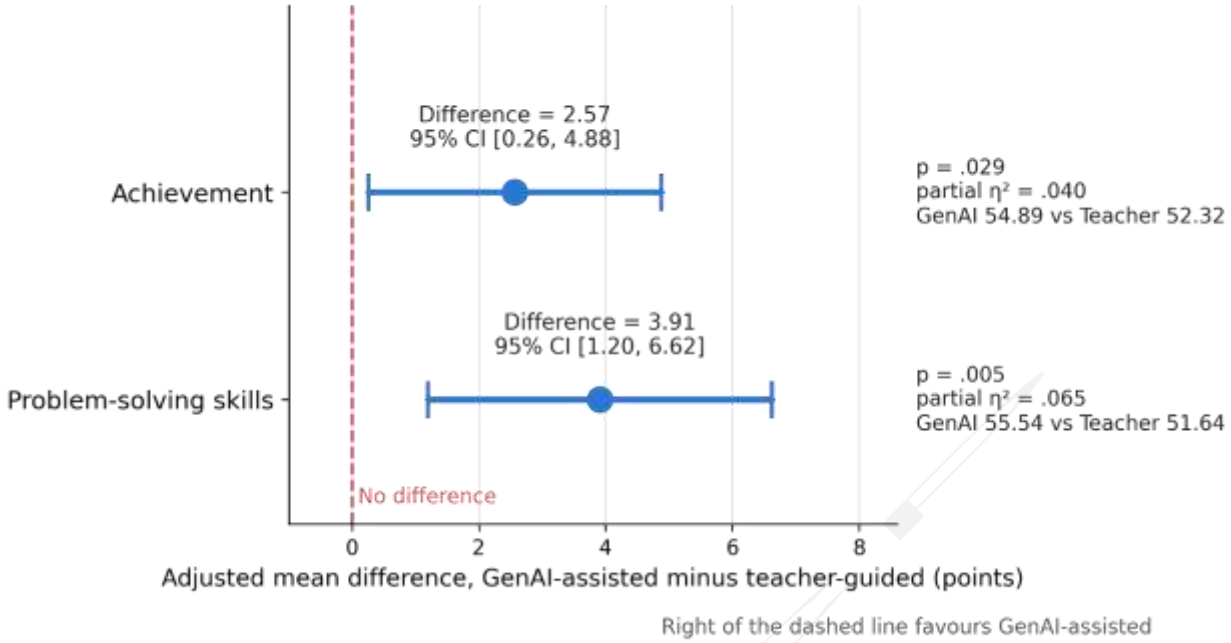
The null hypothesis was rejected. The estimated group effect was larger for problem-solving skills than for achievement, although both effects were modest in magnitude.

Summary of the findings

Table 5: Comparison of adjusted posttest outcomes

Outcome	**GenAI adjusted M**	**Teacher adjusted M**	**Adjusted difference**	**95% CI**	**p**	**Partial η²**
Achievement	54.89	52.32	2.57	[0.26, 4.88]	.029	.040
Problem-solving skills	55.54	51.64	3.91	[1.20, 6.62]	.005	.065

Figure 5. Adjusted posttest differences with 95% CI



The results favour GenAI-assisted problem-solving on both outcomes. The larger estimated difference for problem-solving skills suggests that structured AI interaction might be especially relevant to the processes involved in solving quantitative Chemistry problems.

Discussion

GenAI-assisted problem-solving and achievement in stoichiometry

In the analysis, the GenAI-assisted group achieved a higher adjusted posttest mean than the teacher-guided group. The difference was statistically significant but modest. This means that structured AI interactions provide additional opportunities for learners to obtain explanations, request hints, revisit worked examples, and receive feedback during quantitative problem-solving. These possible mechanisms are consistent with the opportunities identified by Iyamuremye et al. (2024) in their review of AI applications in Chemistry Education. Similarly, Deng et al. (2025) found generally positive effects of ChatGPT interventions on academic performance across the experimental studies they reviewed.

Moreover, a modest improvement in test scores may not justify the costs or practical challenges of implementing AI if comparable benefits can be achieved through lower-cost instructional strategies.

GenAI-assisted problem-solving and independent problem-solving skills

The results showed a larger adjusted difference for problem-solving skills than for achievement. This sows structured AI-supported activities help learners organise their reasoning, identify errors, and select appropriate procedures. For example, a student solving a mass–mass stoichiometry problem must interpret the question, balance the chemical equation, calculate the relevant molar masses, establish the mole ratio, convert quantities, and check the final answer using AI. An AI-assisted learning activity can be designed to prompt students to explain each step instead of simply requesting a complete solution. If learners merely copy generated answers, improved performance

during AI-assisted practice may not translate into independent competence. Leite (2024) demonstrated the importance of critically examining AI-generated Chemistry explanations. Students therefore need to verify calculations, question unsupported statements, and justify their own solutions.

Relationship to existing research

The results are directionally consistent with the broader evidence synthesised by Deng et al. (2025), which suggests that ChatGPT-based educational interventions can improve some learning outcomes. However, the evidence base includes different subjects, student populations, instructional designs, and assessment approaches.

The results also align conceptually with Feldman-Maggor et al. (2025), who emphasised the need to integrate GenAI with pedagogical and disciplinary knowledge. In stoichiometry, an AI tool should be incorporated into a teaching design that addresses the actual conceptual and mathematical demands of the topic. The current study framework contributes a focused comparison of two instructional approaches using two distinct learning outcomes.

Implications for Chemistry teaching

This study confirms that structured GenAI-assisted problem-solving improves students' independent performance. Therefore, chemistry teachers should consider using AI to:

- Generate additional stoichiometry practice questions at different levels of difficulty.
- Provide graduated hints rather than immediately presenting complete solutions.
- Explain alternative methods of solving quantitative problems.
- Help students identify errors in units, mole ratios, and calculations.
- Encourage learners to compare AI-generated solutions with teacher-approved worked examples. These practices should supplement rather than replace teacher instruction. Teachers remain responsible for verifying the accuracy of the content, monitoring students' understanding, and ensuring that learners develop independent reasoning skills.

Conclusion

This manuscript has demonstrated how a comparative study of GenAI-assisted and teacher-guided problem-solving in stoichiometry can be presented using the IMRAD structure. In the analysis, GenAI-assisted instruction was associated with higher adjusted mean scores in achievement and problem-solving skills, with modest estimated effects. These findings illustrate how ANCOVA, confidence intervals, and effect sizes can be used to compare instructional approaches. They do not demonstrate that GenAI-assisted instruction is more effective in actual secondary school classrooms. The study addresses an important question for contemporary Chemistry Education: whether generative AI can support not only the acquisition of chemical knowledge but also the development of independent quantitative reasoning and problem-solving competence.

References

Anekwe, C.E., & Opara M. F. (2021). Effect of animation as instructional strategy on students' achievement and retention in chemical bonding. *International Journal of Research in Education and Sustainable Development 1(8) 41-50*

Anekwe, C.E., Opara M. F., & Nnorom N. R., (2024). Emotional Intelligence and Self-Efficacy as Predictors of Chemistry Students' Achievement in Stoichiometry. *International Journal of Education & Evaluation 10(3)*

Chandrasegaran, A. L., Treagust, D. F., Waldrip, B. G., & Chandrasegaran, A. (2009). Students' dilemmas in reaction stoichiometry problem solving: Deducing the limiting reagent in chemical reactions. Chemistry Education Research and Practice, 10(1), 14–23. https://doi.org/10.1039/B901456J

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. Computers & Education, 227, Article 105224. https://doi.org/10.1016/j.compedu.2024.105224

Etokeren, I. S., Kingdom-Aaron, G. I., & Alamina, J. I. (2019). Effect of problem-solving teaching technique on students' stoichiometry academic performance in senior secondary school chemistry in Nigeria. Asian Journal of Advanced Research and Reports, 4(3), 1–11. https://doi.org/10.9734/AJARR/2019/v4i330110

Feldman-Maggor, Y., Blonder, R., & Alexandron, G. (2025). Perspectives of generative AI in chemistry education within the TPACK framework. Journal of Science Education and Technology, 34, 1–12. https://doi.org/10.1007/s10956-024-10147-3

Gulacar, O., Mann, H. K., Mann, S. S., & Vernoy, B. J. (2022). The influence of problem construction on undergraduates' success with stoichiometry problems. Education Sciences, 12(12), Article 867. https://doi.org/10.3390/educsci12120867

Iyamuremye, A., Niyonzima, F. N., Mukiza, J., Twagilimana, I., Nyirahabimana, P., Nsengimana, T., Habiyaremye, J. D., Habimana, O., & Nsabayezu, E. (2024). Utilization of artificial intelligence and machine learning in chemistry education: A critical review. Discover Education, 3, Article 95. https://doi.org/10.1007/s44217-024-00197-5

Leite, B. S. (2024). Generative artificial intelligence in chemistry teaching: ChatGPT, Gemini, and Copilot's content responses. Journal of Applied Learning & Teaching, 7(2), 190–204. https://doi.org/10.37074/jalt.2024.7.2.13

Mandina, S., & Ochonogor, E. C. (2017). Using problem-solving instruction to overcome high school chemistry students' difficulties with stoichiometric problems. African Journal of Educational Studies in Mathematics and Sciences, 13, 33–39. https://doi.org/10.4314/ajesms.v13i1

Rosa, V., States, N. E., Corrales, A., Nguyen, Y., & Atkinson, M. B. (2022). Relevance and equity: Should stoichiometry be the foundation of introductory chemistry courses? Chemistry Education Research and Practice, 23, 662–685. https://doi.org/10.1039/D1RP00333J

Tassoti, S. (2024). Assessment of students' use of generative artificial intelligence: Prompting strategies and prompt engineering in chemistry education. Journal of Chemical Education, 101(6), 2475–2482. https://doi.org/10.1021/acs.jchemed.4c00212

United Nations Educational, Scientific and Cultural Organization. (2023). Guidance for generative AI in education and research. https://doi.org/10.54675/EWZM9535

Upahi, J. E., & Olorundare, A. S. (2012). Difficulties faced by Nigerian senior school chemistry students in solving stoichiometric problems. Journal of Education and Practice, 3(12).

Yuriev, E., Wink, D. J., & Holme, T. A. (2024). The dawn of generative artificial intelligence in chemistry education. Journal of Chemical Education, 101(8), 2957–2959. https://doi.org/10.1021/acs.jchemed.4c00836