IGNATIAN
INTERNATIONAL JOURNAL FOR MULTIDISCIPLINARY RESEARCH
www.icceph.com
ISSN 2984-9942

EFFECTS OF ARTIFICIAL INTELLIGENCE (AI) ON STUDENT PROBLEM-SOLVING SKILLS

Arlyn Joyce T. Altre, Herby L. Baoing, Jan Marie C. Collado, Jenny Rose E. Lomboy, Arjhay D. Ulnagan

Mathematics Department, Urdaneta City University, Pangasinan, Philippines

ABSTRACT

This study determined the effect of Artificial Intelligence (AI)-assisted pedagogy on the problem-solving skills of first-year Bachelor of Secondary Education students enrolled in Mathematics in the Modern World at Urdaneta City University. The study was completed in May 2026 during School Year 2025-2026. Using a quasi-experimental pre-test and post-test non-equivalent group design, the researchers compared the performance of students taught through conventional pedagogy with those who experienced AI-supported instruction. Problem-solving performance was examined using George Pólya's four-step framework: understanding the problem, planning a solution, carrying out the plan, and checking the solution. Initial results showed that both groups had limited problem-solving skills before instruction, especially in planning, execution, and verification. After the intervention, both groups improved; however, students exposed to AI-assisted pedagogy obtained stronger gains, particularly in carrying out the plan and overall problem-solving performance. The paired-samples t-test indicated significant improvement within the experimental group, while the independent-samples t-test showed that the experimental group outperformed the conventional group in selected post-test measures. The findings suggest that AI tools, when used with teacher guidance and structured problem-solving activities, can support mathematical learning by offering immediate feedback, alternative explanations, and opportunities for solution checking. The study further indicates that AI should be integrated as a cognitive support rather than as a substitute for independent reasoning.

Keywords: *Artificial Intelligence (AI), Problem-Solving Skills, Mathematics Education, AI-Assisted Instruction, Higher-Order Thinking Skills, Educational Technology*

INTRODUCTION

Problem-solving remains one of the most important competencies in mathematics education because it requires learners to interpret situations, connect information, select strategies, and justify solutions. It is not limited to finding an answer; rather, it involves reasoning, decision-making, and applying prior knowledge to unfamiliar or meaningful situations (Abazov, 2016; Oliveri et al., 2017; Mefoh et al., 2017; Sukariasih et al., 2020). When students regularly engage in mathematical problem-solving, they develop habits of analysis, persistence, collaboration, and creative thinking that are useful beyond the classroom (Fitriani et al., 2020; Sholihah & Lastariwati, 2020).

Effective problem-solving also requires a clear process. Polya (1973) organized this process into four steps: understanding the problem, planning, carrying out the plan, and checking the answer. These steps continue to guide mathematics instruction by helping students move from comprehension to strategy selection, execution, and reflection. Related studies also emphasize that students learn mathematics more deeply when they encounter non-routine problems that require sense-making rather than simple memorization of procedures (Hiebert & Grouws, 2007; Sullivan et al., 2015; Lithner, 2017).

As educational technologies develop, problem-solving instruction is increasingly influenced by digital tools. Artificial Intelligence (AI), in particular, has become a relevant innovation because it can provide explanations, generate examples, offer feedback, and support students while they work through mathematical tasks. These possibilities make AI a promising instructional support for strengthening learners' problem-solving performance, provided that its use is guided, ethical, and aligned with learning objectives.

Artificial Intelligence (AI)

Artificial Intelligence refers to computer-based systems designed to perform tasks that normally involve human intelligence, such as learning, reasoning, recognizing patterns, and making decisions (Dignum, 2021; Russell & Norvig, 2021). From its early development in the 1950s, AI has expanded into a multidisciplinary field supported by machine learning, computational modeling, data processing, and predictive systems (Zawacki-Richter et al., 2019).

In education, the growth of Artificial Intelligence in Education (AIED) has introduced new ways to support teaching and learning. AI-powered systems can personalize learning pathways, produce feedback, analyze learner needs, and assist teachers in making instructional decisions (Miao et al., 2021). Generative AI tools such as ChatGPT and Google Gemini have further increased interest in AI by enabling them to respond to prompts, explain concepts, generate examples, and simulate dialogue with users.

Despite these opportunities, AI use also requires caution. Advanced AI models may produce fluent but inaccurate responses, reproduce biases, or encourage

dependence when students use them without reflection. For this reason, the educational value of AI depends on responsible use, verification of outputs, and teacher-designed learning tasks that require students to think critically rather than accept AI-generated answers (Linardatos et al., 2021; von Eschenbach, 2021; Došilović et al., 2018).

Effects of Artificial Intelligence (AI)

AI literacy involves the ability to understand, evaluate, communicate with, and use AI tools appropriately in academic, professional, and everyday settings (Long & Magerko, 2020). In education, AI can help learners process information, explore options, and receive guidance while solving real-world or classroom problems (Chiu et al., 2022).

Large language models and related AI systems can summarize, explain, predict, and generate text or problem-solving suggestions. These functions may support collaboration, creativity, problem-solving, and critical thinking when students are required to examine and justify the outputs they receive (Chakrabarty et al., 2023; Yadav, 2023). AI tools can therefore serve as scaffolds that encourage learners to revisit their ideas, compare solution methods, and check the soundness of their reasoning.

Within educational practice, AI technologies are used to create personalized learning environments, support teachers, and improve institutional processes (Al-Momani & Ramayah, 2024; Chaudhry & Kazim, 2022; Crompton & Burke, 2023). Intelligent tutoring systems, writing assistants, translation tools, and generative chatbots are among the applications that students increasingly encounter. Although these tools expand access to support, they also raise concerns related to data privacy, fairness, accuracy, authorship, and overreliance (Kabudi, 2022).

Relationship of Artificial Intelligence (AI) and Problem-Solving Skills

Generative AI tools have become visible in education because they provide immediate responses and can simulate conversational support. For problem-solving activities, these tools may assist students by restating problems, suggesting representations, generating possible strategies, and providing feedback during solution checking (Zhai, 2022). Such support may be especially useful when learners struggle to move from understanding a problem to executing a solution.

However, AI does not automatically improve problem-solving. Its contribution depends on how students use it and how teachers frame its role. Learners need to evaluate AI responses, compare them with mathematical rules, and justify their own solutions. Without these safeguards, AI may weaken independent reasoning by encouraging students to accept ready-made answers without sufficient cognitive engagement (Lin et al., 2022; Zhai, Wibowo, & Li, 2024; Zhou et al., 2024).

For this reason, the present study treats AI as both a possible cognitive support and a tool requiring careful guidance. It examines whether AI-assisted pedagogy can improve students' performance across Polya's problem-solving steps while also

recognizing the need to maintain students' independent thinking, reflection, and mathematical accountability.

Sustainable Development Goals 4 and 9 and the Use of AI in Education

This study is aligned with Sustainable Development Goal 4, which promotes inclusive and quality education, and Sustainable Development Goal 9, which emphasizes innovation and resilient systems (United Nations, 2015). In mathematics education, AI may contribute to SDG 4 by supporting differentiated instruction, immediate feedback, and learning experiences that address varied student needs (Chiu et al., 2022).

AI integration also has implications for teachers and institutions. Through AI-supported environments, teachers may gain access to additional resources, examples, and learner feedback that can help them design more responsive instruction (Alexander et al., 2024; Viberg et al., 2024). At the same time, ethical implementation is necessary to ensure that technology-enhanced learning remains equitable, transparent, and learner-centered (Hwang & Tu, 2021).

In relation to SDG 9, AI represents an innovation that can strengthen educational systems and promote new ways of addressing learning challenges. Its application in education reflects the broader role of technology in improving access, efficiency, and problem-solving capacity across sectors (Jarrahi, 2018; Jeste et al., 2020; Vinuesa et al., 2020; Goralski & Tan, 2020).

Synthesis of Sub-Themes

The reviewed literature shows that problem-solving is a central mathematical competency developed through structured practice, guided reflection, and opportunities to apply reasoning to meaningful tasks. Conventional teaching can support this development by giving students explanations, examples, and guided exercises. Technology-enhanced environments, on the other hand, can add interactivity, feedback, and flexible access to learning support.

AI extends these possibilities by functioning as an instructional and cognitive support system. It can help students clarify concepts, compare solution methods, and receive feedback as they solve problems. Nevertheless, the literature also warns that AI may create dependence when students use it passively. Thus, effective AI integration should require learners to reason, verify, and explain their answers.

Although many studies discuss AI in education or problem-solving in mathematics, fewer studies directly examine how AI pedagogy affects students' performance across the specific steps of mathematical problem-solving. This gap provides the basis for the present study, which investigates the effects of AI-assisted pedagogy on students' problem-solving skills in Mathematics in the Modern World.

Research Questions

This study aimed to examine the effects of Artificial Intelligence (AI) on students' problem-solving skills. Specifically, it sought to answer the following questions:

1. What is the level of problem-solving skills of students exposed to conventional pedagogy in the pre-test and post-test?
2. What is the level of problem-solving skills of students exposed to AI pedagogy in the pre-test and post-test?
3. Is there a significant difference between the level of problem-solving skills of students in the pre-test and post-test of the conventional and experimental groups?

METHODOLOGY

The study used a quasi-experimental research design, specifically the pre-test and post-test non-equivalent group design. This design was appropriate because the respondents came from existing class sections and random assignment of individual students was not feasible in the natural classroom setting.

Two groups were compared: a conventional group that received regular instruction and an experimental group that experienced AI-assisted pedagogy. Both groups took a pre-test before the intervention and a post-test after the six-week instructional period. This allowed the researchers to examine changes within each group and compare the performance of the two groups after instruction.

Subjects of the Study and Sampling Scheme

The respondents were first-year Bachelor of Secondary Education students majoring in Mathematics and Science at Urdaneta City University during the second semester of Academic Year 2025-2026. Intact group sampling was used to preserve the existing class organization and avoid disruption of regular instruction. A total of 55 students participated in the study. Table 1 shows the distribution of the respondents according to their field of specialization.

Table 1
Distribution of the Respondents along Field of Specialization

Field of Specialization	Number of First Year BSEd
Mathematics	26
Science	29
Total	55

Data Gathering Instrument

The main instruments were researcher-developed pre-test and post-test assessments designed to measure students' problem-solving skills in mathematical

contexts. The test items required students to analyze problems, choose strategies, carry out solutions, and verify answers based on Pólya's problem-solving steps.

To establish content validity, three experts in mathematics education and research evaluated the instruments according to clarity, relevance, organization, and alignment with the study objectives. The research instrument obtained an overall mean of 4.43, interpreted as Valid. The instructional materials obtained an overall mean of 4.8, interpreted as Highly Valid.

A pilot test was conducted among second-year BSEd Mathematics and Science students who were not part of the actual respondents. Based on the item analysis, items 3, 4, 5, 6, 8, 10, 11, 13, and 15 were retained because they were within the acceptable difficulty range. Items that were too easy, difficult, or unclear were revised or removed before the final administration.

The Intervention: Artificial Intelligence (AI) Pedagogy Approach

The intervention was an AI-assisted pedagogical approach designed to support students' mathematical problem-solving skills. The approach integrated ChatGPT and Google Gemini as learning tools that could provide explanations, worked examples, prompts, feedback, and solution-checking support.

AI tools were used within teacher-guided activities rather than as independent substitutes for instruction. Students were asked to solve mathematical problems, consult AI for clarification or feedback when appropriate, discuss their reasoning, and verify the correctness of their solutions. Collaborative activities were also included so that students could compare strategies and improve their understanding.

The AI-assisted pedagogy followed a four-step sequence aligned with Pólya's problem-solving framework.

1. Step 1: AI-Assisted Interactive Motivation Activities. The lesson began with motivation activities supported by AI-generated prompts or visual materials. These activities activated prior knowledge and helped students identify what the problem was asking them to do.

2. Step 2: AI-Guided Discussion. The teacher presented the lesson using reference-based materials enhanced with AI-generated examples and explanations. This step helped students analyze the problem and consider possible strategies.

3. Step 3: AI-Supported Activities. Students solved practice problems while using AI tools for guidance, feedback, and verification. They responded to reflective questions about corrections, understanding, and the usefulness of AI feedback.

4. Step 4: AI-Supported Assignment. Students completed individual tasks and used AI tools to review their solutions. This step emphasized checking, revision, and reflection on the accuracy of answers.

Data Gathering Procedure

The researchers first prepared and submitted a formal request to the Dean of the College of Teacher Education at Urdaneta City University. After approval, they coordinated with the concerned instructors and administered the research instruments to the identified classes.

The pre-test was administered to both the conventional and experimental groups to determine their initial level of problem-solving skills. The instructional phase then lasted for six weeks during the preliminary term and covered selected lessons from Unit 1: The Nature of Mathematics.

The conventional group received teacher-led instruction, guided practice, and regular classroom activities. The experimental group experienced the AI-assisted pedagogical approach, where ChatGPT and Google Gemini were integrated into selected learning activities as support tools for explanation, feedback, and solution verification.

After the intervention, both groups took the post-test. The pre-test and post-test were parallel in content and structure and covered lessons from the three modules. The responses were checked, organized, tabulated, and analyzed using appropriate statistical tools.

The level of problem-solving skills of the respondents was determined using a scoring rubric. Table 2 presents the rubric used to evaluate each step of Polya's problem-solving strategy.

Table 2
Score on Each Step of Polya's Problem-Solving Strategy

Score on each step of Polya's Problem-Solving Strategy	Rating	Descriptions
2	Full Credit	The answer is correct and complete
1	Partial Credit	The answer is partially correct and incomplete
0	No Credit	The answer is incorrect

After obtaining the respondents' total scores, their level of problem-solving skills was classified accordingly. Table 3 shows the corresponding skill level based on the total scores.

Table 3
Level of Problem-Solving Skills Based on Total Score

Score on Each Step of Polya's Problem-Solving Strategy	Total Score	Level of Problem-Solving Skills
16-18	60-72	Highly Skilled
12-15	45-59	Skilled
8-11	30-44	Moderately Skilled
4-7	15-29	Slightly Skilled
0-3	0-14	Not Skilled

Statistical Treatment of Data

Descriptive statistics were used to determine the students' level of problem-solving skills in the pre-test and post-test. The mean described general performance, while the standard deviation and skewness provided information about the distribution and variability of scores.

To compare pre-test and post-test scores within each group, a paired-samples t-test was used. To compare the performance of the conventional and experimental groups, an independent-samples t-test was employed. All tests were interpreted at the 0.05 level of significance.

Scope and Limitations

The study was conducted at Urdaneta City University and focused on first-year Bachelor of Secondary Education students majoring in Mathematics and Science who were enrolled in Mathematics in the Modern World during the second semester of Academic Year 2025-2026. It was limited to determining the effects of AI-assisted pedagogy on students' mathematical problem-solving skills.

The Intervention covered six weeks of preliminary lessons under Unit 1: The Nature of Mathematics. The lessons included Mathematics in Our World, Mathematical Language and Symbols, and Problem-Solving and Reasoning. Topics involved patterns, the Fibonacci sequence, the golden ratio, Pascal's triangle, mathematical expressions and sentences, relations, functions, set theory, logic, inductive and deductive reasoning, proofs, Pólya's problem-solving steps, and recreational mathematics problems.

RESULTS

Level of Problem-Solving Skills Exposed to Conventional pedagogy

The following tables present the level of problem-solving skills of respondents exposed to conventional pedagogy. The data compare students' performance in the pre-test and post-test following instruction using conventional teaching methods.

Table 4
Level of Problem-Solving Skills of Students Exposed to Conventional Pedagogy in the Pre-test and Post-test

Polya's Four Steps in Problem-Solving	Scores	Level of Problem-Solving Skills	Conventional Group			
			Pre-test		Post-test	
			Frequency	%	Frequency	%
Understanding the problem	16 to 18	Highly Skilled	0	0	25	86.21
	12 to 15	Skilled	1	3.45	3	10.34
	8 to 11	Moderately Skilled	5	17.24	1	3.45
	4 to 7	Slightly Skilled	12	41.38	0	0
	0 to 3	Not Skilled	11	37.93	0	0
Planning how to solve it	16 to 18	Highly Skilled	0	0	14	48.28
	12 to 15	Skilled	0	0	7	24.14
	8 to 11	Moderately Skilled	0	0	8	27.59
	4 to 7	Slightly Skilled	11	37.93	0	0
	0 to 3	Not Skilled	18	62.07	0	0
Carrying out the plan	16 to 18	Highly Skilled	0	0	4	13.79
	12 to 15	Skilled	0	0	15	51.72
	8 to 11	Moderately Skilled	0	0	10	34.48
	4 to 7	Slightly Skilled	7	24.14	0	0
	0 to 3	Not Skilled	22	75.86	0	0
Checking if the solution is correct	16 to 18	Highly Skilled	0	0	10	34.48
	12 to 15	Skilled	0	0	10	34.48
	8 to 11	Moderately Skilled	0	0	7	24.14
	4 to 7	Slightly Skilled	1	3.45	2	6.9

Table 5
Descriptive Statistics of the Pre-test and Post-test Problem-Solving Skills of Students Exposed to Conventional Pedagogy

	Range	Min	Max	Mean	SD	Var	Skewness	Kurtosis
Pre-test								
Understanding the problem	10	2	12	5.10	2.782	7.739	.634	-.347
Planning how to solve it	5	0	5	2.90	1.398	1.953	-.562	-.021
Carrying out the plan	6	0	6	2.55	1.594	2.542	.410	-.009
Checking if the solution is correct	4	0	4	.79	1.146	1.313	1.353	.975

Overall	17	2	19	11.34	4.723	22.305	-.167	-.663
Post-test								
Understanding the problem	8	10	18	16.79	1.698	2.884	-2.566	8.558
Planning how to solve it	9	9	18	14.00	3.327	11.071	-.219	-1.510
Carrying out the plan	9	8	17	12.83	2.647	7.005	-.334	-.999
Checking if the solution is correct	11	6	17	13.17	3.496	12.219	-.586	-.779
Overall	29	38	67	56.55	8.638	74.613	-.598	-.778

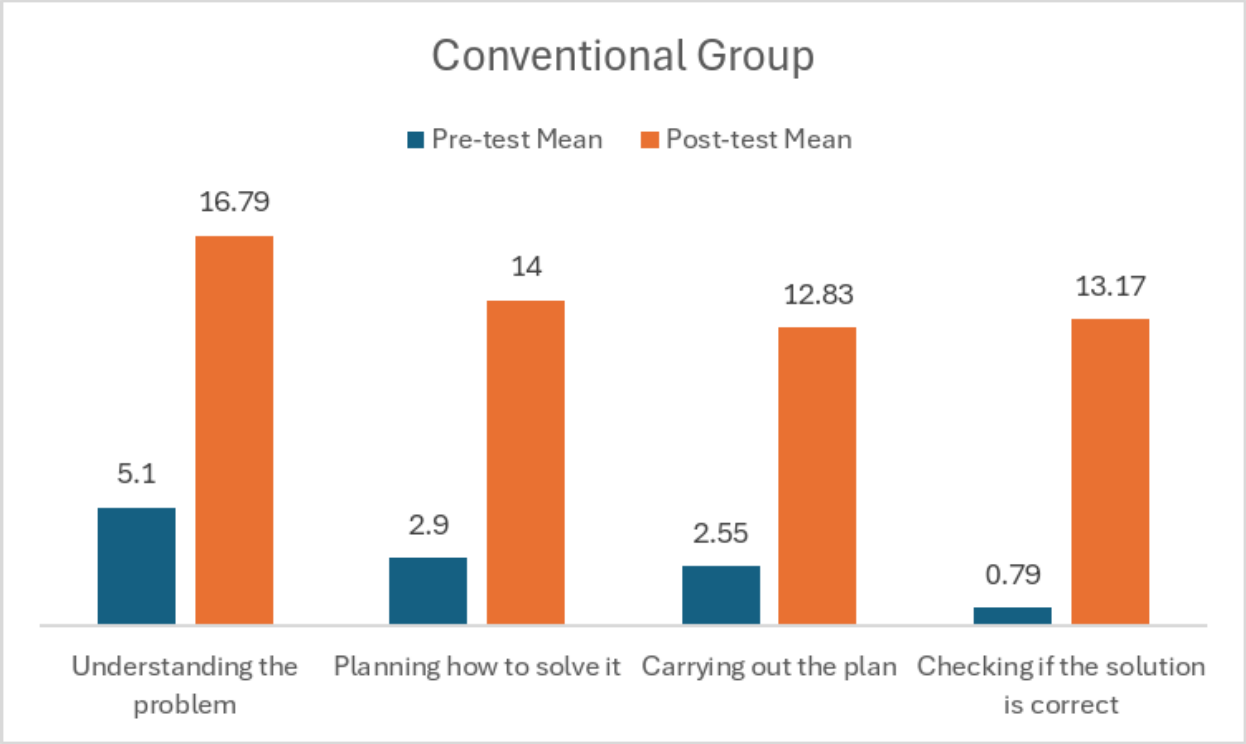


Figure 1. Comparison of Pre-test and Post-test Mean Scores Based on Polya's Problem-Solving Steps (Conventional Group)

Level of Problem-Solving Skills Exposed to AI pedagogy

The following tables present the level of problem-solving skills of respondents exposed to AI pedagogy. The data compare students' pre-test and post-test scores after exposure to AI-based instruction.

Table 6
Level of Problem-Solving Skills of Students Exposed to AI Pedagogy in the Pre-test and Post-test

Polya's Four Steps in Problem-Solving	Scores	Level of Problem-Solving Skills	Experimental Group			
			Pre-test		Post-test	
			Frequency	%	Frequency	%
Understanding the problem	16 to 18	Highly Skilled	2	7.69	24	92.31
	12 to 15	Skilled	3	11.54	2	7.69
	8 to 11	Moderately Skilled	10	38.46	0	0
	4 to 7	Slightly Skilled	5	19.23	0	0
	0 to 3	Not Skilled	6	23.08	0	0
Planning how to solve it	16 to 18	Highly Skilled	0	0	13	50.00
	12 to 15	Skilled	0	0	8	30.77
	8 to 11	Moderately Skilled	0	0	2	7.69
	4 to 7	Slightly Skilled	4	15.38	3	11.54
	0 to 3	Not Skilled	22	84.62	0	0
Carrying out the plan	16 to 18	Highly Skilled	0	0	16	61.54
	12 to 15	Skilled	1	3.85	7	26.92
	8 to 11	Moderately Skilled	1	3.85	3	11.54
	4 to 7	Slightly Skilled	8	30.78	0	0
	0 to 3	Not Skilled	16	61.54	0	0
Checking if the solution is correct	16 to 18	Highly Skilled	0	0	16	61.54
	12 to 15	Skilled	0	0	4	15.38
	8 to 11	Moderately Skilled	1	3.85	4	15.38
	4 to 7	Slightly Skilled	1	3.85	2	7.69
	0 to 3	Not Skilled	24	92.31	0	0

Table 7
Descriptive Statistics of the Pre-test and Post-test Problem-Solving Skills of Students Exposed to AI Pedagogy

	Range	Min	Max	Mean	SD	Var	Skewness	Kurtosis
Pre-test								
Understanding the problem	17	1	18	7.69	4.671	21.822	.367	-.265
Planning how to solve it	5	0	5	2.27	1.251	1.565	.108	-.268
Carrying out the plan	13	0	13	3.31	3.017	9.102	1.602	3.230

Checking if the solution is correct	8	0	8	.88	1.883	3.546	2.789	8.271
Overall	40	2	42	14.12	8.990	80.826	1.339	2.620
Post-test								
Understanding the problem	3	15	18	17.08	.977	.954	-.722	-.493
Planning how to solve it	14	4	18	14.31	3.876	15.022	-1.351	1.083
Carrying out the plan	9	9	18	15.27	2.255	5.085	-1.293	1.463
Checking if the solution is correct	12	6	18	14.81	3.699	13.682	-1.253	.603
Overall	28	44	72	61.50	8.392	70.420	-.927	-.289

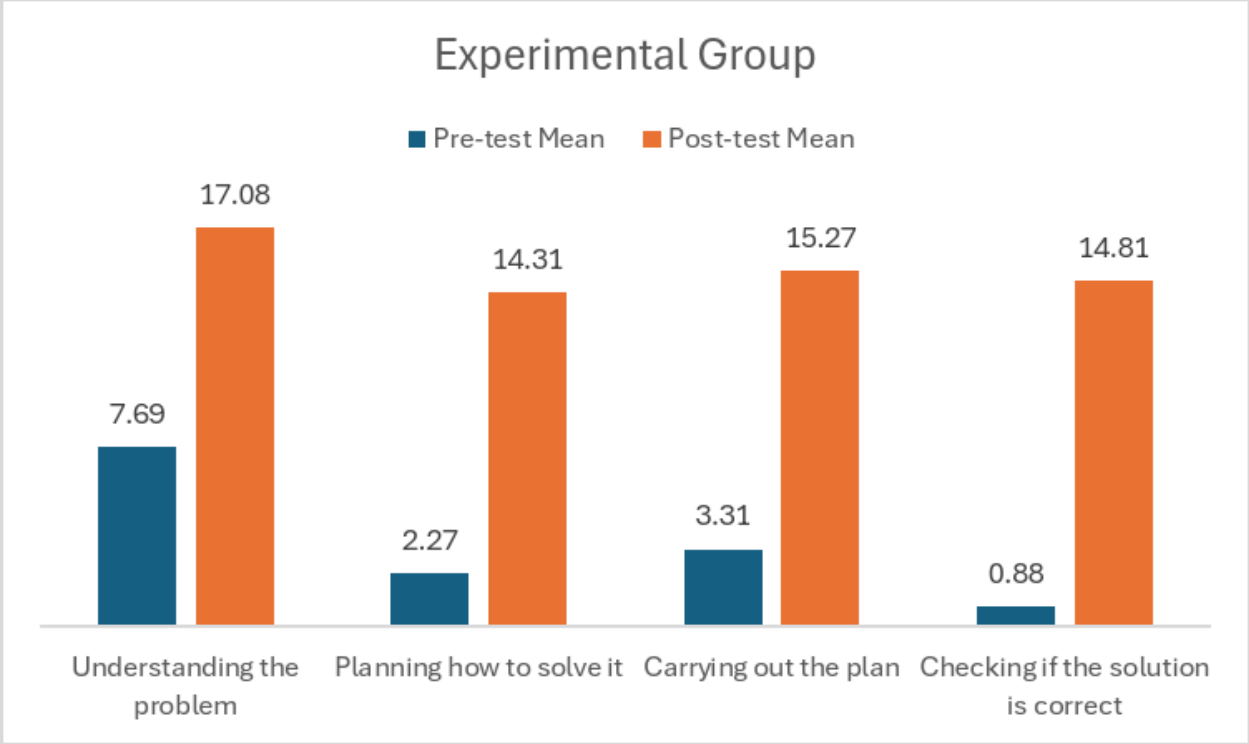


Figure 2. Comparison of Pre-test and Post-test Mean Scores Based on Polya's Problem-Solving Steps (Experimental Group)

Differences between the Level of Problem-solving Skills of Students

The succeeding tables show the differences between the level of problem-solving skills of each group in their pre-test and post-test.

Table 8
Independent Samples t-test Results of the Pre-test Problem-Solving Skills of the Experimental and Conventional Groups

Polya's Four Steps in Problem-Solving	Group	Mean	Mean Difference	LPS	Std. Deviation	t	df	Sig. (2-tailed)	Interpre-tation
Understanding the problem	Conventional	5.10	2.589	SS	2.782	2.528	53	.015	Significant
	Experimental	7.69		MS	4.671				
Planning how to solve it	Conventional	2.90	-.627	NS	1.398	-1.746	53	.087	Not Significant
	Experimental	2.27		NS	1.251				
Carrying out the plan	Conventional	2.55	.756	NS	1.594	1.179	53	.244	Not Significant
	Experimental	3.31		NS	3.017				
Checking if the solution is correct	Conventional	.79	.092	NS	1.146	.220	53	.827	Not Significant
	Experimental	.88		NS	1.883				
Overall	Conventional	11.34	2.771	NS	4.723	1.452	53	.152	Not Significant
	Experimental	14.12		NS	8.990				

Note: p<0.05 is significant
Legend:

Score on Each Step of Polya's Problem-Solving Strategy	Total Score	Level of Problem-Solving Skills (LPS)
16-18	60-72	Highly Skilled (HS)
12-15	45-59	Skilled (S)
8-11	30-44	Moderately Skilled (MS)
4-7	15-29	Slightly Skilled (SS)
0-3	0-14	Not Skilled (NS)

Table 9
Independent Samples t-test Results of the Post-test Problem-Solving Skills of the Experimental and Conventional Groups

Polya's Four Steps in Problem-Solving	Group	Mean	Mean Difference	LPS	Std. Deviation	t	df	Sig. (2-tailed)	Interpre-tation
Understanding the problem	Conventional	16.69	.284	HS	1.698	.748	53	.458	Not Significant
	Experimental	17.08		HS	.977				
Planning how to solve it	Conventional	14.00	.308	S	3.327	.317	53	.753	Not Significant
	Experimental	14.31		S	3.876				
Carrying out the plan	Conventional	12.83	2.442	S	2.647	3.661	53	.001	Significant
	Experimental	15.27		S	2.255				
Checking if the solution is correct	Conventional	13.17	1.635	S	3.496	1.685	53	.098	Not Significant
	Experimental	14.81		S	3.699				
Overall	Conventional	56.55	4.948	S	8.638	2.150	53	.036	Significant
	Experimental	61.50		HS	8.392				

Note: $p<0.05$ is significant
Legend:

Score on Each Step of Polya's Problem-Solving Strategy	Total Score	Level of Problem-Solving Skills (LPS)
16-18	60-72	Highly Skilled (HS)
12-15	45-59	Skilled (S)
8-11	30-44	Moderately Skilled (MS)
4-7	15-29	Slightly Skilled (SS)
0-3	0-14	Not Skilled (NS)

Table 10
Paired Samples t-test Results of the Pre-test and Post-test Problem-Solving Skills of the Conventional Group

Polya's Four Steps in Problem-Solving	Pre-test / Post-test	Mean	Mean Difference	LPS	Std. Deviation	t	df	Sig. (2-tailed)	Interpretation
Understanding the problem	Pre-test Post-test	5.10 16.79	-11.690	SS HS	3.874	-16.250	28	.000	Significant
Planning how to solve it	Pre-test Post-test	2.90 14.00	-11.103	NS S	3.678	-16.259	28	.000	Significant
Carrying out the plan	Pre-test Post-test	2.55 12.83	-10.276	NS S	2.750	-20.121	28	.000	Significant
Checking if the solution is correct	Pre-test Post-test	.79 13.17	-12.379	NS S	3.353	-19.881	28	.000	Significant
Overall	Pre-test Post-test	11.34 56.55	-45.207	NS S	10.220	-23.820	28	.000	Significant

Note: p<0.05 is significant
Legend:

Score on Each Step of Polya's Problem-Solving Strategy	Total Score	Level of Problem-Solving Skills (LPS)
16-18	60-72	Highly Skilled (HS)
12-15	45-59	Skilled (S)
8-11	30-44	Moderately Skilled (MS)
4-7	15-29	Slightly Skilled (SS)
0-3	0-14	Not Skilled (NS)

Table 11
Paired Samples t-test Results of the Pre-test and Post-test Problem-Solving Skills of the Experimental Group

Polya's Four Steps in Problem-Solving	Pre-test / Post-test	Mean	Mean Difference	LPS	Std. Deviation	t	df	Sig. (2-tailed)	Interpre-tation
Understanding the problem	Pre-test Post-test	7.69 17.08	-9.385	MS HS	4.826	-9.916	25	.000	Significant

Planning how to solve it	Pre-test Post-test	2.27 14.31	-12.038	NS S	4.181	-14.683	25	.000	Significant
Carrying out the plan	Pre-test Post-test	3.31 15.27	-11.962	NS S	3.527	-17.294	25	.000	Significant
Checking if the solution is correct	Pre-test Post-test	.88 14.81	-13.923	NS S	3.773	-18.817	25	.000	Significant
Overall	Pre-test Post-test	14.12 61.50	-47.385	NS HS	11.243	-21.490	25	.000	Significant

Note: $p<0.05$ is significant
Legend:

Score on Each Step of Polya's Problem-Solving Strategy	Total Score	Level of Problem-Solving Skills (LPS)
16-18	60-72	Highly Skilled (HS)
12-15	45-59	Skilled (S)
8-11	30-44	Moderately Skilled (MS)
4-7	15-29	Slightly Skilled (SS)
0-3	0-14	Not Skilled (NS)

DISCUSSION

Table 4 shows the level of problem-solving skills of students exposed to conventional pedagogy during the pre-test and post-test, organized according to George Pólya's four steps of problem-solving: understanding the problem, planning a solution, carrying out the plan, and checking if the solution is correct.

During the pre-test, the results reveal that students generally exhibited low levels of problem-solving skills across all four stages. For understanding the problem, most students were categorized as slightly skilled (41.38%) and not skilled (37.93%), while only a few reached the skilled level (3.45%), and none were classified as highly skilled. This indicates that, at the beginning, students had difficulty interpreting and making sense of mathematical problems.

A similar pattern was observed in planning how to solve the problem, where a majority of students (62.07%) were not skilled, while the remaining 37.93% were classified as slightly skilled. This suggests limited ability in identifying and organizing appropriate solution strategies.

In carrying out the plan, most students (75.86%) were not skilled, while 24.14% were slightly skilled, indicating challenges in correctly executing mathematical procedures and operations.

Likewise, in checking if the solution is correct, an overwhelming majority (96.55%) were not skilled, showing a very weak ability in evaluating and verifying their answers.

In contrast, the post-test results demonstrate a significant improvement in students' problem-solving skills across all steps. In understanding the problem, most students (86.21%) reached the highly skilled level, while the rest were classified as skilled (10.34%) and moderately skilled (3.45%). This reflects a stronger ability to comprehend and interpret mathematical problems after instruction.

In planning how to solve the problem, students also showed improvement, with 48.28% reaching the highly skilled level, 24.14% skilled, and 27.59% moderately skilled. This indicates better capability in selecting and organizing solution strategies.

For carrying out the plan, most students were classified as skilled (51.72%) and moderately skilled (34.48%), while 13.79% reached the highly skilled level. This shows improvement in executing mathematical procedures, although fewer students attained the highest level compared to other steps.

Similarly, in checking if the solution is correct, students demonstrated marked progress, with equal proportions achieving highly skilled (34.48%) and skilled (34.48%) levels, while 24.14% were moderately skilled and only 6.9% remained slightly skilled. Notably, none were classified as not skilled, indicating improved ability in evaluating and verifying solutions.

Overall, the results show a clear progression from low to higher levels of problem-solving skills among students exposed to conventional pedagogy. While the pre-test results highlight substantial difficulties across all steps, the post-test results indicate that classroom instruction and practice contributed to measurable improvement in performance. However, differences in performance levels across the four steps suggest that some areas—particularly executing procedures and evaluating solutions—still require further development. This implies that although conventional pedagogy supports improvement through practice, more structured and guided instructional approaches may further enhance students' higher-order problem-solving skills.

Table 5 presents the descriptive statistics of the pre-test and post-test problem-solving skills of students exposed to conventional pedagogy across the four steps in George Pólya's problem-solving process: understanding the problem, planning how to solve it, carrying out the plan, and checking if the solution is correct.

In the pre-test, the overall mean score of 11.34 (SD = 4.72) indicates that students were generally classified under the Not Skilled (NS) level, reflecting limited prior knowledge and underdeveloped problem-solving abilities before the intervention. Across the four steps, students were Slightly Skilled (SS) in understanding the problem (M = 5.10), which suggests that some learners had a basic ability to interpret given problems. However, in planning how to solve it (M = 2.90), carrying out the plan (M = 2.55), and checking if the solution is correct (M = 0.79), students were all categorized as Not Skilled

(NS). This implies that learners initially struggled in identifying appropriate strategies, executing solution procedures, and verifying their answers, with checking being the weakest step.

The pre-test standard deviation shows a moderate dispersion of scores, indicating variability in students' initial problem-solving abilities. In addition, the positively skewed distribution in checking if the solution is correct (skewness = 1.353) suggests that most students obtained low scores in this step, with only a few achieving higher performance.

In the post-test, a marked improvement in students' problem-solving skills was observed. The overall mean score increased to 56.55 (SD = 8.64), corresponding to the Skilled (S) level. This indicates that students improved their problem-solving performance after exposure to conventional pedagogy.

Across the four steps, improvement was consistently observed. Students demonstrated a Highly Skilled (HS) level in understanding the problem (M = 16.79), reflecting strong ability in interpreting and analyzing problems. In planning how to solve it (M = 14.00), carrying out the plan (M = 12.83), and checking if the solution is correct (M = 13.17), students were classified as Skilled (S). These results indicate that learners improved not only in understanding problems but also in selecting strategies, executing procedures, and evaluating their solutions.

The post-test standard deviations indicate some variation in performance among students; however, the higher mean scores show that most learners performed better overall. The negatively skewed distribution in understanding the problem (skewness = -2.566) suggests that scores were concentrated at the higher end, meaning that many students achieved strong performance in this step.

Overall, the findings show a clear and substantial improvement in students' problem-solving skills from pre-test to post-test across all four steps. Although students initially demonstrated limited proficiency—particularly in the steps involving execution and evaluation—they showed improvement through instruction and practice. However, differences across the four steps suggest that certain skills, especially those requiring deeper analysis and evaluation, may still need further strengthening. This indicates that while conventional pedagogy can improve basic problem-solving performance, additional instructional support may still be needed to enhance students' analytical and evaluative skills in mathematics. This improvement is also reflected in Figure 1, which shows an increase in mean scores from pre-test to post-test across all four steps.

Table 6 shows the level of problem-solving skills of students exposed to AI pedagogy during the pre-test and post-test, organized according to George Pólya's four steps of problem-solving: understanding the problem, planning how to solve it, carrying out the plan, and checking if the solution is correct.

In the pre-test, the results show that students demonstrated generally low to moderate levels of problem-solving skills across the four steps. For understanding the

problem, most students were classified as moderately skilled (38.46%) and not skilled (23.08%), while a few reached the skilled (11.54%) and highly skilled (7.69%) levels. This indicates that although some learners had initial comprehension abilities, many still had difficulty fully interpreting mathematical problems.

In planning how to solve it, the majority of students were not skilled (84.62%), while only a small portion were slightly skilled (15.38%). This suggests that students initially struggled in identifying appropriate strategies and organizing solution steps.

A similar pattern was observed in carrying out the plan, where most students were not skilled (61.54%) and slightly skilled (30.78%), with only a few reaching the skilled (3.85%) and moderately skilled (3.85%) levels. This reflects difficulties in accurately executing mathematical procedures.

Likewise, in checking if the solution is correct, a large majority of students (92.31%) were not skilled, indicating weak ability in evaluating and verifying answers.

Overall, the pre-test results show that students had limited competence across all steps of problem-solving before the intervention.

In contrast, the post-test results reveal a substantial and consistent improvement across all four steps. In understanding the problem, most students (92.31%) reached the highly skilled level, while the rest were classified as skilled (7.69%). This indicates a great improvement in students' ability to interpret and analyze problems.

In planning how to solve it, half of the students (50.00%) reached the highly skilled level, followed by skilled (30.77%), while a smaller proportion were moderately skilled (7.69%) and slightly skilled (11.54%). This shows improved ability in selecting and organizing solution strategies.

A more evident improvement was seen in carrying out the plan, where most students were highly skilled (61.54%), followed by skilled (26.92%) and moderately skilled (11.54%). Notably, no students remained in the lower skill levels, indicating improved accuracy in executing procedures.

Similarly, in checking if the solution is correct, most students reached the highly skilled level (61.54%), while others were skilled (15.38%), moderately skilled (15.38%), and slightly skilled (7.69%). Importantly, no students were classified as not skilled in the post-test, showing improved ability in evaluating solutions.

Overall, the findings indicate a clear shift from lower to higher levels of problem-solving skills among students exposed to AI pedagogy. Compared to the pre-test, the post-test results show not only improvement but also a stronger concentration of students in the highest skill level across all four steps. This suggests that AI pedagogy was effective in enhancing students' problem-solving performance, particularly in higher-order skills such as strategy formulation, execution, and evaluation. This implies that integrating

Artificial Intelligence (AI) in mathematics instruction can support learning by providing guided feedback, structured assistance, and interactive experiences that strengthen students' understanding and problem-solving abilities.

Table 7 presents the descriptive statistics of the pre-test and post-test problem-solving skills of students exposed to AI pedagogy across the four steps in George Pólya's problem-solving process: understanding the problem, planning how to solve it, carrying out the plan, and checking if the solution is correct.

In the pre-test, the overall mean score of 14.12 (SD = 8.99) indicates that students were generally classified at the Not Skilled (NS) level, suggesting limited problem-solving ability before the intervention. Across the four steps, students were classified as Moderately Skilled (MS) in understanding the problem (M = 7.69), which shows that some learners had a developing ability to interpret mathematical problems. However, in planning how to solve it (M = 2.27), carrying out the plan (M = 3.31), and checking if the solution is correct (M = 0.88), students were categorized as Not Skilled (NS). This implies that although basic comprehension was present, students still had difficulty selecting strategies, executing procedures, and verifying solutions, with checking identified as the weakest step.

The relatively high standard deviations in the pre-test, particularly in understanding the problem and carrying out the plan, indicate variability in students' initial performance. In addition, the positively skewed distributions in carrying out the plan (skewness = 1.602) and checking if the solution is correct (skewness = 2.789) suggest that many students obtained low scores in these steps, reflecting difficulties in both execution and evaluation.

In the post-test, a marked improvement in students' problem-solving skills was observed. The overall mean score increased to 61.50 (SD = 8.39), corresponding to the Highly Skilled (HS) level. This indicates that students significantly improved their problem-solving performance after exposure to AI pedagogy.

Across the four steps, improvement was consistently observed. Students reached the Highly Skilled (HS) level in understanding the problem (M = 17.08), demonstrating strong ability in interpreting and analyzing mathematical problems. In planning how to solve it (M = 14.31), carrying out the plan (M = 15.27), and checking if the solution is correct (M = 14.81), students were classified as Skilled (S). This shows that learners improved not only in comprehension but also in selecting strategies, executing procedures, and evaluating their solutions.

The post-test standard deviations indicate some variation in performance; however, the higher mean scores suggest that most students improved. The negatively skewed distributions across the steps indicate that scores were concentrated at the higher end, meaning that more students achieved high performance levels. This pattern is especially evident in planning, carrying out, and checking, where many students scored in the upper range.

Overall, the findings show a clear and consistent improvement in students' problem-solving skills from pre-test to post-test across all four steps. These results suggest that AI pedagogy supports students' cognitive development by providing structured guidance aligned with Pólya's framework and allowing more individualized learning pacing. Compared to their initial performance, students demonstrated substantial improvement not only in basic understanding but also in higher-order skills such as strategy formulation, procedural execution, and solution verification. This indicates that AI pedagogy effectively enhances learning and supports students in achieving higher levels of proficiency across all steps of problem-solving. The improvement is further illustrated in Figure 2, which shows a clear increase in mean scores from pre-test to post-test across all four steps.

Table 8 presents the results of the independent-samples t-test comparing the pre-test problem-solving skills of students in the conventional and experimental groups based on George Pólya's four-step framework: understanding the problem, planning, carrying out, and checking the solution.

In terms of understanding the problem, the experimental group (M = 7.69, SD = 4.67), described as moderately skilled (MS), obtained a higher mean score than the conventional group (M = 5.10, SD = 2.78), described as slightly skilled (SS). The computed t-value of 2.528 with a p-value of 0.015 indicates a statistically significant difference between the two groups, which led to the rejection of the null hypothesis. This highlights the role of problem comprehension as a foundational skill in mathematical thinking (Abazov, 2016; Oliveri et al., 2017). In addition, Mefoh et al. (2017) emphasized that understanding the problem serves as the initial step that enables learners to identify relevant information and define task requirements.

On the other hand, for planning, carrying out, checking the solution, and overall problem-solving skills, the p-values (0.087, 0.244, 0.827, and 0.152, respectively) are all above the 0.05 significance level. This indicates no statistically significant differences between the two groups in these areas, leading to the acceptance of the null hypothesis. Laurens et al. (2018) and Simamora et al. (2017) noted that students commonly experience difficulty in organizing and applying solution strategies, particularly in non-routine problems. Likewise, Toll (2017) reported that learners often neglect the evaluation phase, reflecting weak metacognitive regulation during problem solving.

The results suggest that both groups were comparable in their overall problem-solving performance before the intervention. Establishing baseline equivalence is essential in quasi-experimental research to ensure that post-test differences can be attributed to the intervention rather than pre-existing conditions (Shadish et al., 2002).

Despite this similarity, a significant difference was observed in the understanding-the-problem step, where the experimental group performed better. This indicates that the experimental group initially had stronger comprehension skills, which is consistent with Abazov (2016) and Oliveri et al. (2017), who described understanding as a cognitive

process that supports problem interpretation. Mefoh et al. (2017) further stressed that comprehension influences how learners approach and structure problem situations.

However, both groups still demonstrated low performance in planning, carrying out, and evaluating solutions. Laurens et al. (2018) and Simamora et al. (2017) explained that students often struggle to convert understanding into systematic procedures, especially in unfamiliar problem contexts. This suggests a gap between comprehension and execution in problem-solving tasks.

The weak performance in checking the solution indicates limited metacognitive awareness among students. Toll (2017) emphasized that monitoring and evaluating one's solution are essential components of effective problem-solving, yet are often overlooked by learners.

Regarding instructional implications, Yadav (2023) noted that technology-enhanced learning can support more structured reasoning processes. Similarly, Fitriani et al. (2020) and Sholihah and Lastariwati (2020) emphasized that structured problem-solving activities help develop higher-order thinking skills among students.

Furthermore, Chakrabarty et al. (2023) highlighted that AI-supported learning environments can provide scaffolding and adaptive feedback, which may enhance students' reasoning and problem-solving processes.

Table 9 presents the results of the independent-samples t-test comparing the post-test problem-solving skills of students in the conventional and experimental groups across George Pólya's four steps of problem-solving: understanding the problem, planning how to solve it, carrying out the plan, and checking if the solution is correct.

In the post-test results, both groups demonstrated high levels of performance in several areas. For understanding the problem, the experimental group (M = 17.08, SD = 0.98) and the conventional group (M = 16.69, SD = 1.70) were both classified as Highly Skilled (HS), and the difference between them was not statistically significant (p = 0.458). Similarly, in planning how to solve the problem and checking if the solution is correct, both groups were categorized as Skilled (S), with p-values of 0.753 and 0.098, respectively, indicating no significant differences between the groups.

However, a statistically significant difference was found in carrying out the plan, where the experimental group (M = 15.27, SD = 2.255) outperformed the conventional group (M = 12.83, SD = 2.647), with a p-value of 0.001. This led to the rejection of the null hypothesis for this step. In addition, a significant difference was also observed in the overall problem-solving skills, where the experimental group (M = 61.50, SD = 8.39), categorized as Highly Skilled (HS), performed better than the conventional group (M = 56.55, SD = 8.64), categorized as Skilled (S), with p = 0.036.

The findings suggest that both Instructional approaches were effective in developing students' ability to understand mathematical problems, as shown by the high

performance and the absence of a significant difference between the two groups in this step. This indicates that both conventional and AI-assisted instruction can support learners in interpreting problems, consistent with the view that comprehension develops through structured explanations and exposure to learning content (Abazov, 2016; Oliveri et al., 2017). It also suggests that understanding may reach a comparable level once foundational knowledge is adequately developed.

The most notable finding of this study is the significant difference in the carrying out the plan step, where the experimental group performed better. This implies that students who were exposed to AI-assisted instruction demonstrated greater accuracy and effectiveness in executing problem-solving procedures. This step requires applying strategies correctly, and the result suggests that AI tools such as ChatGPT and Google Gemini may have supported learners through step-by-step guidance, feedback, and solution verification. This aligns with Chakrabarty et al. (2023), who noted that AI systems can enhance learning by providing real-time assistance and supporting engagement in complex tasks.

In contrast, no significant differences were found in the planning and checking steps, even though the experimental group obtained slightly higher mean scores. This suggests that while AI may support procedural execution, its influence on higher-order thinking skills such as planning strategies and evaluating solutions may still require more structured instructional design. Planning involves selecting appropriate approaches, while checking requires reflective thinking and self-assessment. As emphasized by Toll (2017), these metacognitive skills are often weak among learners and may not improve without deliberate instructional support.

The significant difference In overall problem-solving performance further supports the effectiveness of AI-assisted instruction. The experimental group's classification as Highly Skilled (HS), compared to the Skilled (S) level of the conventional group, indicates that AI integration contributed to better overall outcomes. This suggests that AI can enhance learning by making instruction more interactive, adaptive, and responsive to student needs (Yadav, 2023; Viberg et al., 2024). Through personalized feedback and guided learning experiences, AI may help students better understand and apply mathematical concepts.

From a theoretical standpoint, these results reinforce George Pólya's problem-solving framework, particularly the importance of the execution step in achieving correct solutions. While understanding and planning are essential, successful problem resolution ultimately depends on accurate execution. The stronger performance of the experimental group in this step suggests that AI can function as a cognitive support tool that helps bridge the gap between understanding and application.

Furthermore, these findings align with the idea that AI can serve as a learning scaffold by reducing cognitive load and guiding students through complex processes (Hwang & Tu, 2021). By assisting learners during execution, AI allows students to focus

more on reasoning and strategy use rather than procedural difficulties, which is especially important in mathematics where execution errors can affect final answers.

Overall, the results in Table 9 show that both instructional approaches improved students' problem-solving skills, but AI-assisted pedagogy had a stronger impact on procedural execution and overall performance. This highlights the potential of AI integration as an effective instructional approach for enhancing mathematical problem-solving, particularly in improving accuracy, efficiency, and application of solutions.

Table 10 presents the results of the paired-samples t-test comparing the pre-test and post-test problem-solving skills of the conventional group across George Pólya's four steps: understanding the problem, planning how to solve it, carrying out the plan, and checking if the solution is correct.

The results show significant improvement across all four steps after the implementation of conventional pedagogy. In understanding the problem, the mean increased from 5.10 (SS) to 16.79 (HS), with a significant difference ($t = -16.250$, $p = 0.000$). In planning how to solve it, the mean rose from 2.90 (NS) to 14.00 (S), while carrying out the plan improved from 2.55 (NS) to 12.83 (S), and checking if the solution is correct increased from 0.79 (NS) to 13.17 (S), all showing statistically significant results ($p = 0.000$).

Overall, the total mean score improved from 11.34 (NS) to 56.55 (S), with a significant t-value of -23.820, confirming a statistically significant improvement in students' problem-solving performance.

These results indicate that conventional pedagogy had a significant positive effect on students' problem-solving skills across all steps. The consistent rejection of the null hypothesis suggests that the improvements observed were due to the instructional intervention rather than chance.

The shift from not skilled and slightly skilled levels in the pre-test to skilled and highly skilled levels in the post-test reflects meaningful development in students' performance. This suggests that structured instruction and guided practice contributed to improved learning outcomes, consistent with Abazov (2016) and Oliveri et al. (2017), who emphasized that problem-solving develops through structured reasoning and guided learning.

Improvement in understanding the problem indicates that classroom instruction helped students better interpret and analyze mathematical tasks, supported by explanations and examples (Mefoh et al., 2017). Gains in planning how to solve it suggest improved strategic use, aligning with Laurens et al. (2018) and Simamora et al. (2017), who highlighted the importance of structured practice in developing problem-solving approaches.

Improvements in carrying out the plan and checking if the solution is correct reflect better procedural accuracy and evaluation skills, supported by practice and feedback. The improvement in checking is particularly important, as this step is often a weak area among students and requires explicit instruction (Toll, 2017).

From a theoretical perspective, the findings support George Pólya's framework, showing that all four steps of problem-solving can be developed through structured instruction. However, since the group reached only the skilled level and not the highly skilled level, it suggests that further enhancement may still be needed for higher-order problem-solving development.

Overall, Table 10 shows that conventional pedagogy significantly improves students' problem-solving skills across all four steps through structured learning, guided practice, and classroom interaction.

Table 11 presents the results of the paired-samples t-test comparing the pre-test and post-test problem-solving skills of the experimental group across George Pólya's four steps: understanding the problem, planning how to solve it, carrying out the plan, and checking if the solution is correct.

The results show substantial improvement across all four steps after the implementation of AI-assisted pedagogy. In understanding the problem, the mean increased from 7.69 (MS) to 17.01 (HS), with a significant difference ($t = -9.916$, $p = 0.000$). In planning how to solve it, the mean rose from 2.27 (NS) to 14.31 (S), while carrying out the plan improved from 3.31 (NS) to 15.27 (S), and checking if the solution is correct increased from 0.88 (NS) to 14.81 (S), all showing statistically significant results ($p = 0.000$).

Overall, the total mean score improved from 14.12 (NS) to 61.50 (HS), with a significant t-value of -21.490, indicating a statistically significant improvement in students' problem-solving performance.

These results indicate that AI-assisted pedagogy had a significant positive effect on students' problem-solving skills across all steps. The consistent rejection of the null hypothesis suggests that the improvements observed were due to the intervention rather than chance, supporting the effectiveness of AI integration in mathematics instruction.

The shift from not skilled and moderately skilled levels in the pre-test to skilled and highly skilled levels in the post-test reflects both quantitative and qualitative improvement in students' performance. As explained by Abazov (2016) and Oliveri et al. (2017), problem-solving is a higher-order cognitive process involving understanding, strategy use, and reasoning. The gains across all steps suggest that AI-supported instruction helped facilitate these processes.

The Improvement in understanding the problem suggests that AI tools helped students better interpret and analyze mathematical tasks. AI-generated explanations,

examples, and representations may have supported students in identifying relevant information and clarifying problem requirements. This aligns with Mefoh et al. (2017) and Sukariasih et al. (2020), who emphasized that meaningful learning occurs when learners connect new information with prior knowledge.

The Increase in planning how to solve it indicates that AI contributed to the development of students' strategic thinking. The shift from not skilled to skilled suggests that AI tools supported learners in organizing their approaches and selecting appropriate strategies. This is consistent with Laurens et al. (2018) and Simamora et al. (2017), who highlighted the role of structured and guided learning in improving problem-solving strategies.

The most notable Improvements were observed in carrying out the plan and checking if the solution is correct, which are essential for procedural accuracy and metacognitive regulation. The gains suggest that AI provided step-by-step guidance, practice opportunities, and feedback that improved accuracy and reduced errors. This is consistent with Chakrabarty et al. (2023) and Yadav (2023), who emphasized that AI supports engagement, accuracy, and critical thinking through interactive learning support.

The Improvement in checking if the solution is correct is particularly important, as this step is often weak among students. The transition from not skilled to skilled indicates improved reflective thinking and self-evaluation. As noted by Toll (2017), this skill often requires explicit scaffolding, which may have been supported through AI-assisted instruction.

From a theoretical perspective, the findings support George Pólya's problem-solving framework, showing that all four steps can be effectively developed through structured instruction. The consistent improvements suggest that AI-assisted pedagogy provides a learning environment that supports understanding, planning, execution, and evaluation.

Overall, the movement from not skilled (NS) to highly skilled (HS) highlights the strong impact of AI integration in enhancing learning outcomes. This aligns with Hwang & Tu (2021) and Viberg et al. (2024), who noted that AI can support learning by offering adaptive, personalized, and interactive experiences that reduce cognitive load and enhance higher-order thinking.

In summary, Table 11 provides strong evidence that AI-assisted pedagogy significantly improves students' problem-solving skills across all steps, demonstrating its effectiveness as an instructional support tool that promotes deeper understanding, strategic reasoning, and reflective learning.

Conclusions

Based on the findings of the study, the following conclusions were drawn:

1. Students initially had limited problem-solving skills, particularly in the higher-order steps of planning, execution, and evaluation.
2. Both conventional and AI-assisted pedagogy improved students' problem-solving skills, but AI-assisted pedagogy produced stronger overall gains.
3. AI-supported instruction was especially helpful in improving students' ability to carry out solution procedures accurately.
4. AI can serve as a cognitive support tool when it is used with teacher guidance, reflective activities, and clear problem-solving steps.
5. The integration of AI in mathematics instruction can contribute to students' critical thinking, reasoning, and metacognitive development, provided that it is used responsibly and does not replace independent thinking.

Recommendations

In light of the findings and conclusion of the study, the following recommendations are proposed:

1. Teachers may integrate AI tools such as ChatGPT and Google Gemini into mathematics lessons as guided supports for explanation, feedback, and solution verification.
2. Students should be taught to use AI responsibly by checking AI-generated outputs, explaining their reasoning, and avoiding dependence on ready-made answers.
3. Urdaneta City University may provide training, digital resources, and policy guidance to support responsible AI integration in teaching and learning.
4. Curriculum planners may include AI-supported problem-solving activities that promote reasoning, justification, reflection, and real-world mathematical application.
5. Future researchers may examine the long-term effects of AI-assisted instruction and explore its influence on creativity, collaboration, self-efficacy, and other subject areas.
6. Policymakers may consider AI-supported educational strategies that are consistent with Sustainable Development Goals 4 and 9, particularly in promoting inclusive, innovative, and technology-enhanced education.

Compliance with Ethical Standards

The study followed ethical guidelines to protect the respondents' rights, safety, and welfare and to ensure integrity, respect, and fairness throughout the research process. The researchers observed Republic Act No. 10173, or the Data Privacy Act of 2012, by protecting respondents' information through coding or anonymization. Access to the data was limited to the research team, and the data were not uploaded or publicly shared. Collected materials, including questionnaires and files, were stored securely and protected from loss, unauthorized access, and tampering and were handled only for research purposes. The study was conducted to contribute to improved instruction while minimizing risks to respondents. Participants were informed about the purpose of the study and their right to decline or withdraw from participation. Confidentiality and anonymity were maintained by avoiding the collection of unnecessary identifying information and ensuring that individual responses could not be traced to specific participants. The researchers committed to accurate recording, analysis, and reporting of

the data without fabrication, falsification, or misrepresentation. The study also disclosed the use of artificial intelligence tools, and AI-generated content was reviewed and verified by the researchers to maintain accuracy and accountability.

Acknowledgments

The researchers extend their heartfelt gratitude to the individuals listed below for their support and prayers, which greatly contributed to the success of this study.

Dr. Amihan April C. Mella-Alcazar, Acting University President, for her wholehearted support and commitment of research endeavors in the University;

Atty. Dar A. Diga, Chief Operating Officer, who has provided exceptional guidance in driving the university's advancement and improvement;

Dr. Josephine S. Lambinicio, Vice President for Academic Affairs, and chairperson of the panel, for her continual and unwavering thoughtfulness, for sharing, motivating, and giving insights to enriched the study;

Dr. Ryan Jayson V. Delos Reyes, Dean of the College of Teacher Education, for his support and encouragement;

Dr. Cristina F. Solomon, member of the panel, for her continuous guidance, support, and encouragement, as well as for generously sharing her expertise to enhance and ensure the success of this study;

Dr. Prescila I. Marcelo, member of the panel, for her unceasing guidance, support, encouragement and sharing of her expertise to improve and make this study thrive;

Sir Vener Abiett L. Castañaga, research instructor, for his professional guidance, valuable suggestions, and unwavering support that greatly contributed to the successful completion of this research;

Dr. Sherelle Lou S. Icutan, adviser, for her expert guidance, invaluable recommendations, and support that significantly contributed to the successful completion of this research study;

Mdm. Graciane Joy De Guzman, critic reader, for patiently reading and checking the format and grammar of this research study;

Sir Rolando D. De Guzman, statistician, for helping the researchers by providing the formula, analyzing and tabulating the data of this study;

The researchers' families, for their love, encouragement, and sacrifices, both moral and financial;

The researchers' friends, for their companionship and inspiration throughout the conduct of this study;

The respondents, for their time, cooperation, and support throughout the conduct of the study; and

Above all, Almighty God, for the blessings, knowledge, and strength that enabled the researchers to complete this research.

REFERENCES

Abazov, R. (2016). Teaching and learning in higher education: A practical guide. Palgrave Macmillan. https://doi.org/10.1057/978-1-137-54856-3

Alexander, B., Ashford-Rowe, K., Barajas-Murphy, N., Dobbin, G., Knott, J., McCormack, M., Pomerantz, J., Seilhamer, R., & Weber, N. (2024). EDUCAUSE horizon report: Teaching and learning edition. EDUCAUSE. https://www.educause.edu/horizon-report

Al-Momani, M. M., & Ramayah, T. (2024). Artificial intelligence adoption in education: A systematic review. Education and Information Technologies. https://doi.org/10.1007/s10639-023-11945-7

Chakrabarty, T., et al. (2023). AI for creativity and learning. Computers in Human Behavior, 140, 107583. https://doi.org/10.1016/j.chb.2022.107583

Chaudhry, M. A., & Kazim, E. (2022). Artificial intelligence in education. Education Sciences, 12(1), 38. https://doi.org/10.3390/educsci12010038

Chiu, T. K. F., et al. (2022). AI literacy in education. Educational Technology & Society, 25(1), 1–15. https://www.jstor.org/stable/48647028

Crompton, H., & Burke, D. (2023). Artificial intelligence in education. British Journal of Educational Technology, 54(2), 342–356. https://doi.org/10.1111/bjet.13207

Dignum, V. (2021). Responsible artificial intelligence. Springer. https://doi.org/10.1007/978-3-030-30371-6

Došilović, F. K., et al. (2018). Explainable AI. MIPRO Proceedings, 210–215. https://doi.org/10.23919/MIPRO.2018.8400040

Fitriani, N., et al. (2020). Problem-solving skills in mathematics. Journal of Physics: Conference Series, 1521, 032012. https://doi.org/10.1088/1742-6596/1521/3/032012

Goralski, M. A., & Tan, T. K. (2020). Artificial intelligence and sustainable development. The International Journal of Management Education, 18(1), 100330. https://doi.org/10.1016/j.ijme.2019.100330

Hiebert, J., & Grouws, D. A. (2007). Mathematics teaching effects. In Second handbook of research on mathematics teaching and learning. https://doi.org/10.1007/978-94-010-9620-9

Hwang, G. J., & Tu, Y. F. (2021). AI roles in education. Educational Technology Research and Development, 69, 1–4. https://doi.org/10.1007/s11423-020-09817-3

Jarrahi, M. H. (2018). AI and future of work. Business Horizons, 61(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Jeste, D. V., et al. (2020). AI and sustainable development. Nature Machine Intelligence, 2, 1–3. https://doi.org/10.1038/s42256-020-0180-9

Kabudi, T. (2022). AI ethics in education. Computers & Education, 182, 104463. https://doi.org/10.1016/j.compedu.2022.104463

Laurens, T., et al. (2018). Problem-solving abilities. Journal of Physics: Conference Series, 1028, 012078. https://doi.org/10.1088/1742-6596/1028/1/012078

Lin, S., et al. (2022). Truthfulness in AI. https://arxiv.org/abs/2207.05221

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable AI: A review of machine learning interpretability methods. Entropy, 23(1), 18. https://doi.org/10.3390/e23010018

Lithner, J. (2017). Mathematical reasoning. Journal of Mathematical Behavior, 48, 1–15. https://doi.org/10.1016/j.jmathb.2017.09.003

Long, D., & Magerko, B. (2020). AI literacy. https://doi.org/10.1145/3313831.3376727

Mefoh, P. C., Nwoke, M. B., Chukwuorji, J. C., & Chijioke, A. O. (2017). Effect of cognitive style and gender on adolescents' problem solving ability. Thinking Skills and Creativity, 25, 47–52. https://doi.org/10.1016/j.tsc.2017.03.002

Miao, F., et al. (2021). AI and education: Guidance for policy-makers. UNESCO. https://unesdoc.unesco.org

Oliveri, M. E., et al. (2017). Problem solving. Educational Measurement, 36(2), 1–10. https://doi.org/10.1111/emip.12121

Polya, G. (1973). How to solve it. Princeton University Press. https://press.princeton.edu

Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson. https://aima.cs.berkeley.edu

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Simamora, R. E., Sidabutar, D. R., & Surya, E. (2017). Improving learning activity and students' problem solving skill through problem based learning (PBL) in junior high school. International Journal of Sciences: Basic and Applied Research, 33(2), 321–331. https://www.gssrr.org/index.php/JournalOfBasicAndApplied/article/view/7368

Sholihah, T. M., & Lastariwati, B. (2020). Problem based learning to increase competence of critical thinking and problem solving. Journal of Education and Learning (EduLearn), 14(1), 148–154. https://doi.org/10.11591/edulearn.v14i1.13772

Sukariasih, L., Tahang, L., Nursalam, L. O., & Fayanto, S. (2020). Description of physics problem-solving in the topic of static fluid: Case study of physics education in Halu Oleo University. Universal Journal of Educational Research, 8(10), 4568–4579. https://doi.org/10.13189/ujer.2020.081025

Sullivan, P., et al. (2015). Challenging tasks. Mathematics Education Research Journal, 27(1), 1–15. https://doi.org/10.1007/s13394-014-0128-8

Toll, C. A. (2017). A problem-solving model for literacy coaching practice. The Reading Teacher, 70(4), 413–421. https://doi.org/10.1002/trtr.1532

United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development. United Nations. https://sdgs.un.org/2030agenda

Viberg, O., Kizilcec, R. F., Wise, A. F., Jivet, I., & Nixon, N. (2024). Advancing equity and inclusion in educational practices with AI-powered educational decision support systems (AI-EDSS). British Journal of Educational Technology, 55(5), 1974–1981. https://doi.org/10.1111/bjet.13507

Vinuesa, R., et al. (2020). AI and SDGs. Nature Communications, 11, 233. https://doi.org/10.1038/s41467-019-14108-y

von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. Philosophy & Technology, 34(4), 1607–1622. https://doi.org/10.1007/s13347-021-00477-0

Yadav, G. G. (2023). Scaling evidence-based instructional design expertise through large language models. arXiv. https://doi.org/10.48550/arXiv.2306.01006

Zawacki-Richter, O., et al. (2019). AI in higher education. International Journal of Educational Technology, 16, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. Smart Learning Environments, 11, Article 28. https://doi.org/10.1186/s40561-024-00316-7

Zhai, X. (2022). AI in STEM education. Computers & Education, 183, 104520. https://doi.org/10.1016/j.compedu.2022.104520

Zhou, M., et al. (2024). AI overreliance. Educational Psychology Review. https://doi.org/10.1007/s10648-023-09789-2

APA Citation:

Altre, A. J. T., Baoing, H. L., Collado, J. M. C., Lomboy, J. R. E., & Ulnagan, A. D. (2026). EFFECTS OF ARTIFICIAL INTELLIGENCE (AI) ON STUDENT PROBLEM-SOLVING SKILLS. Ignatian International Journal for Multidisciplinary Research, 4(10), 512–542. https://doi.org/10.5281/zenodo.23593091

Corresponding author: *arlynjoycealtre11@gmail.com*