Jurnal Pendidikan MIPA
Volume 16. Nomor 4, Tahun 2026 | ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

The Impact of Artificial Intelligence (AI) in Stoichiometry Learning on Problem-Solving Skills of Senior High School Students in Phase F

Sabila Zeolita[1),*], Rayandra Asyhar[1)], Asrial[1)], Muhammad Naswir[1)], Yusnidar[1)], Asmiyunda[1)]

[1)]Pendidikan Kimia, Universitas Jambi

*Coresponding Author: zeolitasabila@gmail.com

ABSTRACT

This research was motivated by students' low problem-solving skills in stoichiometry and the increasing integration of Artificial Intelligence (AI) in learning. Chemistry learning, particularly stoichiometry, still presents difficulties in understanding concepts and determining systematic problem-solving steps. Therefore, this study aimed to analyze the effect of AI use in stoichiometry learning on the problem-solving skills of high school students in phase F. This study employed a quantitative method with a true experimental design using a pretest-posttest control group design. The sample consisted of an experimental group (XI-F4), which received AI, and a control group (XI-F5), which learned without AI. Data were collected through essay tests (pretest and posttest) and a student response questionnaire. The instruments were tested beforehand to ensure validity and reliability. Data analysis included normality, homogeneity, Independent Sample t-test, and N-Gain tests using SPSS Statistics version 29 and Microsoft Excel 2025. The results showed no significant effect of AI on students' problem-solving skills (0.152 > 0.05), although N-Gain and student responses indicated AI's potential as a learning support tool.

Keywords: Artificial Intelligence; Stoichiometry; Problem-Solving Skills

INTRODUCTION

The rapid advancement of technology has transformed educational practices and increased the demand for learning approaches that foster students' 21st-century competencies, particularly critical thinking, creativity, collaboration, and communication. In Indonesia, these demands are reflected in the implementation of the Merdeka Curriculum, which emphasizes student-centered learning and provides greater flexibility for developing students' competencies (Indarta et al., 2022). To support this transformation, the integration of digital technology has become an essential component of classroom learning, enabling more adaptive, interactive, and meaningful learning experiences (Dwita & Zulfitria, 2024). Among recent technological innovations, Artificial Intelligence (AI) has attracted considerable attention because of its ability to provide adaptive learning support through immediate feedback, personalized guidance, and interactive learning experiences. In chemistry education, AI has the potential to assist teachers in facilitating learning while supporting students in understanding complex concepts and developing higher-order thinking skills (Alli, 2025).

Chemistry learning requires students to integrate conceptual understanding, mathematical reasoning, and scientific problem-solving, making it one of the most cognitively demanding subjects at the secondary school level. One of the most challenging chemistry topics is stoichiometry because students must understand quantitative relationships between substances while applying logical and systematic reasoning to solve problems (Borchers et al., 2025). However, many students continue to experience difficulties in solving stoichiometric problems, while teachers often face constraints in providing individualized guidance due to limited instructional time and large class sizes, resulting in learning that tends to remain teacher-centered (Mandina & Kusakara, 2025). Similar conditions were identified through interviews with chemistry teachers at SMAN 17 Muaro Jambi, which revealed that students in Phase F still relied heavily on teacher explanations and had limited opportunities for independent learning. These findings were reinforced by a student needs analysis showing that 83% of students preferred step-by-step explanations, 76% considered technology-based media helpful for understanding chemistry concepts, 80% reported greater motivation when learning through interactive media, and 73% expressed the need for tutor-like learning media that could support independent thinking. These findings

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

indicate the need for AI-assisted learning that provides adaptive guidance while encouraging students to develop problem-solving skills more independently.

Previous studies have reported various educational benefits of AI in chemistry learning. Borchers et al. (2025) found that AI-based intelligent tutoring systems supported students in developing problem-solving strategies during stoichiometry learning. However, the study primarily emphasized the development of learning strategies rather than experimentally measuring students' problem-solving skills as the main learning outcome. Hadibarata et al. (2025) demonstrated that AI-assisted blended learning significantly improved students' conceptual understanding of chemistry. Nevertheless, the study did not specifically examine whether AI also improved students' problem-solving skills. Alli (2025) reported that AI enhanced students' learning motivation and provided more personalized learning experiences, whereas Mandina & Kusakara (2025) highlighted the positive contribution of AI to students' engagement during chemistry learning. Although these studies demonstrate the educational potential of AI, they mainly focus on conceptual understanding, motivation, engagement, and learning experiences, leaving empirical evidence regarding students' problem-solving skills relatively limited.

Several important research gaps therefore remain. First, relatively few studies have specifically investigated the effect of AI on students' problem-solving skills, although problem solving is one of the essential competencies expected in chemistry learning. Second, limited studies have focused specifically on stoichiometry, despite this topic requiring systematic reasoning and quantitative analysis. Third, empirical evidence regarding AI implementation within the context of the Merdeka Curriculum, particularly for Phase F senior high school students in Indonesia, remains limited (Nja et al., 2024). Finally, previous studies have rarely compared AI as a learning assistant with conventional classroom learning using a true experimental design, making the effectiveness of AI-assisted learning difficult to evaluate objectively (Kumar & Kumar, 2025). These limitations indicate the need for further empirical investigation to clarify the contribution of AI-assisted learning to students' problem-solving skills in chemistry.

Accordingly, this study investigates the effect of AI-assisted learning on the problem-solving skills of Phase F high school students in stoichiometry using a true experimental pretest-posttest control group design. Unlike previous studies, this research specifically evaluates students' problem-solving skills based on Polya's problem-solving framework rather than focusing primarily on conceptual understanding or learning motivation. The novelty of this study lies in examining the use of AI as a learning support tool for developing students' problem-solving skills in stoichiometry through four problem-solving stages: understanding the problem, devising a plan, carrying out the plan, and reviewing the solution. Furthermore, this study examines AI-assisted learning within the context of the Merdeka Curriculum and directly compares learning with AI assistance and learning without AI assistance under comparable instructional conditions. By integrating AI-assisted stoichiometry learning, Polya-based problem-solving skills, Phase F learning, and a direct experimental comparison between the two learning conditions, this study provides empirical evidence that extends previous research on AI integration in chemistry education. The findings contribute to the literature on AI integration in chemistry education (Suhailam Pullanikkattil, 2025) and provide evidence-based insights for chemistry teachers and schools regarding the use of AI as a learning support tool to facilitate students' problem-solving skills (Chiu, 2021).

METHOD

Research Design

This study employed a quantitative approach to examine the effect of Artificial Intelligence (AI) use in stoichiometry learning on students' problem-solving skills. The experimental design applied was a pretest-posttest control group design, which enabled the comparison of learning outcomes between an experimental group receiving AI-assisted learning and a control group receiving equivalent instruction without AI integration. In the experimental class, AI was integrated as a learning support tool during the problem-solving process, particularly when students developed and implemented solution strategies for stoichiometric problems. Students worked collaboratively in groups and used AI to consult problem-solving steps, formulate prompts, obtain guidance, and check their reasoning, while the teacher monitored the learning process, guided group discussions, and ensured that students continued to perform the stoichiometric calculations and evaluate the accuracy of AI-generated responses independently. AI assistance was provided during the investigation stage of the two learning meetings,

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

with approximately 40 minutes allocated in the first meeting and 35 minutes in the second meeting for guided problem-solving activities. Students were also instructed to complete the posttest independently without AI assistance. This design was selected to examine the effect of AI-assisted learning on students' problem-solving skills by comparing learning outcomes between students who received AI assistance and those who received equivalent instruction without AI integration (Sugiyono, 2021). The research design applied in this study is presented in Table 1.

Table 1. True Experimental Design (Pretest-Posttest Control Group)

Sampling	Group	Pretest	Treatment	Posttest
Random	Experiment	Y1	X	Y2
Random	Control	Y1	-	Y2

Research Site and Participants

This research was conducted at SMAN 17 Muaro Jambi, Jambi, Indonesia, during the 2025/2026 academic year in stoichiometry learning activities. The research subjects consisted of Phase F senior high school students in grade XI, involving two classes selected using a random sampling technique. The experimental class was XI F4, consisting of 34 students, and received learning treatment integrated with Artificial Intelligence (AI). Meanwhile, the control class was XI F5, also consisting of 34 students, and participated in the same learning process without AI assistance while following equivalent instructional objectives. In total, the study involved 68 students as research participants.

Research Procedures and Instruments

The research was conducted in three stages: preparation, implementation, and final evaluation. During the preparation stage, learning materials and research instruments were designed and prepared prior to classroom implementation. In the implementation stage, both groups were first administered a pretest to assess students' initial problem-solving abilities before the learning treatment. The experimental group then participated in stoichiometry learning integrated with Artificial Intelligence (AI), in which AI functioned as a learning support tool by providing step-by-step guidance, immediate feedback, and assistance for independent learning. Meanwhile, the control group studied the same material without the use of AI while following similar learning objectives. After the completion of the learning process, both groups were administered a posttest to examine changes in students' problem-solving abilities following the treatment. The overall research procedure is presented in Figure 1.

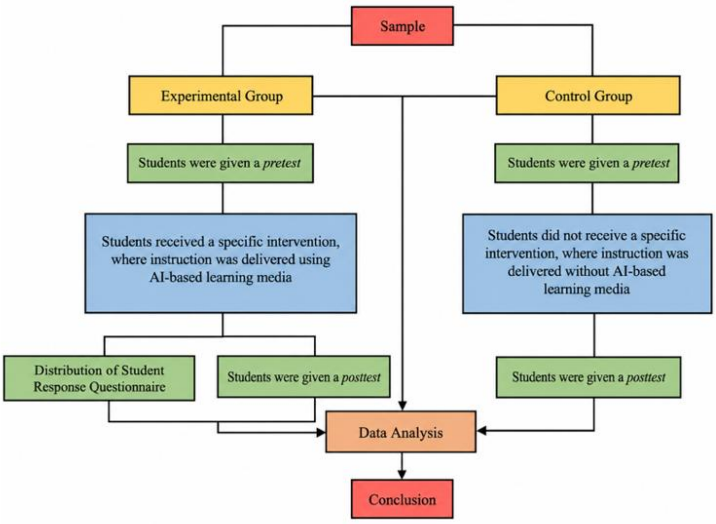

Figure 1. Research Procedure

The instruments used in this study consisted of an essay test and a student response questionnaire. The essay test was developed based on four problem-solving indicators adapted from Polya's framework, namely understanding the problem, planning a solution, implementing the solution plan, and evaluating the final answer. A complete description of the research instruments is presented in Table 2.

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

Table 2. Research Instruments and Data Collection Procedures

Research Activity	Data Source	Data Collection Technique	Instrument
Preliminary Observation	Teacher	Interview	Interview Guide
	Students	Questionnaire Distribution	Student Learning Needs Questionnaire
Students Problem-Solving Skills	Students	Administering a problem-solving skills test	Essay Test (pretest–posttest)
Research Response Evaluation	Students	Questionnaire Distribution	Student Response Questionnaire

Before being used in the study, the essay test instrument underwent expert validation, followed by empirical validity and reliability testing, to ensure its suitability for data collection. The expert validation was conducted by a chemistry education lecturer who evaluated the instrument in terms of content suitability, alignment with the learning objectives and problem-solving indicators, clarity and appropriateness of the questions, language use, and the overall presentation and construction of the test items. Following expert validation, the instrument was empirically tested on a non-sample class. The empirical validity test was conducted using the Pearson product–moment correlation coefficient at a significance level of 5% (α = 0.05). The results indicated that all five essay items were valid, as each item met the required validity criterion, with the obtained correlation coefficient exceeding the critical r-value of 0.3388. These findings indicate that each test item was capable of measuring the intended research variable appropriately. The complete results of the validity test are presented in Table 3.

Table 3. Instrument Validity Test Results

Item Number	Correlation Coefficient (r)	Critical r-Value	Interpretation
Item No. 1	0,692	0,3388	Valid
Item No. 2	0,682	0,3388	Valid
Item No. 3	0,550	0,3388	Valid
Item No. 4	0,720	0,3388	Valid
Item No. 5	0,543	0,3388	Valid

Furthermore, reliability testing was conducted to examine the internal consistency of the instrument using Cronbach's Alpha coefficient analysis, resulting in a coefficient value of 0.609. Based on the reliability criteria adopted in this study, a coefficient value between 0.60 and 0.80 indicates a high degree of reliability (Guilford, as cited in Rezki Putri Juliani & Erita, 2023). Accordingly, the instrument was categorized as having a high degree of reliability and demonstrated sufficient internal consistency for use in the research. The complete results of the reliability test are presented in Table 4.

Table 4. Instrument Reliability Test Results

Statistic	Cronbach's Alpha Value	Interpretation
Reliability Test	0,609	High Reliability

Data Analysis Technique

Data analysis began with instrument validity and reliability testing, followed by prerequisite tests consisting of the Shapiro–Wilk normality test and the Levene homogeneity test. Improvements in students' problem-solving skills were analyzed using the N-Gain test, while the effect of AI-assisted learning was examined using the Independent Sample t-Test at a significance level of 5%. Student response questionnaire data were analyzed descriptively using percentage scores and interpreted according to response category criteria.

RESULT AND DISCUSSION

Result of Student Activity Observation

The implementation of research in the experimental class and the control class is not only focused on the final learning outcomes, but also pays attention to how the learning process takes place during the research. During the learning activities, observers observe student activities to see student involvement in each stage of learning that takes place in both research classes. This observation data provides an overview of the development

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

of student activities during the learning process before further analysis of the research data is carried out. The results of observation data on student activities during learning can be seen in Table 5.

Table 5. Results of Student Activity Observation

No	Student Activity Indicators	Meeting I		Meeting II	
		Experimental Class (%)	Control Class (%)	Experimental Class (%)	Control Class (%)
1	Pupils' attention to learning	60	60	92.5	90
2	Students' readiness to participate in learning	60	55	100	87.5
3	Active participation of students in learning	60	60	92.5	85
4	Students' ability to understand problems	55	57.5	90	77.5
5	Students' ability to make plans	57.5	57.5	90	65
6	Students' ability to implement the plan	62.5	60	90	72.5
7	Students' ability to review	52.5	50	70	65
8	Students' ability to work together in groups	67.5	60	90	90
9	Students' courage in expressing their opinions	57.5	52.5	87.5	77.5
10	Reflection and motivation of students towards learning	67.5	57.5	87.5	77.5
Mean (%)		60	57	89	78.75

The results of observations presented in Table 5. show that there is a change in student activities in both classes during the learning process. These changes can be seen from the percentage of student activity that has increased at each meeting in each research class. In the experimental class, the percentage of student activity at the first meeting was 60% and increased to 89% at the second meeting. Meanwhile, in the control class, the percentage of student activity at the first meeting was 57% and increased to 78.75% at the second meeting. The results indicate that student involvement in both classes improved throughout the learning process, with a greater increase observed in the experimental class. The comparison of student activity percentages in both classes is illustrated in Figure 2.

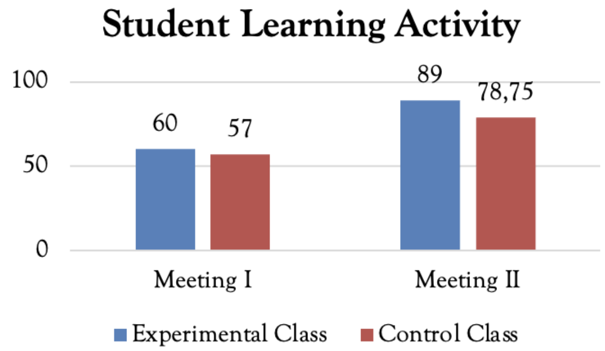

Figure 2. Percentage of Student Activites Based on Observation Results

Observation results showed increased student activity in both classes from the first to the second meeting, with greater improvement in the experimental class. Better performance was also observed across all problem-solving indicators, including understanding the problem, developing a plan, implementing the plan, and reviewing the results. However, these observations alone could not confirm the treatment effect. Therefore, pretest and posttest data were further analyzed using statistical tests to determine the significance of the differences between the two classes.

Description of Essay Test Results Data

In this study, the data used were in the form of essay test results in the form of pretest and posttest. The essay test instrument for five questions was prepared based on problem solving skill indicators according to Polya in (Anugraheni & Kristen Satya Wacana, 2019). Before use, the instrument has gone through content validity validation by experts to ensure the suitability of the material and the accuracy of the measurement to the ability to be studied. Furthermore, a validity test was carried out on the question item using product moment correlation with the help of SPSS Statistics version 29 which showed that all items met the validity criteria at the level of significance α=5%. The reliability test was also carried out using Cronbach's Alpha coefficient and a value of

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

0.609 was obtained which met the reliability criteria, so that the instrument was declared consistent. Thus, all questions in the essay test are suitable for use as a data collection tool in research.

At the beginning of the study, pretests were administered to the experimental and control classes to assess students' initial problem-solving skills in stoichiometry before the learning treatment. The learning process was conducted over two meetings in each class, with time allocations of 90 minutes (2 × 45 minutes) and 135 minutes (3 × 45 minutes), respectively. After the intervention, a posttest was administered to evaluate students' problem-solving skills. The improvement in students' performance was analyzed by comparing the pretest and posttest scores of both classes, and the results are summarized in Table 5.

Table 6. Recapitulation of Pretest, Posttest, and Essay Test Results Improvement

Student Code	Experimental Class			Control Class		
	Pretest	Posttest	Increment (Δ) Posttest - Pretest	Pretest	Posttest	Increment (Δ) Posttest - Pretest
M-1	20	60	40	20	55	35
M-2	15	55	40	15	45	30
M-3	40	65	25	35	50	15
M-4	10	35	25	25	45	20
M-5	20	45	25	15	25	10
M-6	40	65	25	50	65	15
M-7	25	40	15	20	50	30
M-8	45	60	15	30	40	10
M-9	45	80	25	40	55	15
M-10	50	70	20	20	70	50
M-11	35	55	20	10	40	30
M-12	5	45	40	30	75	45
M-13	50	90	25	20	45	25
M-14	45	60	15	50	65	15
M-15	55	75	20	35	60	25
M-16	10	55	45	45	65	20
M-17	10	50	40	65	85	20
M-18	35	50	15	15	35	20
M-19	15	60	45	50	60	10
M-20	40	70	25	25	45	20
M-21	5	40	35	30	40	10
M-22	20	50	30	45	60	15
M-23	10	45	35	30	50	20
M-24	60	75	15	20	30	10
M-25	5	35	30	5	45	40
M-26	55	90	35	35	60	25
M-27	20	45	25	30	45	15
M-28	45	85	40	40	50	10
M-29	35	70	20	45	75	30
M-30	30	75	45	55	70	15
M-31	15	40	25	65	75	10
M-32	50	70	15	20	45	25
M-33	35	70	35	45	65	20
M-34	45	70	25	45	80	35
$\bar{x}$	30.58823529	60.14705882	29.55882353	33.08823529	54.85294118	21.76470588

The data in Table 6 show that the mean ($\bar{x}$) pretest scores of the experimental and control classes were relatively similar, indicating comparable initial problem-solving skills before the learning treatment. After the intervention, the experimental class achieved a higher mean posttest score than the control class, indicating better descriptive performance in solving stoichiometric problems. In addition, the posttest score distribution shifted toward higher score ranges in both classes, with a greater proportion of students in the 61–100 range in the experimental class. These findings indicate an improvement in students' problem-solving skills following the

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

learning process, although the difference between the two classes was further evaluated through statistical hypothesis testing.

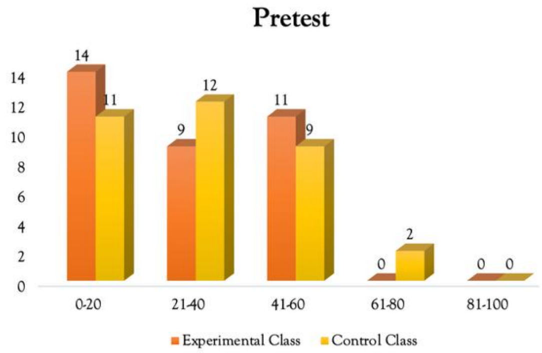

Figure 3. Frequency Distribution Diagram of Pretest Scores for Experimental Class and Control Class

Based on Figure 3, the distribution shows that the majority of students in both classes were in the 0–60 range before the learning treatment was given. The similar distribution pattern of scores between the experimental and control classes indicates that the initial problem-solving skills of students in both groups were at comparable levels.

Figure 4. Frequency Distribution Diagram of Posttest Scores for Experimental Class and Control Class

Furthermore, based on Figure 4, the distribution of posttest scores shows a change in the pattern of score distribution in both classes. This change in score distribution indicates that after the learning implementation, the distribution of student scores shifted to a higher range compared to the initial conditions. The increase in the number of students in the 61–100 score range in the experimental class indicates higher learning outcomes compared to the control class. These data indicate an increase in students' problem-solving skills after the learning treatment, as reflected in the shift in score distribution to a higher category.

Students' Problem Solving Skills

The cognitive assessment of students in this study was analyzed based on problem solving skill indicators which include the ability to understand problems, formulate plans, implement plans, and review. Analysis of each indicator is carried out to identify the achievement of students' abilities at each stage of problem solving during learning. The recapitulation of students' cognitive achievements based on problem solving skill indicators is presented in detail in Table 7.

Table 7. Results of Students' Cognitive Assessment Based on Problem Solving Skill Indicators

Problem Solving Skill Indicators	Questions	Experimental Classes	Control Class
Understanding the problem	1	55%	54%
Creating a plan	2	80%	78%
Execute the plan	3, 4	58%	51%
Reviewing	5	49%	40%
Quantity		242	223
Mean ($\bar{x}$)		60%	56%

 ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

Based on the data presented in Table 7, students' problem-solving skills varied across the four indicators in both the experimental and control classes. The highest achievement was observed in the developing a plan indicator, with percentages of 80% and 78%, respectively, while the implementing the plan indicator also showed higher achievement in the experimental class (58%) than in the control class (51%). In contrast, the understanding the problem indicator showed only a slight difference between the experimental (55%) and control (54%) classes. The lowest achievement in both classes was found in the review indicator, with percentages of 49% and 40%, respectively. Overall, the experimental class achieved a slightly higher average percentage (60%) than the control class (56%). However, this difference was not statistically significant based on the Independent Sample t-test (Sig. = 0.152 > 0.05), indicating that AI-assisted learning showed a positive descriptive tendency but did not significantly improve students' problem-solving skills.

Testing Requirements Analysis

Before hypothesis testing, the data were subjected to prerequisite tests to determine their suitability for parametric statistical analysis. Data processing and analysis were performed using SPSS Statistics version 29. The prerequisite tests included normality and homogeneity tests to evaluate the data distribution and the equality of variances between groups. The results of these tests served as the basis for selecting the appropriate statistical analysis in the subsequent hypothesis testing stage.

Normality Test

Before hypothesis testing, the data were subjected to a normality test as a prerequisite for parametric statistical analysis to ensure that the data were normally distributed. The normality test was performed using the Shapiro–Wilk test in SPSS Statistics version 29 because the sample size consisted of fewer than 50 respondents. The decision was based on the significance value (Sig.), where the data were considered normally distributed if Sig. > 0.05 and not normally distributed if Sig. < 0.05. The results of the pretest and posttest normality tests are presented in Table 8.

Table 8. Results of Pretest and Posttest Data Normality Test for Experimental and Control Class

Normality Test		Pretest		Posttest	
		Experiments	Controls	Experiments	Controls
Shapiro-Wilk	Statistic	0.944	0.965	0.961	0.974
	df	34	34	34	34
	Sig.	0.079	0.345	0.262	0.577
Significance Level(α)		0.05		0.05	
Conclusion		Normal Distributed Data	Normal Distributed Data	Normal Distributed Data	Normal Distributed Data

The results of the normality test in Table 8 show that the pretest and posttest data in both the experimental and control classes have a significance value greater than 0.05, so that all research data are declared to be normally distributed. In the pretest data, the experimental class obtained a significance value of 0.079 and the control class of 0.345. Meanwhile, in the posttest data, the significance value of the experimental class was 0.262 and the control class was 0.577. These results show that the distribution of data in each group meets the assumption of normality as one of the conditions for using parametric statistical analysis. With the fulfillment of these assumptions, hypothesis testing in this study can be continued using parametric tests in the form of independent sample t-tests to analyze the differences in problem solving skills between the experimental class and the control class.

Homogeneity Test

After the normality test, a homogeneity test was conducted as a prerequisite for parametric statistical analysis to determine whether the experimental and control classes had equal variances. The pretest and posttest data were analyzed using Levene's test in SPSS Statistics version 29 to assess the homogeneity of variance between the two groups before hypothesis testing. The basis for the homogeneity test decision is presented as follows:

- If the value of Sig. (p-value) > 0.05, then H_0 is accepted (homogeneous variance).
- If the value of Sig. (p-value) ≤ 0.05, then H_0 is minus (non-homogeneous variance).

The results of the homogeneity test calculation in this study can be seen in Table 9.

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

Table 9. Homogeneity Test Results on Learning Outcome Scores (Posttest) of Experimental Classes and Control Classes

		Levene Statistic	df1	df2	Sig.
Value	Based on Mean	0.096	1	66	0.758
	Based on Median	0.086	1	66	0.770
	Based on Median and with adjusted df	0.086	1	65.959	0.770
	Based on trimmed mean	0.094	1	66	0.760

Based on the homogeneity test results presented in Table 9, the significance value in the Based on Mean section was 0.758, which was greater than 0.05. Therefore, H_0 was accepted, indicating that the variances of the experimental and control classes were homogeneous. Since the homogeneity assumption was satisfied, the data met the prerequisites for parametric statistical analysis and were suitable for subsequent hypothesis testing to examine the effect of AI use on students' problem-solving skills in stoichiometry.

Hypothesis Testing (Independent t-test)

The results of the prerequisite test analysis showed that the pretest and posttest data in both the experimental and control classes met the criteria of normality and homogeneity. This condition shows that the distribution of research data is in the normal category and has the same variance between groups, so the data is suitable for analysis using parametric statistics. Based on the fulfillment of these two assumptions, hypothesis testing in this study was carried out using the Independent Sample t-test with the help of SPSS Statistics software version 29.

According to Buchori et al., (2023) decision-making in the hypothesis test is carried out through a comparison between values $t_{calculated}$ and as well as based on significance values ($t_{critical}$Sig. (2-tailed)). The test decision based on the comparison of t-values is determined by the following conditions:

- If the value $\pm t_{calculated} < \pm t_{critical}$, then H_0 is accepted and H_a is rejected.
- If the value $\pm t_{calculated} > \pm t_{critical}$, then H_0 is rejected and H_a is accepted.

In addition, decision-making can also be determined based on the significance value (p-value) in Sig. (2-tailed) with the following criteria:

- If the significance value $p > 0.05$, then H_0 is accepted.
- If the significance value $p < 0.05$, then H_0 is rejected.

These criteria are used as a basis for determining whether or not there is a difference in the results of students' problem-solving skills between the experimental class and the control class on the stoichiometric material. Thus, the hypothesis in this study is formulated as follows:

H_0 : There is no effect of the use of AI in stochiometry learning on the improvement of problem solving skills of high school students in Phase F at SMAN 17 Muaro Jambi.

H_a : There is an effect of the use of AI in stoichiometric learning on the improvement of problem solving skills of high school students in Phase F at SMAN 17 Muaro Jambi.

Table 10. Hypothesis Test Results on Pretest and Posttest Scores of Experimental and Control Class

Statistics	Value	
	Pretest	Posttest
Sig. (2-tailed)	0.525	0.152
Significance Level (α)	0.05	
Significance Value	0.525 > 0.05	0.152 > 0.05
Conclusion	H_0 accepted	H_a rejected

The results of the hypothesis test in Table 10 show that the significance value of the pretest (Sig. (2-tailed)) is greater than the significance level α = 0.05, which is 0.525 > 0.05, so that the null hypothesis (H_0) is accepted and the alternative hypothesis (H_a) is rejected. These results showed that there was no difference in students' problem-solving skills between the experimental class and the control class before the learning treatment was

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

given. This condition indicates that the initial ability of students in both research groups is at a relatively equal level because the AI-assisted learning process has not been implemented.

Meanwhile, the test results on the posttest data showed that the significance value (Sig. (2-tailed)) is greater than the significance level of $\alpha = 0.05$, i.e. $0.152 > 0.05$, so H_0 is accepted and H_a is rejected. These results showed that there was no significant difference in students' problem-solving skills between the experimental class and the control class after the learning treatment was given. Thus, the use of AI in stoichiometry learning has not shown a significant effect on improving students' problem-solving skills based on the results of hypothesis testing carried out.

N-Gain

After the hypothesis test was carried out, data analysis was continued with the N-Gain test to determine the level of improvement in students' problem-solving skills in the experimental class and the control class after the implementation of learning. The N-Gain test is carried out based on the difference in pretest and posttest scores obtained by students in each class. According to Hake in Kolopita et al. (2022), the N-Gain value obtained is then categorized based on the level of increase listed in Tabel 11.

Table 11. N-Gain Score Result Criteria

Normality Value Gain	Criteria
$0.70 \leq n \leq 1.00$	Height
$0.30 \leq n < 0.70$	Medium
$0.00 \leq n < 0.30$	Low

The N-Gain Score criteria in Table 11 are used to determine the level of problem solving skills of students after the learning treatment is given. Based on these criteria, the results of the N-Gain calculation in the experimental class and the control class are presented in Table 12.

Table 12. N-Gain Test Calculation Results

No	Problem Solving Skill Indicators	Experimental Classes				Control Class			
		Pretest	Posttest	N-Gain	Categories	Pretest	Posttest	N-Gain	Categories
1	Understanding the Problem	35.3	55	0.3	Medium	41.2	54	0.2	Low
2	Creating a plan	48.5	80	0.6	Medium	54.4	78	0.5	Medium
3	Execute the plan	20.6	58.1	0.5	Medium	22.4	51.1	0.4	Medium
4	Reviewing	27.9	49	0.3	Medium	25	40	0.2	Low

The N-Gain results presented in Table 12 showed that students' problem-solving skills improved in both the experimental and control classes after the learning process. In the experimental class, all indicators were categorized as moderate, whereas the control class showed two indicators in the moderate category and two in the low category. These results indicate a more consistent improvement in the experimental class. However, this difference was not statistically significant based on the Independent Sample t-test. Therefore, the N-Gain analysis describes the extent of students' improvement in each class following the learning intervention.

Analysis of Student Response Questionnaire Data

Analysis of student response questionnaire data was carried out in an experimental class to determine students' responses to the use of AI in stoichiometric learning. The questionnaire data was analyzed using Microsoft Excel 2025 software based on the results of students' answers to each indicator that had been determined in the research instrument. The results of the analysis of the students' responses are presented in Tabel 13.

Table 13. Results of Data Analysis of Student Response Questionnaire to the Use of AI in Stoichiometric Learning

No	Indicator	Percentage	Categories
1	Utilization of AI in learning	83%	Excellent

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

No	Indicator	Percentage	Categories
2	Student engagement in AI-assisted learning	89%	Excellent
3	Students' interaction and perception of the use of AI	81%	Excellent

The results of the analysis of student response questionnaire data in Table 13 show that the use of AI in stoichiometric learning obtained a very good category in all indicators measured. The indicator of the use of AI in learning obtained a percentage of 83%, while the indicator of student involvement in AI-assisted learning obtained the highest percentage, which was 89%. Meanwhile, the indicators of students' interaction and perception of the use of AI obtained a percentage of 81%. These results show that the use of AI in the learning process received a positive response from students, both in terms of media utilization, involvement during learning, and perceptions of the application of AI in stoichiometric materials.

The findings revealed a notable pattern in the development of students' problem-solving skills. Although the experimental class demonstrated greater improvement than the control class based on the posttest and N-Gain results, the difference was not sufficiently large to produce a statistically significant effect. This pattern suggests that AI-assisted learning may support students' learning progress at a descriptive level, but its use does not automatically lead to substantial improvements in complex problem-solving skills. The absence of a statistically significant effect may be related to the limited duration and intensity of the intervention, as developing problem-solving skills requires repeated engagement in understanding problems, developing strategies, implementing solutions, and reviewing the results. Thus, the findings indicate that AI may provide useful support during the learning process, but its contribution to problem-solving development depends on how consistently students engage in the complete problem-solving process and how effectively AI is integrated into instruction. This finding is consistent with (Vernanda et al., 2025), who reported that AI use does not necessarily produce significant improvements in learning outcomes under all instructional conditions.

Analysis of the problem-solving indicators provides further insight into the pattern of the overall findings. The experimental class achieved higher percentages than the control class across all four stages of Polya's problem-solving process: understanding the problem (55% vs. 54%), devising a plan (80% vs. 78%), carrying out the plan (58% vs. 51%), and reviewing the solution (49% vs. 40%). The highest achievement in both classes was observed in devising a plan, indicating that students were relatively more capable of formulating solution strategies after identifying the problems. In contrast, reviewing showed the lowest achievement in both classes, suggesting that students were less accustomed to evaluating the accuracy of their procedures and final answers. Although the experimental class achieved a higher overall average (60%) than the control class (56%), the relatively modest differences across the indicators suggest that AI assistance did not produce a substantial shift in students' problem-solving performance. The relatively larger differences observed in carrying out the plan and reviewing the solution may indicate that AI was more useful in providing procedural guidance and supporting students in checking their reasoning, whereas understanding the problem and devising a plan require more fundamental cognitive processes that may develop through repeated independent practice.

The absence of a statistically significant effect may be related to the limited duration and intensity of the intervention. Problem-solving skills require repeated practice and sustained engagement with diverse problems, whereas the AI-assisted learning intervention in this study was implemented over only two learning meetings. Moreover, AI functioned as a supporting tool rather than a substitute for students' cognitive processes. Kaliisa et al. (2025) emphasized that the educational value of AI depends on its integration into appropriate instructional processes and guidance. In addition Li et al. (2025) highlighted that the benefits of AI may be limited when students rely on generated answers without engaging in the underlying reasoning process. This issue is particularly relevant to the present study, as some students tended to seek immediate answers rather than critically examine and develop their own solutions. The findings are also consistent with Şimşek et al. (2025), who reported that AI use does not automatically produce significant improvements in learning outcomes, while Setyawan et al. (2025) emphasized the continuing importance of teacher guidance in AI-assisted learning. Therefore, the non-significant effect observed in this study may be explained by the limited intervention period, the need for repeated problem-solving practice, and differences in the quality of students' engagement with AI assistance.

This pattern can be further understood through the perspectives of scaffolding theory and constructivism. From a scaffolding perspective, AI can provide temporary guidance and feedback that assists students in

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

navigating complex problem-solving tasks. However, the effectiveness of such support depends on whether students gradually internalize the supported processes and become increasingly independent. In the present study, AI provided additional assistance during the problem-solving process, but the limited intervention period may have restricted students' opportunities to transform this assistance into independent competence. From a constructivist perspective, AI functions most effectively as a learning resource that supports students in actively constructing knowledge through problem solving, discussion, and evaluation rather than replacing their cognitive processes. Students were therefore required to formulate prompts, interpret AI-generated responses, perform stoichiometric calculations, and evaluate the accuracy of the solutions. Consequently, the non-significant effect observed in this study suggests that AI support alone is insufficient to produce substantial changes in complex problem-solving skills when students have limited time and opportunities to engage in sustained, independent problem solving.

The N-Gain results showed that all indicators in the experimental class were in the moderate category, while in the control class, some indicators remained in the low category, particularly in the aspects of understanding problems and reviewing. The greatest improvement occurred in the planning indicator, while reviewing remained a weakness for students due to their lack of familiarity with evaluating the process and results of problem solving.

Furthermore, the questionnaire results showed that students responded very positively to the use of AI in learning, with 83% responding to the aspect of AI utilization, 89% to student engagement, and 81% to interaction and perceptions of AI. These findings indicate that AI is well-received as a learning medium because it helps students understand the material, increases learning engagement, and facilitates problem solving, although it has not yet had a significant impact on problem-solving skills.

CONCLUSION

This study contributes empirical evidence to the growing body of research on AI integration in chemistry education by demonstrating that the use of AI as a learning support tool does not automatically produce a statistically significant improvement in students’ problem-solving skills in stoichiometry. Although the experimental class showed greater descriptive improvement and students responded positively to AI-assisted learning, the findings indicate that AI should not be regarded as a standalone solution for developing problem-solving skills. Instead, its contribution is more appropriately understood as a complementary form of support that can provide guidance and feedback when integrated with problem-based learning, active student engagement, and sustained teacher facilitation. Theoretically, this study extends the discussion of AI-assisted chemistry learning by shifting attention beyond conceptual understanding and learning engagement toward the development of Polya-based problem-solving skills in stoichiometry within Phase F learning. Practically, the findings suggest that teachers should integrate AI strategically as a scaffold for students' problem-solving processes while maintaining students' active cognitive involvement and preventing dependence on automatically generated answers. This study was limited to one school, a relatively small sample, a two-meeting intervention, and a limited AI implementation context; therefore, the findings should be interpreted within these conditions. Future research should employ longer interventions, larger and more diverse samples, and varied AI integration strategies to further examine the conditions under which AI can effectively support the development of students' problem-solving skills in chemistry education.

References

Alli, A. O. (2025). A Review of the Potential of Artificial Intelligence (AI) in Enhancing Chemistry Education. *International Journal of Recent Innovations in Academic Research*, *9*(2), 31–43. https://doi.org/10.5281/zenodo.15202897

Anugraheni, I., & Kristen Satya Wacana, U. (2019). Pengaruh Pembelajaran Problem Solving Model Polya Terhadap Kemampuan Memecahkan Masalah Matematika Mahasiswa. *Jurnal Pendidikan*, *4*(1), 1–6. https://doi.org/10.26740/jp.v4n1.p1-6

Borchers, C., Fleischer, H., Yaron, D. J., McLaren, B. M., Scheiter, K., Aleven, V., & Schanze, S. (2025). Problem-Solving Strategies in Stoichiometry Across Two Intelligent Tutoring Systems: A Cross-National Study. *Journal of Science Education and Technology*, *34*(2), 384–400. https://doi.org/10.1007/s10956-024-10197-7

ISSN: 2088-0294 | e-ISSN: 2621-9166
https://doi.org/10.37630/jpm.v16i4.4845

Chiu, W. K. (2021). Pedagogy of emerging technologies in chemical education during the era of digitalization and artificial intelligence: a systematic review. In *Education Sciences* (Vol. 11, Number 11). MDPI. https://doi.org/10.3390/educsci11110709

Dwita, R., & Zulfitria. (2024). TEKNOLOGI PENDIDIKAN DALAM KURIKULUM MERDEKA BELAJAR: MEMBANGUN MASA DEPAN PENDIDIKAN YANG INKLUSIF DAN BERDAYA SAING. *Cendikia: Jurnal Pendidikan Dan Pengajaran*, *2*(6), 26–34.

Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic Review of Artificial Intelligence in Education: Trends, Benefits, and Challenges. Multimodal Technologies and Interaction, 9(8). https://doi.org/10.3390/mti9080084

Hadibarata, T., Hidayat, T., & Khamidun, M. H. (2025). Blended Learning for Stoichiometry and Mass Balance in Environmental Chemistry. *Acta Pedagogia Asiana*, *4*(2), 86–100. https://doi.org/10.53623/apga.v4i2.651

Indarta, Y., Jalinus, N., Waskito, W., Samala, A. D., Riyanda, A. R., & Adi, N. H. (2022). Relevansi Kurikulum Merdeka Belajar dengan Model Pembelajaran Abad 21 dalam Perkembangan Era Society 5.0. *EDUKATIF : JURNAL ILMU PENDIDIKAN*, *4*(2), 3011–3024. https://doi.org/10.31004/edukatif.v4i2.2589

Kaliisa, R., Misiejuk, K., López-Pernas, S., & Saqr, M. (2025). How does artificial intelligence compare to human feedback? A meta-analysis of performance, feedback perception, and learning dispositions. Educational Psychology. https://doi.org/10.1080/01443410.2025.2553639

Kolopita, C. P., Katili, M. R., Mohammad, R., & Yassin, T. (2022a). INVERTED: Journal of Information Technology Education PENGARUH MEDIA PEMBELAJARAN TERHADAP HASIL BELAJAR SISWA PADA MATA PELAJARAN KOMPUTER DAN JARINGAN DASAR. INVERTED (Journal of Information Technology Education), 2(1), 1–12. http://ejurnal.ung.ac.id/index.php/inverted

Kumar, S., & Kumar. (2025). Shodh Sari-An International Multidisciplinary Journal @2025 International Council for Education Research and Training. *Shodh Sari-An International Multidisciplinary Journal*, *04*(02), 380–396. https://doi.org/10.59231/SARI7830

Mandina, S., & Kusakara, R. (2025). The Application of AI in Chemistry Learning: Experiences of Secondary School Students in Zimbabwe. *European Journal of Mathematics and Science Education*, *6*(1), 1–15. https://doi.org/10.12973/ejmse.6.1.1

Nja, C., Edet Uwe, U., Obi, C., Edet, U., & Inwang, V. (2024). Artificial Intelligence Tools of Personalized Learning and Intelligent Tutoring System as Correlates of Students Motivation in Chemistry. *AJSTME Journal*, *10*(1), 27–32. https://www.ajstme.com.ng

Rezki Putri Juliani, S., & Erita, J. (2023). Analisis Validitas dan Reliabilitas Instrumen Penilaian Kemampuan Berpikir Kritis dalam Konteks Sekolah Menengah. *Journal of Educational Integration and Development*, 3(3), 2023.

Setyawan, A., Zakiah Sholiha, E., Meliyani, H., & Aulia, Z. (2025). Dampak Negatif Ketergantungan Berlebihan pada Artificial Intelligence (AI) terhadap Dampak Negatif Ketergantungan Berlebihan pada Artificial Intelligence (AI) terhadap Kemampuan Berpikir Kritis dan Pemecahan Masalah Siswa Sekolah Menengah. JISEP : Jurnal Ilmu Sosial, Ekonomi Dan Pendidikan, 1(2), 31–44. https://jurnal.suriaacademicpress.com/index.php/JISEP

Şimşek, A. C., Anders, G., Göth, J., Specht, L., & Huff, M. (2025). Is ChatGPT a good study companion? The role of AI-generated summaries and reflective prompts in learning from educational videos. Computers and Education: Artificial Intelligence, 9. https://doi.org/10.1016/j.caeai.2025.100512

Suhailam Pullanikkattil, R. Y. C. S. B. (2025). *Artificial Intelligence Based Personalized Learning for Chemical Engineering Education* (1st ed.). CRC Press.

Vernanda, C., Citra Dewi, V., & Eka Jayanti, W. (2025). PENGARUH ARTIFICIAL INTELLEGENCE (AI) TERHADAP KETERAMPILAN BERFIKIR KRITIS PELAJAR ATAU MAHASISWA DALAM MENYELESAIKAN MASALAH. In Journal of Information Systems Management and Digital Business (JISMDB (Vol. 2).