

AI-Assisted Mathematics Learning: Jeneponto Salt Production Ethnomathematics and Junior High Students' Problem- Solving Ability

Syaharuddin^{1*}, Sumiati S.², Nurqiyamah Hamid³, Hasmawati⁴, Muh. Usman Taufik⁵,
Adryan Putra Pratama⁶

^{1*2,3,5}Institut Turatea Indonesia, Jeneponto, Indonesia;

⁴Education and Culture Office of Jeneponto, Indonesia

⁶Department of Informatics and Computer Engineering, Universitas Negeri Makassar,
Indonesia

ARTICLE INFO

Article History:

Received :

Revised :

Accepted :

Keyword:

AI chatbot;
ethnomathematics; salt
production; mathematical
problem-solving ability;
mathematics learning

*Correspondence Author:
syahar2@gmail.com

ABSTRACT

Mathematical problem-solving ability among junior high school students remains a challenge in mathematics learning, while the integration of artificial intelligence with local cultural contexts has received limited attention when examined simultaneously. This study aimed to analyze the improvement in the mathematical problem-solving ability of ninth-grade students through Problem-Based Learning (PBL) assisted by an AI chatbot in the context of salt-production ethnomathematics in Jeneponto, both overall and at each problem-solving stage. The study employed a quantitative approach with a one-group pretest-posttest design involving 30 ninth-grade students from a junior high school in Jeneponto Regency who were selected purposively. Mathematical problem-solving ability was measured using an ethnomathematics-based contextual essay test and analyzed using the Shapiro-Wilk test, paired-samples t-test, Cohen's dz, and N-Gain. The mean score increased from 33.13 on the pretest to 80.80 on the posttest. This increase was statistically significant, $t(29) = 17.419$, $p < .001$, with a very large effect size (Cohen's $dz = 3.18$). The N-Gain of 0.707 indicated a high level of improvement. Analysis by problem-solving stage showed improvement across all indicators, with the largest absolute increase occurring at the carrying out the solution stage (D3), followed by the checking the result stage (D4). The findings indicate that the integration of PBL, AI chatbot support, and salt-production ethnomathematics has the potential to support contextual mathematics learning and the development of students' problem-solving ability.

ABSTRAK

Kemampuan pemecahan masalah matematis siswa sekolah menengah pertama masih menjadi tantangan dalam pembelajaran matematika, sementara integrasi kecerdasan buatan dengan konteks budaya lokal masih terbatas ketika dikaji secara simultan. Penelitian ini bertujuan menganalisis peningkatan kemampuan pemecahan masalah matematis siswa kelas IX melalui Problem-Based Learning (PBL) berbantuan

Kata Kunci:

kecerdasan buatan;
etnomatematika; produksi
garam; pemecahan masalah
matematis; Problem-Based
Learning

chatbot kecerdasan buatan dalam konteks etnomatematika produksi garam di Jeneponto, baik secara keseluruhan maupun pada setiap tahap pemecahan masalah. Penelitian menggunakan pendekatan kuantitatif dengan desain one-group pretest–posttest yang melibatkan 30 siswa kelas IX dari salah satu sekolah menengah pertama di Kabupaten Jeneponto yang dipilih secara purposif. Kemampuan pemecahan masalah matematis diukur menggunakan tes uraian kontekstual berbasis etnomatematika dan dianalisis melalui uji Shapiro–Wilk, uji t berpasangan, ukuran efek Cohen’s d , dan N-Gain. Rata-rata skor meningkat dari 33,13 pada pretest menjadi 80,80 pada posttest, dengan peningkatan yang signifikan dan ukuran efek yang sangat besar. Nilai N-Gain sebesar 0,707 menunjukkan peningkatan pada kategori tinggi. Analisis berdasarkan tahap pemecahan masalah menunjukkan peningkatan pada seluruh indikator, dengan peningkatan absolut terbesar pada tahap melaksanakan strategi (D3), diikuti tahap memeriksa hasil (D4). Temuan menunjukkan bahwa integrasi PBL, dukungan chatbot kecerdasan buatan, dan etnomatematika produksi garam berpotensi mendukung pembelajaran matematika kontekstual dan pengembangan kemampuan pemecahan masalah matematis siswa.

How to Cite:

Syahrudin, S., Sumiati, S., Hamid, N., Hasmawati, H., Taufik, M. U., & Pratama, A. P. (2026). AI-Assisted Mathematics Learning: Jeneponto Salt Production Ethnomathematics And Junior High Students’ Problem-Solving Ability. *Paedagogia: Jurnal Pendidikan*, 15(2), 267-288. <https://doi.org/10.24239/pdg.Vol15.Iss2.1478>

INTRODUCTION

Mathematical problem-solving ability is a central competence in mathematics learning because students need to understand situations, formulate mathematical relationships, select appropriate strategies, carry out solutions, and evaluate the results rather than merely obtain final answers. This need is reflected in the PISA 2025 results, in which Indonesian students’ mean mathematics score was 364, compared with the OECD average of 463, and only 17% of Indonesian students reached at least Level 2 proficiency, compared with 65% across OECD countries. These results reinforce the need to strengthen students’ ability to interpret situations, represent information mathematically, plan solution strategies, and evaluate the reasonableness of their results.¹

One persistent difficulty in mathematical problem solving occurs when students must transform contextual information into mathematical representations or models. Research on Indonesian students’ algebraic word-problem solving has shown difficulties in mathematizing situations, particularly

¹ OECD, *PISA 2025 Results (Volume I): Future-Ready Students* (OECD Publishing, 2026), <https://doi.org/10.1787/73451bc5-en>.

in constructing appropriate mathematical models². Problem solving therefore involves more than applying a known procedure after an equation has been provided; it requires students to formulate an approach, coordinate strategies and procedures, and evaluate the resulting solution^{3,4,5}. These demands make stage-based analysis relevant for identifying not only whether students improve, but also which parts of the problem-solving process remain challenging.

Ethnomathematics provides one way to connect formal mathematical reasoning with students' experiences and community practices. Cultural practices can provide meaningful situations from which students identify quantities, relationships, patterns, and strategies before expressing them in formal mathematical language. Previous evidence indicates positive effects of ethnomathematics-based learning on mathematical problem-solving ability and mathematical literacy, although the magnitude of these effects can vary according to the cultural context, learning design, and instructional materials^{6,7}. Thus, the educational value of ethnomathematics does not lie simply in placing cultural objects or activities into mathematics problems, but in deliberately transforming culturally situated practices into mathematical activities.

In Jeneponto, salt production provides an authentic context involving quantities of production, land area, prices, product quality, labor, and income that can be connected to mathematical modeling. Previous research in Jeneponto demonstrated that salt production can function as a source of contextual problems for developing mathematical problem-solving ability through a learning design integrating problem-based learning and a STEAM perspective⁸. That study also showed that the checking stage remained relatively weak after learning,

² Al Jupri and Paul Drijvers, "Student Difficulties in Mathematizing Word Problems in Algebra," *Eurasia Journal of Mathematics, Science and Technology Education* 12, no. 9 (2016): 2481–502, <https://doi.org/10.12973/eurasia.2016.1299a>.

³ Manuel Santos-Trigo, "Problem Solving in Mathematics Education: Tracing Its Foundations and Current Research-Practice Trends," *ZDM—Mathematics Education* 56, no. 2 (2024): 211–22, <https://doi.org/10.1007/s11858-024-01578-8>.

⁴ Septian Dika Maulana and Cecep Anwar Hadi Firdos Santosa, "Analisis Kemampuan Pemecahan Masalah Matematika Siswa SMA Pada Kurikulum Merdeka Dan Kurikulum 2013 Berdasarkan Teori Polya," *Jurnal Pendidikan Matematika: Judika Education* 7, no. 1 (2024): 93–105, <https://doi.org/10.31539/judika.v7i1.8341>.

⁵ Francis Aposika, "Investigating the Relationship between Problem-Solving Skills and Academic Performance in STEM Subjects among Secondary School Students," *Discover Education* 5 (2026): 425, <https://doi.org/10.1007/s44217-026-01487-w>.

⁶ Sri Apriatni et al., "The Influence of Ethnomathematics Based Learning on Mathematics Problem-Solving Ability: A Meta-Analysis," *Jurnal Pendidikan Matematika (JUPITEK)* 5, no. 1 (2022): 23–33, <https://doi.org/10.30598/jupitekvol5iss1pp23-33>.

⁷ Ryan Angga Pratama and Tuğba Yanpar Yelken, "Effectiveness of Ethnomathematics-Based Learning on Students' Mathematical Literacy: A Meta-Analysis Study," *Discover Education* 3 (2024): 202, <https://doi.org/10.1007/s44217-024-00309-1>.

⁸ Syaharuddin et al., "Students' Mathematical Problem-Solving Ability Based on Polya's Stages: A Study of Local Salt-Production Ethnomathematics-Based Learning in Jeneponto," *Journal of Local Wisdom in Education* 2, no. 1 (2026): 38–56, <https://doi.org/10.65508/01s42c69>.

suggesting that familiarity with a cultural context alone does not necessarily ensure that students sustain their reasoning through verification. Research using ethnomathematics-based contextual learning has likewise indicated the potential of cultural contexts to support mathematical learning⁹. These findings suggest a need for additional instructional support that can accompany students across the problem-solving process, particularly when they need to plan, execute, and verify their solutions.

Artificial intelligence chatbots offer a potential form of such process-oriented support. Research on educational chatbots and AI-assisted mathematics learning indicates that chatbots can function as tutoring or dialogic support, provide immediate feedback, respond to students' learning needs, and assist mathematical thinking¹⁰¹¹¹²¹³. Research on ChatGPT and large language models further highlights opportunities for personalization and interaction, while systematic reviews emphasize that the educational value of AI depends on how the technology is integrated into instruction¹⁴¹⁵¹⁶. Therefore, the relevant question is not simply whether AI can provide mathematical assistance, but how its assistance can be structured so that students remain actively engaged in mathematical reasoning.

This consideration becomes particularly important when AI chatbots are integrated with problem-based learning. Previous research has reported positive learning outcomes from problem-based learning assisted by AI chatbots, while research in junior high school mathematics in Jeneponto has examined an AI

⁹ Saprina Maulida and Lisa Dwi Afri, "Implementasi Model Pembelajaran Contextual Teaching and Learning (CTL) Berbasis Etnomatematika Terhadap Kemampuan Literasi Numerasi Siswa," *Jurnal Pendidikan Matematika: Judika Education* 8, no. 4 (2025): 626–36, <https://doi.org/10.31539/apdh6c08>.

¹⁰ Antonín Jančařík et al., "Using AI Chatbot for Math Tutoring," *Journal of Education Culture and Society* 14, no. 2 (2023): 285–96, <https://doi.org/10.15503/jecs2023.2.285.296>.

¹¹ Chukwunonso W. Okonkwo and Abejide Ade-Ibijola, "Chatbots Applications in Education: A Systematic Review," *Computers and Education: Artificial Intelligence* 2 (2021): 100033, <https://doi.org/10.1016/j.caeai.2021.100033>.

¹² Nguyen Van Doc et al., "Teaching Mathematics with the Assistance of an AI Chatbot to Enhance Mathematical Thinking Skills for High School Students," *International Journal of Current Science Research and Review* 6, no. 12 (2023): 8574–80, <https://doi.org/10.47191/ijcsrr/V6-i12-102>.

¹³ Y. P. W. Laja et al., "Chatbot Artificial Intelligence (AI) to Supporting Differentiated Mathematics Learning," *Jurnal Pendidikan MIPA* 26, no. 1 (2025): 118–30, <https://doi.org/10.23960/jpmipa.v26i1.pp118-130>.

¹⁴ Enkelejda Kasneci et al., "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education," *Learning and Individual Differences* 103 (2023): 102274, <https://doi.org/10.1016/j.lindif.2023.102274>.

¹⁵ Muhammad Turmuzi et al., "ChatGPT in School Mathematics Education: A Systematic Review of Opportunities, Challenges, and Pedagogical Implications," *Teaching and Teacher Education* 170 (2026): 105286, <https://doi.org/10.1016/j.tate.2025.105286>.

¹⁶ Chung Kwan Lo et al., "The Influence of ChatGPT on Student Engagement: A Systematic Review and Future Research Agenda," *Computers & Education* 219 (2024): 105100, <https://doi.org/10.1016/j.compedu.2024.105100>.

chatbot within an NHT-based PBL model¹⁷¹⁸. At the same time, studies and reviews have identified concerns regarding inaccurate AI-generated information, dependence on AI-generated answers, reduced cognitive engagement under some conditions, and the need for teacher supervision and critical evaluation¹⁹²⁰²². These findings support a pedagogical position in which a chatbot is used as structured assistance within learning rather than as a substitute for students' mathematical decision making.

Taken together, the existing literature points to three partially connected lines of research. First, ethnomathematics can provide meaningful cultural contexts for mathematical learning and problem solving. Second, AI chatbots can provide interactive assistance during learning, but their educational value depends on pedagogical design and supervision. Third, problem-based learning can organize students' engagement with contextual problems and has been combined with AI chatbot support in mathematics education. However, based on the literature reviewed for this study, these strands have not yet been sufficiently connected in the specific context of Jeneponto salt-production ethnomathematics. In particular, the existing Jeneponto salt-production study examined problem solving through a culturally grounded PBL–STEAM design, whereas the existing local AI study examined chatbot-supported mathematics learning through an NHT-based PBL approach²³²⁴. The intersection of salt-production ethnomathematics, AI chatbot support, and stage-level analysis of mathematical problem solving therefore remains insufficiently documented.

The novelty of the present study lies in connecting these components within a single problem-solving learning design: salt production is positioned as

¹⁷ Setyo Ajie Wibowo and Intan Faizah, "Pengaruh Model Pembelajaran Berbasis Masalah Berbantuan Chatbot AI Terhadap Higher Order Thinking Skill," *Jurnal Nyamadassana: Jurnal Penelitian Pendidikan, Sosial Dan Keagamaan* 2, no. 2 (2023): 73–85, <https://doi.org/10.59291/jnd.v2i2.36>.

¹⁸ Husriani Husain et al., "Efektivitas Chatbot-AI Dalam Pembelajaran Matematika Melalui Model NHT Berbasis PBL Di SMPN 3 Binamu," *Mandalika Mathematics and Education Journal* 8, no. 1 (2026): 193–208, <https://doi.org/10.29303/jm.v8i1.11074>.

¹⁹ Kasneci et al., "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education."

²⁰ Daniel Desta Avia Difaputra Purnomo et al., "Islamic Education: Examining AI Use, Reliance, and Perceived Validity among Tarbiyah Students," *Paedagogia: Jurnal Pendidikan* 15, no. 1 (2026): 173–90, <https://doi.org/10.24239/pdg.Vol15.Iss1.1436>.

²¹ Amirhossein Tilili et al., "What If the Devil Is My Guardian Angel: ChatGPT as a Case Study of Using Chatbots in Education," *Smart Learning Environments* 10, no. 1 (2023): 15, <https://doi.org/10.1186/s40561-023-00237-x>.

²² Meyla Tri Fadila et al., "Mathematics in the AI Era: The Effectiveness of ChatGPT in Helping Students Solve HOTS Problems," *AlphaMath: Journal of Mathematics Education* 11, no. 2 (2025): 191–204, <https://doi.org/10.30595/alphamath.v11i2.28261>.

²³ Syaharuddin et al., "Students' Mathematical Problem-Solving Ability Based on Polya's Stages: A Study of Local Salt-Production Ethnomathematics-Based Learning in Jeneponto."

²⁴ Husain et al., "Efektivitas Chatbot-AI Dalam Pembelajaran Matematika Melalui Model NHT Berbasis PBL Di SMPN 3 Binamu."

the source of contextual mathematical problems, PBL organizes students' problem-solving activities, Pólya's stages provide a structure for examining changes in the problem-solving process, and the AI chatbot is positioned as an interactive support mechanism rather than a provider of final solutions. The study therefore examines not only overall changes in students' mathematical problem-solving ability but also changes at the stages of understanding the problem (D1), planning a strategy (D2), carrying out the solution (D3), and checking the result (D4)²⁵. Based on this framework, the study addresses two questions: (1) how does students' mathematical problem-solving ability change after participating in AI chatbot-assisted learning in the context of salt-production ethnomathematics in Jeneponto; and (2) how large is the improvement in students' mathematical problem-solving ability overall and at each problem-solving stage following the learning intervention?

METHODS

This study employed a quantitative approach with a one-group pretest–posttest design. The research design consisted of an initial measurement (O_1), AI chatbot-assisted learning (X), and a final measurement (O_2). This design was used to identify changes in students' mathematical problem-solving ability after they participated in the learning sequence. Because the study did not include a comparison group, changes in scores were interpreted as changes occurring after the learning intervention and were not used to conclude that the AI chatbot alone caused the improvement in students' ability. The study was conducted from April to September 2026, while the learning implementation took place from July 20 to August 20, 2026, during the first semester of the 2026/2027 academic year. The study consisted of 10 meetings: one meeting for the pretest, eight learning meetings, and one meeting for the posttest. Each learning meeting lasted 2×40 minutes²⁶.

The participants were 30 ninth-grade students from one junior high school in Jeneponto Regency, South Sulawesi, who participated as an intact class. The students were 13–14 years old, consisting of 11 male and 19 female students. The class was selected purposively based on the teacher's readiness to implement the learning activities and the availability of supporting facilities at the school. All students completed the entire research sequence, including the pretest, eight learning meetings, and posttest; therefore, all 30 students' data were included in the analysis.

The learning focused on the topic of Systems of Linear Equations in Two Variables (SPLDV) by integrating the ethnomathematical context of salt production and sales in Jeneponto, particularly salt-production activities in Arungkeke Village. The learning was implemented using a Problem-Based

²⁵ Pólya, *How to Solve It: A New Aspect of Mathematical Method*.

²⁶ William R. Shadish et al., *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Houghton Mifflin, 2002).

Learning (PBL) approach, while the AI chatbot was used as an interactive companion during the problem-solving process. PBL served as the main learning framework through which students were oriented to contextual problems, organized for investigation, guided in developing solutions, and engaged in presenting and evaluating their problem-solving processes. The learning activities were designed so that students did not merely use the cultural context as an illustration but identified information from salt-production practices, determined variables, formulated mathematical models, solved the models, and interpreted the results. The problem-solving process was directed through four stages: understanding the problem (D1), planning a strategy (D2), carrying out the solution (D3), and checking the result (D4), based on Pólya's stages of problem solving²⁷.

The AI chatbot was configured using the same instructional framework throughout the study and functioned primarily as an interactive companion rather than a provider of final answers. Assistance was intended to be provided progressively through three levels: open-ended questions to encourage students to identify the problem, directed questions to help students determine the next step, and one concrete hint when students continued to experience difficulties. If students remained unable to proceed after the staged assistance, the teacher provided further support. The chatbot was designed to support students' thinking processes across stages D1–D4, with the intended interaction protocol emphasizing prompts and hints rather than direct provision of complete solutions. The chatbot was used only during learning activities and was not permitted during either the pretest or posttest. Interaction records were retained as supporting documentation to examine how the chatbot was actually used during the learning process.

Classroom implementation was also documented through structured observations conducted during the eight learning meetings. The observation sheet consisted of 22 indicators covering the general learning process, PBL activities, the D1–D4 problem-solving stages, scaffolding, AI chatbot use, collaboration, ethnomathematical context, communication, evaluation, time management, student independence, contextual interpretation, and meeting-specific learning activities. Each indicator was rated on a four-point scale to describe the extent to which the planned learning activities were implemented. These observations were used as supporting evidence of the implementation of the learning activities and were analyzed descriptively across the eight learning meetings.

The main study data consisted of students' mathematical problem-solving ability, measured using a contextual essay test. The instrument consisted of two items, each assessing the problem-solving process based on the four Pólya stages: understanding the problem (D1), planning a strategy (D2), carrying out the solution (D3), and checking the result (D4). Each stage was scored from 0 to 4, giving a maximum score of 16 for each item and 32 for the entire instrument. The

²⁷ Pólya, *How to Solve It: A New Aspect of Mathematical Method*.

scores were then converted to a 0–100 scale. The pretest and posttest used equivalent isomorphic forms rather than identical items, maintaining comparable indicators, context, and difficulty while using different numerical presentations. Classroom observations and documentation were used as supporting data to provide information about the implementation of the learning activities, whereas the analysis of changes in problem-solving ability was based on the test results.

The content validity of the instrument was examined through expert judgment by three validators, consisting of two mathematics education lecturers and one junior high school mathematics teacher. The validators' ratings were analyzed using Aiken's V coefficient to quantify expert judgments of the instrument's content validity. The Aiken's V values ranged from 0.778 to 1.000, with a mean value of 0.870. The validators' assessments and feedback were used to make minor revisions, particularly regarding clarity and precision of wording, before the instrument was administered in the trial. The instrument trial involved 30 students from another school who were not included in the study participants. The internal reliability of the instrument, based on the eight scoring components, yielded a Cronbach's alpha coefficient of 0.805. This coefficient was used to indicate the internal consistency of the instrument scores in accordance with the concept of the alpha coefficient developed by Cronbach²⁸²⁹.

Data analysis was conducted in several stages. Descriptive statistics were used to describe the pretest and posttest scores, including the mean, standard deviation, minimum, and maximum scores. The normality of the distribution of difference scores between the posttest and pretest was examined using the Shapiro–Wilk test. Changes in scores before and after the learning intervention were analyzed using a paired-samples t-test. The magnitude of the change was also reported using Cohen's d for paired data (dz), calculated based on the mean difference score and its standard deviation. In addition, improvement in ability was analyzed using normalized gain (N-Gain), calculated for each student by considering the initial score and the maximum possible score, and then summarized at the group level. The analysis was also conducted across the four problem-solving stages (D1–D4) to examine changes at each stage of the problem-solving process³⁰³¹³².

²⁸ Lewis R. Aiken, "Three Coefficients for Analyzing the Reliability and Validity of Ratings," *Educational and Psychological Measurement* 45, no. 1 (1985): 131–42, <https://doi.org/10.1177/0013164485451012>.

²⁹ Lee J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika* 16, no. 3 (1951): 297–334, <https://doi.org/10.1007/BF02310555>.

³⁰ S. S. Shapiro and M. B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)," *Biometrika* 52, nos. 3–4 (1965): 591–611, <https://doi.org/10.1093/biomet/52.3-4.591>.

³¹ Daniël Lakens, "Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-Tests and ANOVAs," *Frontiers in Psychology* 4 (2013): 863, <https://doi.org/10.3389/fpsyg.2013.00863>.

³² Richard R. Hake, "Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses," *American Journal of Physics* 66, no. 1 (1998): 64–74, <https://doi.org/10.1119/1.18809>.

The study also considered research ethics by obtaining permission from the relevant authorities, approval from the school principal, and consent from the teacher, students, and students' parents or guardians. Students' identities were not included in the reporting of the research findings. During the use of the AI chatbot, students were not instructed to enter personal data or information that could identify them. All research data were used solely for scientific analysis and reporting.

RESULTS

Classroom observations indicated that the planned learning activities were implemented at a generally high level across the eight learning meetings. The implementation score increased from 84.09% in the first and second meetings to 85.23% in the third and fourth meetings, 86.36% in the fifth and sixth meetings, and 87.50% in the seventh and eighth meetings. The mean implementation score across the eight meetings was 85.80%, categorized as very good. The observations documented that the ethnomathematical salt-production context was consistently used to connect SPLDV with real-life situations, while students' use of the AI chatbot became progressively more directed during the learning sequence. Early meetings showed that students still needed guidance in formulating specific prompts and organizing their investigation, whereas later meetings showed more purposeful use of the chatbot for clarifying steps, checking procedures, and supporting reflection.

Interaction records provided additional evidence that students actually used the chatbot during the learning activities. Across the recorded interactions, students engaged with the chatbot at different problem-solving stages: identifying known and unknown information at D1, clarifying coefficients and selecting strategies at D2, receiving prompts to continue procedural steps at D3, and checking solutions by substituting values back into the equations at D4. The records therefore show that the chatbot was used as part of the problem-solving process rather than only as a source of final answers. However, the interaction records also showed occasional responses in which the chatbot provided explicit mathematical representations or intermediate calculations. Thus, the observed use was predominantly supportive and process-oriented, although the intended level of scaffolding was not maintained uniformly across all interactions.

Figure 1 displays two screenshots of a chat interface showing interactions between a student and a chatbot (Tanya ChatGPT) during a problem-solving activity. The chatbot provides guidance and feedback on the student's work.

Left Screenshot:

- Student:** Kami sudah menyusun dua persamaan, Tolong tanyakan kepada kami apa yang sebaiknya kami perhatikan pada kedua persamaan ini.
- Chatbot:** Bagus, Kelompok 1. Kalian sudah menyusun dua persamaan dan memilih metode eliminasi. Sekarang, coba perhatikan apakah setiap angka dan koefisien pada kedua persamaan benar benar sesuai dengan jumlah karung yang dibeli serta total pembayaran masing-masing pembeli? Bagian mana yang ingin kalian periksa terlebih dahulu?
- Student:** menurut saya sudah sesuai
- Chatbot:** Bagus, Kelompok 1. Kalau kalian yakin sudah sesuai, sekarang lanjutkan dengan rencana metode eliminasi yang kalian pilih. Menurut kalian, variabel mana yang akan kalian hilangkan terlebih dahulu, dan apa yang perlu dilakukan pada kedua persamaan agar variabel tersebut dapat dihilangkan?
- Student:** (Handwritten notes showing equations and calculations)
- Chatbot:** Bagus, Kelompok 1. Kalian sudah sampai pada nilai y melalui eliminasi. Sekarang, satu langkah lagi berikutnya: gunakan nilai y yang sudah kalian peroleh ke salah satu persamaan awal untuk mencari nilai x . Menurut kalian, lebih mudah menggunakan persamaan Pak Sabang atau persamaan Pak Rahman? Mengapa?
- Student:** (Handwritten notes showing calculations for x)
- Chatbot:** Bagus, Kelompok 1. Kalian sudah mencoba memeriksa persamaan kedua dengan cara substitusi. Namun, coba bandingkan kembali baris kedua perhitungan kalian dengan persamaan kedua di atasnya pada persamaan kedua, berapa koefisien y yang seharusnya ikut dikalikan dengan nilai y ?
- Student:** (Handwritten notes showing calculations for x)
- Chatbot:** Bagus, Kelompok 1, kalian sudah menyadari bahwa pemeriksaan perlu dilakukan pada kedua persamaan. Namun, coba perhatikan lagi persamaan kedua: di sana tertulis $3x + 4y$, sedangkan pada baris substitusi kalian menulis $2(Rp50.000)$. Menurut kalian, angka berapa yang seharusnya digunakan untuk mengalikan nilai y pada persamaan kedua?

Right Screenshot:

- Student:** Benar, Kelompok 1. Kalian memilih variabel x untuk dihilangkan, dan koefisien x pada kedua persamaan memang sudah sama. Sekarang pikirkan: operasi apa yang dapat kalian lakukan pada kedua persamaan agar kedua suku x saling menghilangkan?
- Chatbot:** mengurangi persamaan pertama dengan persamaan kedua
- Student:** (Handwritten notes showing equations and calculations)
- Chatbot:** Bagus, Kelompok 1. Mengurangkan kedua persamaan merupakan langkah yang sesuai dengan rencana eliminasi variabel x . Sekarang, coba kalian lakukan pengurangan tersebut sendiri pada LKPD. Setelah dikurangkan, apa yang terjadi pada suku yang memuat x ?
- Student:** (Handwritten notes showing calculations for x)
- Chatbot:** yang saya ketahui hanya sampai disini, apa yang harus saya lakukan berikutnya?
- Student:** (Handwritten notes showing calculations for x)
- Chatbot:** Bagus, Kelompok 1. Kalian sudah sampai pada nilai y melalui eliminasi. Sekarang, satu langkah lagi berikutnya: gunakan nilai y yang sudah kalian peroleh ke salah satu persamaan awal untuk mencari nilai x . Menurut kalian, lebih mudah menggunakan persamaan Pak Sabang atau persamaan Pak Rahman? Mengapa?
- Student:** (Handwritten notes showing calculations for x)
- Chatbot:** Bagus, Kelompok 1. Sekarang koefisien y pada persamaan kedua sudah kalian gunakan dengan benar, yaitu 4. Namun, coba perhatikan baris hasil perkalian terakhir: jika 4 dikalikan dengan nilai y yang kalian peroleh, apakah hasilnya benar-benar Rp100.000,00? Coba hitung kembali bagian 4 x Rp50.000,00, lalu tuliskan hasilnya sendiri.
- Student:** (Handwritten notes showing calculations for x)
- Chatbot:** Bagus, Kelompok 1. Sekarang pemeriksaan persamaan kedua sudah konsisten: hasil substitusi kalian menghasilkan total pembayaran yang sesuai dengan soal. Berarti nilai x dan y kalian lulus pemeriksaan pada kedua persamaan.

Figure 1. Examples of Student–Chatbot Interactions Across Problem-Solving Stages

Note. The excerpts are authentic student–chatbot interaction records collected during the learning activities. Student identifiers have been removed to maintain anonymity.

Descriptive analysis showed an improvement in students' mathematical problem-solving ability after participating in the learning activities. The mean score increased by 47.67 points, from 33.13 on the pretest to 80.80 on the posttest, while score variability decreased from an SD of 15.42 to 8.21. The minimum and maximum scores also increased from 9–69 to 63–97, respectively. A summary of the descriptive statistics is presented in Table 1.

Table 1. Descriptive Statistics of Students' Mathematical Problem-Solving Ability

Measurement	N	Mean	SD	Minimum	Maximum
Pretest	30	33.13	15.42	9	69
Posttest	30	80.80	8.21	63	97
Difference	30	47.67	14.99	22	81

To clarify the changes in the mean scores and variability in students' mathematical problem-solving ability before and after the learning intervention, the pretest and posttest scores are compared in Figure 2.

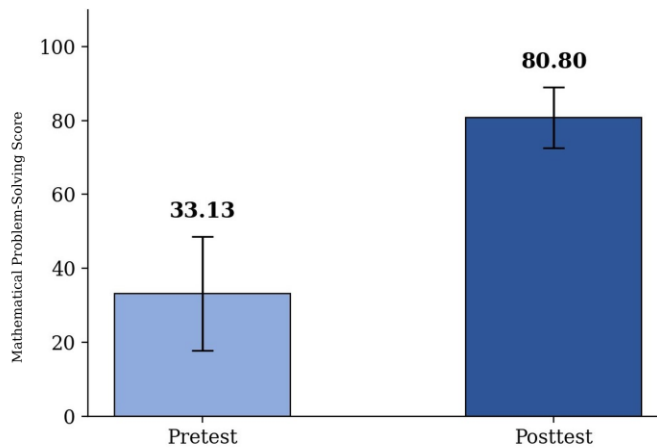


Figure 2. Mean Pretest and Posttest Scores of Students' Mathematical Problem-Solving Ability

Normality was then assessed using the difference scores between the *posttest* and *pretest*. The Shapiro–Wilk test yielded $W = 0.979$ with $p = 0.802$. A p -value greater than 0.05 indicates that the difference scores were normally distributed. Therefore, the analysis of score differences was conducted using a *paired-samples t-test*. The results of the normality test are presented in Table 2.

Table 2. Normality Test Results for Difference Scores

Data	W	N	P	Decision
Posttest – Pretest	0.979	30	0.802	Normal

Based on the normality test results presented in Table 2, the difference scores were normally distributed; therefore, a paired-samples *t-test* was conducted to examine the difference in students' mathematical problem-solving ability scores before and after the learning intervention. The test showed a statistically significant difference, $t(29) = 17.419$, $p < .001$. The mean difference in scores was 47.67 points, with a 95% confidence interval of [42.07, 53.26]. The effect size for paired data was Cohen's $d_z = 3.18$. The results of the difference test and effect size are presented in Table 3.

Table 3. Results of the Difference Test and Effect Size

Analysis	Value
Mean difference	47.67
95% CI	[42.07, 53.26]
$t(29)$	17.419
p	< 0.001
Cohen's d_z	3.18

Further analysis was conducted to determine the level of improvement in students' mathematical problem-solving ability based on N-Gain. The mean individual N-Gain was 0.707. A total of 15 students (50%) were in the moderate category and 15 students (50%) were in the high category, with no students falling into the low category. The distribution of N-Gain categories for students' mathematical problem-solving ability is presented in Table 4.

Table 4. Distribution of N-Gain Categories for Mathematical Problem-Solving Ability

Category	N-Gain Range	n	Percentage
Low	< 0.30	0	0%
Moderate	$0.30 - < 0.70$	15	50%
High	≥ 0.70	15	50%
Total		30	100%

The distribution of students across the three N-Gain categories is further visualized in Figure 3 to illustrate the composition of improvement at the individual level.

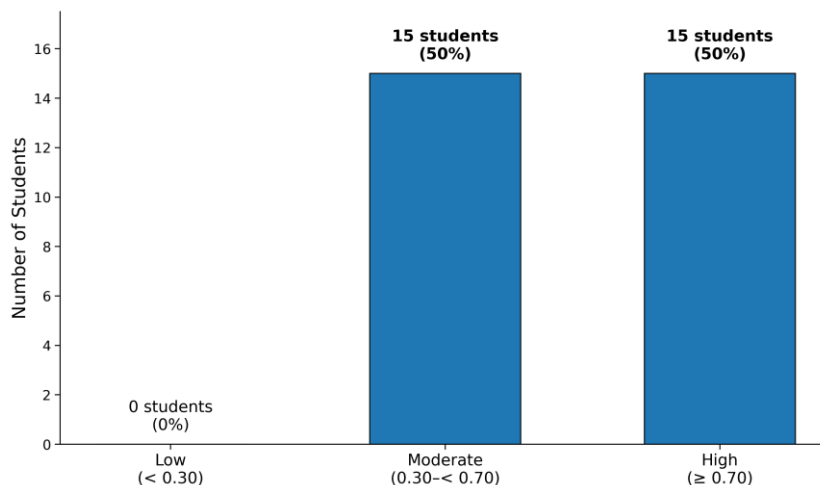


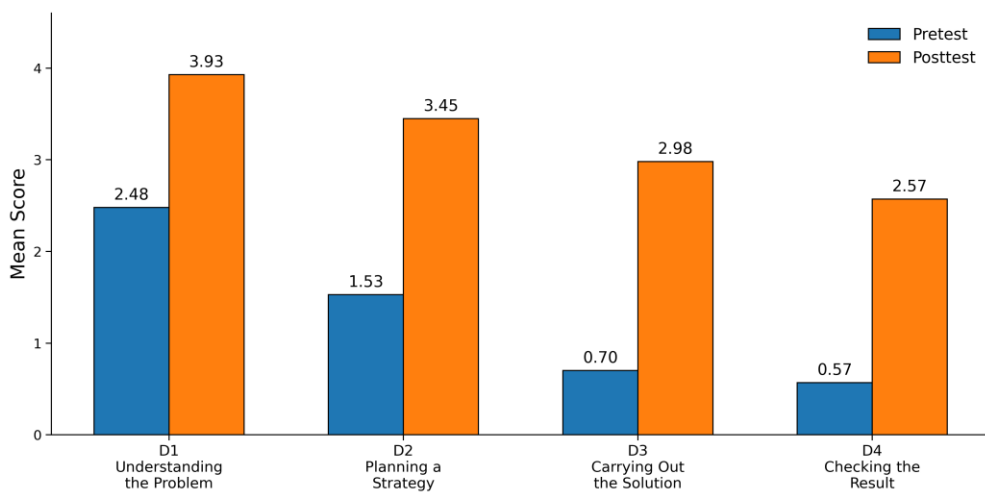
Figure 3. Distribution of N-Gain Categories for Mathematical Problem-Solving Ability

Further analysis was conducted across the four problem-solving stages to distinguish changes in raw scores from normalized improvement. The mean D1 score increased from 2.48 to 3.93, an absolute increase of 1.45 points, with an N-Gain of 0.956 (high). D2 increased from 1.53 to 3.45, an absolute increase of 1.92 points, with an N-Gain of 0.777 (high). D3 increased from 0.70 to 2.98, representing the largest absolute increase of 2.28 points, while its N-Gain was 0.692 (moderate). D4 increased from 0.57 to 2.57, an absolute increase of 2.00 points, with an N-Gain of 0.583 (moderate). Thus, the largest absolute gain occurred at D3, whereas the highest normalized gain occurred at D1. Changes across the four problem-solving stages are presented in Table 5.

Table 5. Changes in Scores at Each Problem-Solving Stage

Stage	<i>Pretest</i>	<i>Posttest</i>	Increase	N-Gain	Category
Understanding the problem (D1)	2.48	3.93	1.45	0.956	High
Planning a strategy (D2)	1.53	3.45	1.92	0.777	High
Carrying out the solution (D3)	0.70	2.98	2.28	0.692	Moderate
Checking the result (D4)	0.57	2.57	2.00	0.583	Moderate

To clarify the pattern of changes at each stage, a visual comparison of the mean pretest and posttest scores for D1–D4 is presented in Figure 4.

**Figure 4. Mean Pretest and Posttest Scores at Each Problem-Solving Stage**

DISCUSSION

The results of this study showed changes in students' mathematical problem-solving ability after they participated in learning that integrated the ethnomathematical context of Jenepono salt production, Problem-Based Learning (PBL), and an AI chatbot. The classroom observations also indicated that the learning activities were implemented at a very good level, with an average implementation score of 85.80% across the eight learning meetings. The progressive increase in implementation scores from 84.09% in the early meetings to 87.50% in the final meetings suggests that the learning sequence became increasingly established during implementation. These changes were observed not only in the overall scores but also across the four problem-solving stages: understanding the problem (D1), planning a strategy (D2), carrying out the solution (D3), and checking the result (D4). This pattern indicates that the learning experience provided opportunities for students to move from

understanding the situation to planning, carrying out, and evaluating solutions. Conceptually, this pattern is consistent with Pólya's framework, which views problem solving as a sequence of understanding the problem, devising a plan, carrying out the plan, and looking back at the result³³.

The overall change in scores can be understood from the characteristics of the problems used. Salt production was not merely presented as an illustration of the problems but served as the source of situations that students had to translate into mathematical representations. Students first encountered quantitative information that had meaning in community life, then determined the relevant quantities, established relationships among those quantities, and used them to formulate SPLDV models. This mechanism is relevant to findings that difficulties in mathematical problem solving often arise during the mathematizing process, particularly when students have to transform information from real-life situations into mathematical representations and models³⁴. Thus, the cultural context in this study functioned as a starting point for constructing mathematical meaning rather than merely as a narrative setting.

These findings strengthen empirical evidence regarding the potential of ethnomathematics-based learning. The meta-analysis by Apriatni et al. reported a strong combined effect on mathematical problem-solving ability, while Pratama and Yanpar Yelken found a positive effect of ethnomathematics-based learning on mathematical literacy, with variations influenced by regional characteristics, learning variables, and instructional materials^{35,36}. The present findings extend these results to the context of local salt production in Jeneponto and to an analysis based on problem-solving stages. This means that the contribution of cultural contexts can be considered not only in terms of improvements in final scores but also in their potential use for organizing mathematical thinking processes from understanding a situation to evaluating a solution.

The D1–D4 analysis provides a more detailed picture of changes in students' ability and distinguishes absolute score gains from normalized improvement. The mean D1 score increased from 2.48 to 3.93, an absolute increase of 1.45 points, with an N-Gain of 0.956 (high). D2 increased from 1.53 to 3.45, an absolute increase of 1.92 points, with an N-Gain of 0.777 (high). D3 increased from 0.70 to 2.98, representing the largest absolute increase of 2.28 points, while its N-Gain was 0.692 (moderate). D4 increased from 0.57 to 2.57, an absolute increase of 2.00 points, with an N-Gain of 0.583 (moderate). Thus, the largest absolute change occurred in the execution stage (D3), whereas the highest normalized improvement occurred in understanding the problem (D1).

³³ Pólya, *How to Solve It: A New Aspect of Mathematical Method*.

³⁴ Jupri and Drijvers, "Student Difficulties in Mathematizing Word Problems in Algebra."

³⁵ Apriatni et al., "The Influence of Ethnomathematics Based Learning on Mathematics Problem-Solving Ability: A Meta-Analysis."

³⁶ Pratama and Yanpar Yelken, "Effectiveness of Ethnomathematics-Based Learning on Students' Mathematical Literacy: A Meta-Analysis Study."

This distinction suggests that the magnitude of raw-score change and the proportion of possible improvement were not identical across the four stages. The pattern also indicates that the changes extended beyond students' initial ability to identify information to their ability to implement strategies and review their results.

The D4 pattern needs to be considered alongside previous research conducted in Jeneponto. That study used PBL with a salt-production context and STEAM integration and reported improvements in problem-solving ability, but the checking stage remained relatively the weakest stage under the final learning condition. In the present study, D4 also did not reach the final level achieved by D1 and D2, but its absolute increase of 2.00 points was greater than the increases in D1 and D2. The observed improvement in D4 may be related to the opportunities provided during the learning process for students to revisit their procedures and verify their results. The classroom observations and interaction records are consistent with this interpretation because they documented repeated prompting to check procedures and results. However, this should be understood as a possible learning mechanism rather than evidence that the chatbot independently caused the improvement in D4, because the one-group pretest–posttest design did not separate the effects of the chatbot from PBL, the ethnomathematical context, or other learning factors³⁷³⁸.

The position of the AI chatbot is important in explaining this learning design. The chatbot was intended to function as an interactive companion that provided staged assistance through open-ended questions, directed questions, and concrete hints. The interaction records provide concrete evidence of this function: students were prompted to identify information at D1, consider mathematical strategies at D2, continue procedural steps at D3, and verify their solutions at D4. These records show that the chatbot was actually incorporated into the problem-solving process rather than being used only as a source of final answers. At the same time, the records revealed occasional responses in which the chatbot provided explicit mathematical representations or intermediate calculations. Therefore, the chatbot functioned predominantly as structured process support, but the intended level of scaffolding was not maintained uniformly across all interactions. This finding reinforces the importance of monitoring AI-generated responses rather than assuming that a predefined prompt will always produce the intended pedagogical behavior³⁹⁴⁰.

The present findings are broadly consistent with studies that have examined AI chatbots as part of problem-based or mathematics learning.

³⁷ Syaharuddin et al., "Students' Mathematical Problem-Solving Ability Based on Polya's Stages: A Study of Local Salt-Production Ethnomathematics-Based Learning in Jeneponto."

³⁸ Shadish et al., *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*.

³⁹ Jančařík et al., "Using AI Chatbot for Math Tutoring."

⁴⁰ Laja et al., "Chatbot Artificial Intelligence (AI) to Supporting Differentiated Mathematics Learning."

Wibowo and Faizah reported improvements in higher-order thinking through problem-based learning assisted by an AI chatbot⁴¹, while Kasneci et al. highlighted the potential of large language models for personalized and interactive learning alongside concerns about accuracy, dependence, and the need for supervision⁴². Husain et al. similarly reported differences between PBL supported by an AI chatbot and PBL without chatbot support in junior high school mathematics. The present study differs in the way these components are functionally connected: salt production provides the mathematical context, PBL organizes the learning activity, Pólya's stages structure the problem-solving process, and the chatbot provides staged interaction during students' work⁴³.

The intended restriction of the chatbot is also important from a pedagogical perspective. The chatbot was configured to guide rather than replace students' mathematical reasoning, with students expected to determine representations, select strategies, perform calculations, and verify their results. However, the interaction records show that this restriction was not perfectly maintained in every response⁴⁴. This observation is consistent with concerns in the literature regarding the accuracy and pedagogical reliability of AI-generated responses. Consequently, the educational value of the chatbot depends not only on the prompt or protocol but also on teacher supervision, students' critical evaluation, and opportunities to verify AI-generated information⁴⁵.

The need for human supervision is particularly important because AI-generated outputs cannot always be accepted as accurate information. Previous research has emphasized the need for critical evaluation, while studies of ChatGPT in education have highlighted both its opportunities and its limitations. In the present study, teacher supervision remained important when students encountered difficulties or when chatbot responses required clarification^{46,47}. Thus, the integration of AI did not replace the pedagogical function of the teacher but positioned the chatbot as one source of assistance within a learning structure designed and monitored by the teacher.

The main contribution of this study lies in the functional relationship among the learning components. The ethnomathematical context of salt

⁴¹ Wibowo and Faizah, "Pengaruh Model Pembelajaran Berbasis Masalah Berbantuan Chatbot AI Terhadap Higher Order Thinking Skill."

⁴² Kasneci et al., "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education."

⁴³ Husain et al., "Efektivitas Chatbot-AI Dalam Pembelajaran Matematika Melalui Model NHT Berbasis PBL Di SMPN 3 Binamu."

⁴⁴ Turmuzi et al., "ChatGPT in School Mathematics Education: A Systematic Review of Opportunities, Challenges, and Pedagogical Implications."

⁴⁵ Lo et al., "The Influence of ChatGPT on Student Engagement: A Systematic Review and Future Research Agenda."

⁴⁶ Purnomo et al., "Islamic Education: Examining AI Use, Reliance, and Perceived Validity among Tarbiyah Students."

⁴⁷ Tlili et al., "What If the Devil Is My Guardian Angel: ChatGPT as a Case Study of Using Chatbots in Education."

production provides meaningful mathematical problems; PBL organizes students' activities for investigating and solving those problems; Pólya's stages provide a structure for observing the problem-solving process; and the AI chatbot provides interactive support when students encounter difficulties. With this architecture, AI does not function independently as an additional technology but is positioned within a learning sequence that has clear cognitive objectives, a defined cultural context, and teacher supervision. This functional integration extends previous research that has often examined ethnomathematics and AI-supported learning as separate approaches.

The N-Gain distribution, consisting of 15 students in the moderate category and 15 students in the high category, indicates that improvement was observed across all students analyzed, although the magnitude of improvement was not identical. This pattern provides descriptive evidence that changes occurred consistently within the observed research group. However, these results cannot be generalized to a broader student population. The study involved only one intact class at one school and did not use a comparison group. Changes in scores may therefore be associated with the learning experience but may also have been influenced by students' development, prior learning, or other factors during the study period. The findings are consequently more appropriately interpreted as initial empirical evidence of changes following implementation of the learning design rather than as an estimate of the independent causal effect of the AI chatbot.

From an educational perspective, the findings indicate that an AI chatbot can be positioned to expand students' opportunities to receive assistance during problem solving while the teacher remains the primary guide of the pedagogical process. The teacher designs contextual problems, establishes rules for chatbot interaction, monitors the learning process, intervenes when necessary, and ensures that students verify their solution steps and results. Theoretically, the study supports the use of cultural contexts as sources of mathematical problems that can be processed through mathematical modeling and problem solving. Practically, the staged-assistance design provides one possible way for teachers to use chatbots without transferring responsibility for mathematical reasoning to AI. These implications should nevertheless be interpreted proportionally to the one-group research design. Future studies should use comparison groups, involve multiple schools, and analyze students' interaction records in greater depth to examine which forms of chatbot assistance are associated with changes at different problem-solving stages.

CONCLUSION

This study showed changes in students' mathematical problem-solving ability after participating in learning that integrated Jeneponto salt-production ethnomathematics, Problem-Based Learning (PBL), Pólya's problem-solving stages, and AI chatbot support. The mean score increased from 33.13 on the

pretest to 80.80 on the posttest, with a mean difference of 47.67 points, $t(29) = 17.419$, $p < .001$, and Cohen's $d_z = 3.18$. The mean N-Gain of 0.707 was in the high category, with 15 students in the moderate category and 15 in the high category. At the problem-solving stage level, all four stages increased. The largest absolute increase occurred in D3, at 2.28 points, followed by D4 at 2.00 points, D2 at 1.92 points, and D1 at 1.45 points, whereas the highest normalized improvement was observed in D1 (N-Gain = 0.956).

The contribution of this study lies in a learning design that connects salt production as a source of meaningful mathematical problems, PBL as the organizer of learning activities, Pólya's stages as a structure for the problem-solving process, and the AI chatbot as staged interactive support under teacher supervision. The interaction records provide evidence that the chatbot was incorporated into students' problem-solving activities, although the intended level of scaffolding was not maintained uniformly across all interactions. These findings therefore provide initial empirical evidence of changes in students' mathematical problem-solving ability following implementation of the learning design, rather than evidence of an independent causal effect of the AI chatbot.

The findings suggest that AI-supported mathematics learning can be designed to connect local cultural contexts with structured problem solving while maintaining the teacher's role in monitoring and evaluating the learning process. However, the one-group pretest–posttest design, single-class setting, and absence of a comparison group limit causal interpretation and generalizability. Future research should employ comparison groups, involve multiple schools, and analyze chatbot interaction records more systematically to examine how different forms of AI assistance relate to changes at each problem-solving stage.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Ministry of Higher Education, Science, and Technology (Kemdiktisaintek) for providing research funding support. The authors also thank the Education and Culture Office of Jenepono Regency for institutional support and facilitation of the research implementation, the UPT junior high school where the study was conducted for granting permission and providing access to conduct the research, and the mathematics teacher who assisted with and implemented the learning activities throughout the study.

AUTHOR CONTRIBUTIONS

Syharuddin: Conceptualization, methodology, development of the ethnomathematics- and AI chatbot-based learning design, data analysis, and manuscript writing and finalization.

Sumiati S.: Literature review, learning model design, development of learning materials, and manuscript review.

Nurqiyamah Hamid: Learning implementation, data collection, analysis of students' mathematical problem-solving ability, and manuscript review.

Hasmawati: Validation of the learning model and learning materials, interdisciplinary contribution, documentation, and manuscript writing and review.

Muh. Usman Taufik: Instrument preparation, learning implementation, data collection and processing, documentation, and manuscript review under the guidance of the research team.

Adryan Putra Pratama: Technical support for the implementation of learning technology, preparation and management of learning devices and AI chatbot access, technical documentation of technology use, and review of the technological aspects of the manuscript.

FUNDING

This research was funded by the Ministry of Higher Education, Science, and Technology through the 2026 Research Grant, Beginner Lecturer Research (PDP) Scheme, based on Decree Number 196/KPA/C3/KPT/2026 and Agreements/Contracts Number 288/C3/DT.05.00/PL-BARU/2026, 306/LL9/PPIP.PL-MONO/2026, and 15/K-LPPM-INTI/JP/IV/2026.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Aiken, Lewis R. "Three Coefficients for Analyzing the Reliability and Validity of Ratings." *Educational and Psychological Measurement* 45, no. 1 (1985): 131–42. <https://doi.org/10.1177/0013164485451012>.
- Aposika, Francis. "Investigating the Relationship between Problem-Solving Skills and Academic Performance in STEM Subjects among Secondary School Students." *Discover Education* 5 (2026): 425. <https://doi.org/10.1007/s44217-026-01487-w>.
- Apriatni, Sri, Syamsuri, Hepsi Nindiasari, and Sukirwan. "The Influence of Ethnomathematics Based Learning on Mathematics Problem-Solving Ability: A Meta-Analysis." *Jurnal Pendidikan Matematika (JUPITEK)* 5, no. 1 (2022): 23–33. <https://doi.org/10.30598/jupitekvol5iss1pp23-33>.
- Cronbach, Lee J. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16, no. 3 (1951): 297–334. <https://doi.org/10.1007/BF02310555>.
- Doc, Nguyen Van, Nguyen Thi Hoai Nam, Ngo Tu Thanh, and Nguyen Minh Giam. "Teaching Mathematics with the Assistance of an AI Chatbot to Enhance Mathematical Thinking Skills for High School Students."

- International Journal of Current Science Research and Review* 6, no. 12 (2023): 8574–80. <https://doi.org/10.47191/ijcsrr/V6-i12-102>.
- Fadila, Meyla Tri, Reni Untarti, Akhmad Jazuli, and Anton Jaelani. “Mathematics in the AI Era: The Effectiveness of ChatGPT in Helping Students Solve HOTS Problems.” *AlphaMath: Journal of Mathematics Education* 11, no. 2 (2025): 191–204. <https://doi.org/10.30595/alphamath.v11i2.28261>.
- Hake, Richard R. “Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses.” *American Journal of Physics* 66, no. 1 (1998): 64–74. <https://doi.org/10.1119/1.18809>.
- Husain, Husriani, Syaharuddin, Yasriuddin, and Sri Wahyuni Agustina. “Efektivitas Chatbot-AI Dalam Pembelajaran Matematika Melalui Model NHT Berbasis PBL Di SMPN 3 Binamu.” *Mandalika Mathematics and Education Journal* 8, no. 1 (2026): 193–208. <https://doi.org/10.29303/jm.v8i1.11074>.
- Jančařík, Antonín, Jakub Michal, and Jarmila Novotná. “Using AI Chatbot for Math Tutoring.” *Journal of Education Culture and Society* 14, no. 2 (2023): 285–96. <https://doi.org/10.15503/jecs2023.2.285.296>.
- Jupri, Al, and Paul Drijvers. “Student Difficulties in Mathematizing Word Problems in Algebra.” *Eurasia Journal of Mathematics, Science and Technology Education* 12, no. 9 (2016): 2481–502. <https://doi.org/10.12973/eurasia.2016.1299a>.
- Kasneci, Enkelejda, Kathrin Seßler, Stefan Küchemann, et al. “ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education.” *Learning and Individual Differences* 103 (2023): 102274. <https://doi.org/10.1016/j.lindif.2023.102274>.
- Laja, Y. P. W., E. B. S. Delvion, and M. M. L. Siahaan. “Chatbot Artificial Intelligence (AI) to Supporting Differentiated Mathematics Learning.” *Jurnal Pendidikan MIPA* 26, no. 1 (2025): 118–30. <https://doi.org/10.23960/jpmipa.v26i1.pp118-130>.
- Lakens, Daniël. “Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-Tests and ANOVAs.” *Frontiers in Psychology* 4 (2013): 863. <https://doi.org/10.3389/fpsyg.2013.00863>.
- Lo, Chung Kwan, Khe Foon Hew, and Morris Siu-yung Jong. “The Influence of ChatGPT on Student Engagement: A Systematic Review and Future Research Agenda.” *Computers & Education* 219 (2024): 105100. <https://doi.org/10.1016/j.compedu.2024.105100>.
- Maulana, Septian Dika, and Cecep Anwar Hadi Firdos Santosa. “Analisis Kemampuan Pemecahan Masalah Matematika Siswa SMA Pada

- Kurikulum Merdeka Dan Kurikulum 2013 Berdasarkan Teori Polya.” *Jurnal Pendidikan Matematika: Judika Education* 7, no. 1 (2024): 93–105. <https://doi.org/10.31539/judika.v7i1.8341>.
- Maulida, Saprina, and Lisa Dwi Afri. “Implementasi Model Pembelajaran Contextual Teaching and Learning (CTL) Berbasis Etnomatematika Terhadap Kemampuan Literasi Numerasi Siswa.” *Jurnal Pendidikan Matematika: Judika Education* 8, no. 4 (2025): 626–36. <https://doi.org/10.31539/apdh6c08>.
- OECD. *PISA 2025 Results (Volume I): Future-Ready Students*. OECD Publishing, 2026. <https://doi.org/10.1787/73451bc5-en>.
- Okonkwo, Chukwunonso W., and Abejide Ade-Ibijola. “Chatbots Applications in Education: A Systematic Review.” *Computers and Education: Artificial Intelligence* 2 (2021): 100033. <https://doi.org/10.1016/j.caeai.2021.100033>.
- Pólya, George. *How to Solve It: A New Aspect of Mathematical Method*. 2nd ed. Princeton University Press, 1973.
- Pratama, Ryan Angga, and Tuğba Yanpar Yelken. “Effectiveness of Ethnomathematics-Based Learning on Students’ Mathematical Literacy: A Meta-Analysis Study.” *Discover Education* 3 (2024): 202. <https://doi.org/10.1007/s44217-024-00309-1>.
- Purnomo, Daniel Desta Avia Difaputra, Sri Fariyati, Abdullah Arif, and Baidar MM Hasan. “Islamic Education: Examining AI Use, Reliance, and Perceived Validity among Tarbiyah Students.” *Paedagogia: Jurnal Pendidikan* 15, no. 1 (2026): 173–90. <https://doi.org/10.24239/pdg.Vol15.Iss1.1436>.
- Santos-Trigo, Manuel. “Problem Solving in Mathematics Education: Tracing Its Foundations and Current Research-Practice Trends.” *ZDM—Mathematics Education* 56, no. 2 (2024): 211–22. <https://doi.org/10.1007/s11858-024-01578-8>.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin, 2002.
- Shapiro, S. S., and M. B. Wilk. “An Analysis of Variance Test for Normality (Complete Samples).” *Biometrika* 52, nos. 3–4 (1965): 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>.
- Syahrudin, Siti Rahaimah Binti Ali, Agustan Syamsuddin, Nursakiah, and Randy Saputra Mahmud. “Students’ Mathematical Problem-Solving Ability Based on Polya’s Stages: A Study of Local Salt-Production Ethnomathematics-Based Learning in Jenepono.” *Journal of Local Wisdom in Education* 2, no. 1 (2026): 38–56. <https://doi.org/10.65508/01s42c69>.

- Tlili, Amirhossein, B. Shehata, M. A. Adarkwah, et al. "What If the Devil Is My Guardian Angel: ChatGPT as a Case Study of Using Chatbots in Education." *Smart Learning Environments* 10, no. 1 (2023): 15. <https://doi.org/10.1186/s40561-023-00237-x>.
- Turmuzy, Muhammad, Syahrul Azmi, and Ni Made Intan Kertiyani. "ChatGPT in School Mathematics Education: A Systematic Review of Opportunities, Challenges, and Pedagogical Implications." *Teaching and Teacher Education* 170 (2026): 105286. <https://doi.org/10.1016/j.tate.2025.105286>.
- Wibowo, Setyo Ajie, and Intan Faizah. "Pengaruh Model Pembelajaran Berbasis Masalah Berbantuan Chatbot AI Terhadap Higher Order Thinking Skill." *Jurnal Nyanadassana: Jurnal Penelitian Pendidikan, Sosial Dan Keagamaan* 2, no. 2 (2023): 73–85. <https://doi.org/10.59291/jnd.v2i2.36>.