

Students' Mathematical Problem-Solving Ability Based on Polya's Stages: A Study of Local Salt-Production Ethnomathematics-Based Learning in Jeneponto

Syahrudin^{1,2*}, Siti Rahaimah Binti Ali³, Agustan Syamsuddin⁴, Nursakiah⁴, Randy Saputra Mahmud⁵

¹Doctoral Program in Education, Universitas Muhammadiyah Makassar, Makassar, Indonesia

²Department of Mathematics Education, Institut Turatea Indonesia, Jeneponto, Indonesia

³Faculty of Human Development, Universiti Pendidikan Sultan Idris, Tanjung Malim, Malaysia

⁴Department of Mathematics Education, Universitas Muhammadiyah Makassar, Makassar, Indonesia

⁵Doctoral School of Education, University of Szeged, Szeged, Hungary

Article Info

Article history:

Received Mar 18, 2026

Revised Jun 18, 2026

Accepted Jun 26, 2026

Published Online Jun 30, 2026

Keywords:

Ethnomathematics

Local wisdom

Mathematical modeling

Problem-solving

Salt-production context

ABSTRACT

Contextual mathematics instruction is important when pupils struggle to translate everyday situations into algebraic models. However, the role of local knowledge in supporting stages of mathematical problem solving remains underexplored. This study examined changes in year 9 pupils' mathematical problem-solving performance across Polya's four stages after a local salt-production ethnomathematics learning sequence in Jeneponto, Indonesia. A pre-experimental one-group pretest-posttest design involved 32 pupils in three problem-based learning sessions integrating STEAM perspectives. Two constructed-response tasks on systems of linear equations in two variables used authentic salt-production scenarios to assess problem understanding, equation formulation, solution execution, and verification. Responses were scored with an expert-reviewed stage-specific rubric and analyzed using descriptive statistics, paired-samples tests, standardized effect sizes, and normalized gain. Mean scores increased significantly at every stage and overall, with a medium overall normalized gain of 0.51. Gains were highest for devising a plan (0.66) and carrying out the plan (0.53), while looking back remained the weakest posttest stage (45.31%; gain = 0.45). Posttest performance also differed significantly between the two task items, suggesting unequal task demands. By combining an authentic coastal livelihood context with stage-level measurement, the study shows how local wisdom can function as an epistemic bridge to mathematical modeling while highlighting verification as a distinct instructional need. The study contributes a culturally grounded approach for diagnosing problem-solving processes and planning mathematics support.

© 2026 The Author(s). This is an open access article under the [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding author:

Syahrudin

Email: syahrud2@gmail.com

I. INTRODUCTION

Mathematical problem-solving is a central goal of mathematics education because learners must interpret a situation, represent its relationships, choose and carry out a strategy, and judge whether the resulting answer is reasonable. These demands are particularly evident in lower-secondary algebra, where pupils must translate narrative information into symbolic form. Indonesian studies continue to report difficulty with this transition: many pupils can perform procedures once an equation is given but struggle to mathematize a word problem independently (Jupri & Drijvers, 2016). The same concern appears in large-scale assessment. The Programme for International Student Assessment (PISA) shows that Indonesian learners continue to find tasks involving mathematical modeling and the interpretation of real-world situations challenging (OECD, 2023), while district-level education-report data show a decline in the proportion of junior-secondary pupils in Jeneponto Regency who meet the minimum numeracy standard (Kemendikdasmen, 2026).

More practice alone is unlikely to resolve this difficulty. Algebraic modeling involves making sense of quantities and relationships before they are expressed formally. In Jeneponto, salt production is a familiar coastal livelihood shaped by land area, salt quantities and grades, prices, labor, and income. These practices provide a concrete basis for discussing linear relationships without using culture merely as decoration. Ethnomathematics views such practices as situated mathematical knowledge and encourages careful connections between community practices and formal school mathematics (D'Ambrosio, 2018). Recent work likewise treats ethnomathematics as a resource for relevance, participation, and conceptual discussion, while emphasizing that its educational value depends on deliberate instructional design rather than on cultural reference by itself (Batiibwe, 2025).

Polya's four-stage framework offers a practical way to examine where that support occurs: understanding the problem, devising a plan, carrying out the plan, and looking back (Polya, 1981). Unlike a single total score, the framework can reveal where a pupil's reasoning breaks down. A pupil may understand a familiar situation yet fail to construct equations, or may complete an algorithm without checking whether the answer is sensible in context. Earlier research in Jeneponto found that successful pupils moved through these broad stages when solving systems of linear equations, whereas lower performance was associated with limited conceptual understanding (Syaharuddin, 2019). A systematic review of algebraic thinking similarly identifies symbolization, generalization, and representation as persistent difficulties across educational levels (Sibgatullin et al., 2022).

Instruction, then, needs to combine cultural relevance with explicit support for modeling and reflection. Problem-Based Learning is well suited to this purpose because it centers learning on a meaningful problem, collaborative inquiry, and justification of proposed solutions. A systematic review of technology-supported Problem-Based Learning in mathematics education reports potential benefits for problem solving and critical thinking, while also stressing the importance of carefully designed tasks and teacher preparation (Kirabo et al., 2024; Noralam et al., 2025). The learning sequence in this study was also organized from a science, technology, engineering, arts, and mathematics perspective. Salt production served as the common context linking measurement, proportional and algebraic reasoning, representation, practical design, and communication rather than as a set of separate disciplinary activities. Recent reviews note that interdisciplinary mathematics approaches require a coherent conceptual rationale and assessments aligned with the intended thinking processes (Deák & Kumar, 2024; Fiskawarni et al., 2025; Martín-Cudero et al., 2024).

Indonesian research provides several examples of culturally grounded mathematics learning. Buginese food practices have been used to explore geometry and values (Pathuddin et al., 2021), while batik patterns have supported culturally situated geometry learning (Prahmana & D'Ambrosio, 2020). Local-wisdom modules have been developed to support character formation and learning outcomes (Salmia et al., 2024a, 2024b), and familiar games and joyful mathematics activities have been associated with engagement and early numeracy (Kurniati et al., 2025; Salmia & Malik, 2025). *Sipakatau* and *Sipakalebbi* values have also been incorporated into elementary character-education approaches (Salmia et al., 2024a). In addition, a classroom study of integrated Problem-Based Learning reported gains in critical-thinking performance, although its one-group design calls for cautious interpretation (Nurfadila et al., 2023). Together, these studies show that local knowledge can be used productively in instruction. They do not, however, identify which stage of formal problem solving changes when pupils work with a local livelihood context, nor do they separate progress in mathematical modeling from progress in verification.

This gap has both theoretical and practical consequences. Theoretically, research on local wisdom in education needs to explain how cultural context mediates specific cognitive actions, rather than simply whether learners enjoy familiar materials. Practically, teachers need evidence that points to where follow-up support is needed: understanding the context, translating it into equations, carrying out a plan, or evaluating the solution. Meta-analytic evidence suggests that ethnomathematics-based learning can support mathematical outcomes, but it also shows

substantial variation in tasks, measures, and implementation conditions (Pratama & Yanpar Yelken, 2024). Research on mathematical word problems likewise shows that wording, contextual cues, and task structure influence performance, so two contextual items should not be assumed to place identical demands on pupils (Vessonen et al., 2024). Studies of cultural pedagogy make a related point: local values become educationally meaningful when they are tied to an explicit learning process rather than simply named as background context (Kadir & Rosmini, 2025). What remains unclear is the pathway from context to formal reasoning, whether a familiar livelihood supports only initial comprehension, extends to equation formation and execution, or also helps pupils evaluate a solution. Answering this question requires stage-level evidence rather than an aggregate achievement score alone.

This study examined a three-session local salt-production ethnomathematics learning sequence with Year 9 pupils in Jeneponto. Three questions guided the analysis: (1) How did pupils' overall mathematical problem-solving performance, and their performance at each Polya stage, change from pretest to posttest? (2) How did the distribution of pupils across low, medium, and high problem-solving categories change before and after the sequence? and (3) Did the two planned parallel salt-production tasks produce comparable posttest performance? By examining both stage-specific outcomes and task-level variation, the study develops an operational account of local wisdom as a bridge between familiar livelihood knowledge and formal algebraic reasoning. Given the one-group pretest–posttest design, the analysis focuses on descriptive and inferential evidence of within-class change rather than attributing the observed differences solely to the intervention.

II. METHODS

A pre-experimental one-group pretest–posttest design was used (Creswell & Creswell, 2018; Shadish et al., 2002). The same class completed the same instrument before and after the instructional sequence, with no separate comparison group. This design was selected because the study aimed to describe the magnitude and stage-specific pattern of change within one class rather than compare the intervention with an alternative approach. It allows each pupil's performance to be compared directly across the two measurement occasions and makes changes across the four problem-solving stages visible. Because there was no comparison group, however, the observed differences cannot be attributed exclusively to the learning sequence.

Participants were 32 Year 9 pupils from one mathematics class at a state junior secondary school in Jeneponto Regency, South Sulawesi, during the 2025/2026 academic year. Every

pupil completed both the pretest and posttest and was included in the analysis. The dataset therefore contained 32 matched pretest–posttest observations, and all paired analyses were based on the same pupils at both measurement points.

Before instruction, pupils completed a written pretest assessing baseline mathematical problem-solving ability using a task that required them to formulate and solve an SPLDV (*Sistem Persamaan Linear Dua Variabel*, or a system of linear equations in two variables). They then participated in three Problem-Based Learning sessions, each lasting 2×40 minutes, that incorporated a STEAM approach and used salt-production practices from the Arungkeke coastal community as the context. The sequence followed three instructional phases that corresponded to [Polya's \(1981\)](#) four problem-solving stages and also formed the assessment framework.

In Session 1, pupils examined authentic production and pricing data from two salt-farming plots and worked in groups to identify the mathematical problem. Session 2 focused on modeling and solution: pupils formulated an SPLDV from the context and solved it using substitution, elimination, or a combination of methods, with teacher support provided through guiding questions. Session 3 focused on verification, interpretation, and presentation. Pupils returned their solutions to the salt-production context, judged their reasonableness, developed a simple recommendation for the farming group, and presented their work to the class. Verification received explicit attention because pupils commonly overlook this stage. Immediately after Session 3, pupils completed a posttest that was structurally equivalent to the pretest.

Both the pretest and posttest consisted of two written items asking pupils to formulate and solve an SPLDV in a salt-production scenario. Each item was structured around [Polya's \(1981\)](#) four stages, with a separate response field for each stage: D1, recording known and asked information and defining variables x and y ; D2, formulating two equations and selecting and justifying a solution method; D3, applying the chosen method to obtain numerical values for x and y ; and D4, substituting the values into both equations, interpreting the meaning of x and y in context, judging the reasonableness of the solution, and stating a final conclusion. Each stage was scored analytically on a 0 – 4 scale using a rubric ranging from “no response” to “complete and contextually accurate.” Content validity was established through expert judgment by three validators (Aiken's $V > 0.80$), and independent scoring by two raters confirmed inter-rater reliability (Cohen's $\kappa \geq 0.70$). This validation procedure is consistent with established practice in developing local wisdom–based teaching modules in the South Sulawesi

context (Salmia et al., 2024a). Using the same stage structure at both measurement points provided a directly comparable basis for examining change.

Salt production in Arungkeke is a long-established economic practice in Jeneponto's coastal communities. It involves managing salt ponds (*tammang*), whose quality varies with sunlight exposure and seawater flow, and selling salt by the sack at prices that depend on quality. The practice naturally yields two unknown quantities – the numbers of sacks from each plot or quality grade – linked through two equations representing total production and total sales value. This made the context structurally appropriate for an SPLDV without artificially imposing a cultural setting on the mathematics. That structural fit, rather than cultural labeling alone, guided the choice of context. It also allowed numerical solutions to be evaluated against plausible production and sales conditions, an important feature of D4, in which pupils were expected to judge whether their answers were reasonable in context.

Two trained raters independently scored every pretest and posttest response using the same rubric. They compared their initial ratings only after scoring independently, discussed discrepancies, and agreed on a final consensus score for each response. This procedure was intended to reduce subjective bias, especially at D2 and D4, where scoring depends on qualitative judgments about pupils' reasoning rather than a simple right-or-wrong decision. The consensus scores were then entered into the dataset for statistical analysis.

Analyses were conducted for each Polya stage and for the overall KPM (*Kemampuan Pemecahan Masalah* or mathematical problem-solving ability) score. Pretest and posttest performance was summarized with descriptive statistics (mean and standard deviation, reported as a percentage of the maximum possible score). At each time point, pupils were also classified as high ($\geq 66\%$), medium (46–65%), or low ($\leq 45\%$) to describe changes in the score distribution. Gain scores (posttest minus pretest) met the Shapiro–Wilk normality assumption ($W = .976, p = .679$), so paired-samples t-tests were used to examine pretest–posttest differences overall and at each stage. Cohen's d_z was reported as the standardized paired effect size (Cohen, 1988). A Wilcoxon signed-rank test on overall KPM scores provided a distribution-free robustness check. The magnitude of stage-level improvement was also expressed as normalized gain ($N\text{-Gain} = (\text{posttest} - \text{pretest}) / (\text{maximum score} - \text{pretest})$), categorized as high (≥ 0.7), medium (0.3 – 0.7), or low (< 0.3) following Hake's (1998) convention. As a supplementary comparison of the two planned parallel tasks, posttest scores for items 1 and 2 were analyzed using a paired-samples t-test. All analyses were conducted in IBM SPSS Statistics version 26.

Ethical clearance covered the classroom intervention and assessment procedures,

including obtaining informed consent from pupils and their parents or guardians, as well as permission from the school principal. Two brief excerpts from posttest responses are included in the Results to illustrate the stage-level patterns identified quantitatively. They are used only as confirmatory examples and do not constitute a systematic qualitative analysis, which falls outside the scope of the study's one-group quantitative design. Their inclusion therefore does not change the quantitative design or the basis of the reported findings. The procedure of this study is summarized in Figure 1.



Figure 1. Research design and procedural stages of the study

III. RESULTS

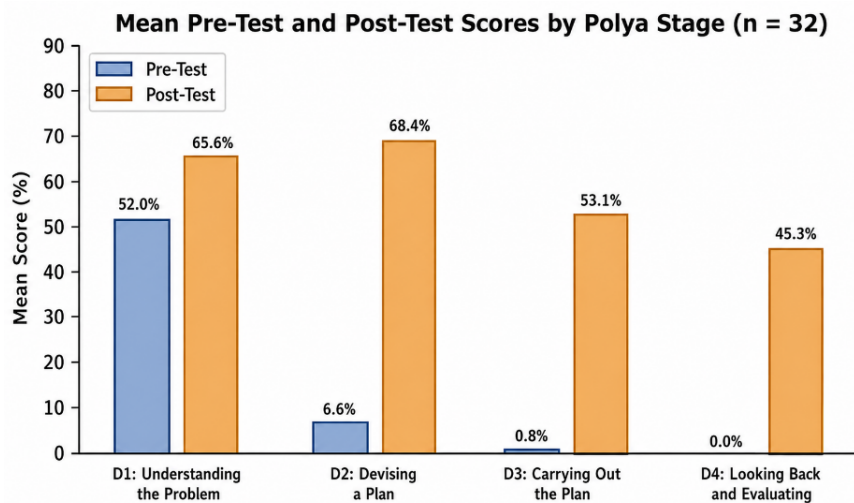
The analysis first examined changes in pupils' mathematical problem-solving performance across Polya's four stages and in overall KPM from pretest to posttest. Table 1 presents the descriptive and inferential statistics for these comparisons, including mean scores, paired-samples t -tests, effect sizes, and normalized gains.

Table 1. Pretest, posttest, and N-gain scores by Polya stage

Stage	Pretest M (%)	Posttest M (%)	t(31)	p	dz	N-gain	Category
D1: Understanding the problem	51.95	65.62	4.50	< .001	0.80	0.29	Low
D2: Devising a plan	6.64	68.36	18.09	< .001	3.20	0.66	Medium
D3: Carrying out the plan	0.78	53.12	16.63	< .001	2.94	0.53	Medium
D4: Looking back	0.00	45.31	14.33	< .001	2.53	0.45	Medium
Overall KPM	14.84	58.11	18.32	< .001	3.24	0.51	Medium

Note. dz = Cohen's d for paired samples. N-Gain categories follow [Hake \(1998\)](#): high ≥ 0.7 , medium 0.3–0.7, low < 0.3

Figure 2 visually illustrates the contrasting performance profiles across Polya's four stages before and after the learning sequence. At pretest, understanding the problem (D1) was clearly the strongest stage at 51.95%, whereas performance in D2, D3, and D4 was very low, at 6.64%, 0.78%, and 0.00%, respectively. At posttest, the profile became more balanced, with devising a plan (D2) reaching the highest mean score of 68.36%, followed by D1 at 65.62%, D3 at 53.12%, and D4 at 45.31%. The visual pattern therefore highlights substantial progress in equation formulation and solution execution, while also showing that looking back remained the least developed stage after instruction. Because the four stages began from markedly different baseline levels, the visual differences should be interpreted together with the corresponding N-Gain values rather than from posttest scores alone.

**Figure 2.** Mean pretest and posttest scores by Polya stage

D4 had the lowest posttest mean (45.31%) and a lower normalized gain (0.45) than D2 and D3. Verification therefore remained comparatively weak in this class, although the data do not explain why or show that reflection is inherently more difficult to learn. The D4 posttest mean also remained below the study's threshold for the medium category. Many pupils still needed support in substituting their answers back into the equations, interpreting the resulting solution, and judging its plausibility. Verification and reflection thus remain clear priorities for subsequent instruction.

Table 2. Distribution of pupils across KPM categories, pretest and posttest

KPM Category	Pretest n (%)	Posttest n (%)
High ($\geq 66\%$)	0 (0.0%)	9 (28.1%)
Medium (46–65%)	0 (0.0%)	16 (50.0%)
Low ($\leq 45\%$)	32 (100.0%)	7 (21.9%)

Table 2 shows a clear redistribution of scores. All 32 pupils were in the low category at pretest. By posttest, 9 pupils (28.1%) were classified as high, 16 (50.0%) as medium, and 7 (21.9%) remained low. This represents substantial movement out of the low category at the group level, but it should be read descriptively rather than as a causal estimate of intervention effectiveness. The seven pupils who remained low also show that improvement was not uniform across the class.

The distribution adds information that the mean scores alone cannot show. At posttest, 25 of 32 pupils (78.1%) reached at least the medium category, indicating that the observed gains were not confined to a small number of high scorers. At the same time, the seven pupils who remained in the low category point to meaningful variation in change and a need for differentiated follow-up support. This type of distributional evidence is useful because similar class means can mask differences in the number of pupils who cross practically meaningful performance thresholds. As with the mean results, however, the category shifts remain descriptive because no comparison group was included.

Table 3. Pretest, posttest, and N-gain scores for item 1 and item 2

Task item	Pretest M (%)	Posttest M (%)	N-Gain	Category
Item 1	16.99	68.16	0.62	Medium
Item 2	12.70	48.05	0.41	Medium

Note. Paired-samples *t*-test between item 1 and item 2 posttest scores: $t(31) = 4.37, p < .001$

Table 3 compares posttest performance on the two tasks that were intended to be parallel; it is not a test of instrument consistency. Pupils scored higher on item 1 ($M = 68.16\%$) than on item 2 ($M = 48.05\%$), $t(31) = 4.37$, $p < .001$. The significant difference suggests that the items were not empirically equivalent in difficulty or interpretive demand in this administration, a point that should be considered when the combined score is interpreted. Both items yielded medium N-Gain values, although the gain was larger for item 1. The item-level difference therefore remains important when the test is treated as a single overall measure.

Two brief posttest excerpts illustrate the contrast already visible in the quantitative results: D1 was comparatively well developed at baseline, whereas D4 remained weak after instruction. These examples are included only to make the scoring dimensions for understanding and verification more concrete. They were not selected to represent every response type, and they do not form a systematic qualitative analysis of pupils' written strategies or errors across the four Polya stages. Figures 3 and 4 were retyped and translated into English by the authors from the pupils' original handwritten Indonesian responses; the layout follows the original answer sheet.

At D1, pupil S9 defined the unknowns as “*jumlah karung garam kualitas 1*” (number of quality 1 sacks) and “*jumlah karung garam kualitas 2*” (number of quality 2 sacks), identifying the quantities to be found rather than merely repeating the unit prices given in the task. As Figure 3 shows, the response extracted the relevant quantities before equation formation and linked x and y directly to the two salt grades. This example is consistent with the relatively high D1 pretest mean of 51.95% reported in Table 1 and illustrates the scoring criterion for well-developed problem comprehension.

D1: Understanding the Problem

Given: known data | required: to be determined | variables x and y

Given: the price of salt: quality 1 = 60,000 / sack
quality 2 = 50,000 / sack
total price of salt = 22,000,000
total number of sacks = 400
y = number of quality 1 sacks, x = number of quality 2 sacks.

Figure 3. Posttest D1 response illustrating well-developed problem comprehension, consistent with the relatively high D1 pretest baseline

At D4, the pattern was different. Several posttest responses marked the verification box “✓ Benar” (correct) even when the computed values had not been substituted or the substituted values did not satisfy the equation being checked, and the accompanying interpretation fields were completed with entries that did not agree with the solution. Pupil S24’s response in Figure 4 provides one example. In both equation checks, the blanks intended for the computed values contain the variable names x and y , so the equations are restated rather than tested, yet both are marked correct. The interpretation fields describe $x = 60,000$ and $y = 30,000$ as numbers of sacks, whereas the final conclusion reports 60 and 30 sacks; the condition $x + y = \text{total}$ is affirmed, although neither pair of values sums to 900; and the stated total revenue of Rp 22,500,000 repeats the right-hand side of the second equation, although 60 and 30 sacks would yield Rp 5,100,000. Such responses show partial engagement with verification: pupils appeared to follow the checking routine and filled in every field but did not complete a valid substitution or a coherent contextual interpretation. This pattern is consistent with D4 having the lowest posttest mean and the lowest N-Gain among the three stages that began near zero. It also helps explain why D4 remained the weakest posttest stage despite improvement elsewhere. Marking an answer as correct and filling in the interpretation fields, without demonstrating that the substitution works or that the interpretation is consistent with the computed values and the context, does not meet the full looking-back criterion.

D4: Looking Back and Evaluating

Check Equation 1:	$\underline{\quad x \quad} + \underline{\quad y \quad} = \underline{\quad 900 \quad}$ ☒ Correct ☐ Incorrect
Check Equation 2:	$60,000(\underline{\quad x \quad}) + 50,000(\underline{\quad y \quad}) = \text{Rp } \underline{\quad 22,500,000 \quad}$ ☒ Correct ☐ Incorrect
Understanding the meaning of the x -variable:	$x = 60,000$, meaning <u>Number of quality 1 sacks</u>
Understanding the meaning of the y -variable:	$y = 30,000$, meaning <u>Number of quality 2 sacks</u>
Solution feasibility:	$x > 0$? <u>Yes</u> $y > 0$? <u>Yes</u> $x + y = \text{total}$? <u>Yes</u> Conclusion on feasibility: Because the values of x and y are positive, the solution is feasible.
Final conclusion:	Sell <u>60</u> sacks of quality 1 and <u>30</u> sacks of quality 2, resulting in revenue of <u>50</u> sacks. Total revenue = Rp <u>22,500,000</u>

Figure 4. Posttest D4 response illustrating superficial verification, consistent with D4’s comparatively modest N-Gain

IV. DISCUSSION

One of the clearest findings is that change differed across the four problem-solving stages. The salt-production context gave pupils access to familiar quantities, prices, and production relationships before these were represented symbolically. This familiarity may help account for the relatively high D1 baseline, but it does not show that familiarity by itself produced mathematical understanding. A more cautious interpretation is that local wisdom served as an epistemic resource: it made the quantities and relationships available for shared attention, while explicit instruction remained necessary to turn them into formal algebraic models. Seen this way, the stage-level results sharpen the general claim that contextual familiarity matters by indicating where it appears most useful and where formal support is still needed.

The largest observed gains occurred in devising and carrying out a plan, which matches the instructional emphasis on identifying unknowns, constructing equations, comparing strategies, and executing solutions. This pattern is consistent with research showing that pupils often struggle with the mathematization of algebraic word problems rather than with routine calculation alone (Jupri & Drijvers, 2016; Sibgatullin et al., 2022). It also complements studies that deliberately connect Indonesian cultural practices with disciplinary mathematics, including Buginese food practices and batik patterns (Pathuddin et al., 2021; Prahmana & D'Ambrosio, 2020). The present data cannot separate the contribution of the cultural context from Problem-Based Learning or the explicit stage prompts. What they do show is that these elements can be combined in a coherent sequence. The stage-specific gains are therefore better understood as an outcome observed under the combined instructional design rather than as evidence for any one component in isolation.

Looking back, however, the posttest stage remained the least developed. Verification asks learners to substitute a solution, determine whether it satisfies both equations, and judge whether it is plausible in the original context. These are acts of monitoring and sense-making, not simply an additional calculation. The pattern is consistent with scholarship that treats metacognitive monitoring as a distinct dimension of mathematical problem solving (Santos-Trigo, 2024; Schoenfeld, 2016). Instruction can respond by giving checking its own time and prompts: pupils can return each answer to the salt-production situation, substitute it into both equations, and explain why the resulting number of sacks, price, or income is plausible or implausible.

The category distribution also prevents the group means from being read too broadly. Although 78.1% of pupils achieved medium or high performance on the posttest, 7 remained

in the low category. Culturally familiar tasks therefore did not remove every barrier to formal mathematical reasoning. This is consistent with [Batiibwe's \(2025\)](#) argument that ethnomathematics creates pedagogical opportunities whose value depends on how they are designed and implemented. For pupils who continue to struggle, follow-up instruction may need shorter contextual tasks, explicit equation templates, peer explanations, and a gradual withdrawal of scaffolds, rather than simply more exposure to cultural content.

The item-level findings add a methodological caution. Item 1 and item 2 were designed as planned-parallel tasks, yet posttest performance differed significantly. Similar contexts therefore do not guarantee psychometric equivalence. Differences in wording, order, numerical relationships, or the salience of contextual cues may have changed the demands of the two items. This interpretation is consistent with evidence that task characteristics influence mathematical word-problem performance ([Vessonen et al., 2024](#)). Future studies should pilot the items, counterbalance their order, document readability and contextual familiarity, and report item-level reliability evidence.

The results also help clarify how local wisdom relates to interdisciplinary design. A science, technology, engineering, arts, and mathematics perspective should do more than add a label to a culturally contextualized worksheet. Its instructional value lies in providing a rationale for examining production conditions, representing relationships, communicating solutions, and evaluating consequences within the same phenomenon. Recent reviews similarly emphasize coherent design and teacher readiness over collections of disconnected activities ([Deák & Kumar, 2024](#); [Kirabo et al., 2024](#); [Martín-Cudero et al., 2024](#); [Noralam et al., 2025](#)). Here, the local livelihood provided the shared phenomenon around which those activities were organized. Comparable ethnoscience work mapping a traditional Indonesian game onto physics concepts suggests that the same design principle can extend beyond mathematics to other STEM subjects ([Fiskawarni et al., 2025](#)).

This interpretation also gives local wisdom a more precise theoretical role. In this study, it is not a cultural illustration placed in front of a conventional exercise. Instead, it functions as a representational and epistemic bridge between community livelihood knowledge and algebraic modeling. Pupils first make sense of quantities and relationships in a familiar practice, formalize them as equations, solve the equations, and then return to the original practice to judge whether the solution is plausible. The Polya-stage rubric makes this sequence visible and helps identify where that bridge is strong and where it needs further support. In this sense, the study links a cultural resource to observable cognitive processes rather than relying on a general

claim of contextual relevance.

The educational implications are correspondingly specific. Teachers can draw on community livelihoods as sources of authentic data and discussion, but each task should include clear prompts for selecting information, forming equations, solving them, and checking the result. Curriculum developers can support this work through validated local data sources, guidance for task design, and analytic scoring rubrics. Further research should examine the sequence with comparison groups, larger and more varied samples, delayed posttests, and psychometrically evaluated item banks. These studies would help establish the durability and relative effectiveness of the observed pattern while respecting local communities' authority over the representation of their knowledge.

V. CONCLUSION

This study found clear within-class improvements in pupils' mathematical problem-solving performance following the local salt-production ethnomathematics learning sequence. Overall, KPM increased from 14.84% at pretest to 58.11% at posttest, with a medium normalized gain of 0.51, and significant improvements were observed across all four Polya stages. The largest gains occurred in devising a plan (D2; N-Gain = 0.66) and carrying out the plan (D3; N-Gain = 0.53), while looking back (D4) remained the weakest posttest stage despite showing improvement. The performance distribution also changed substantially, with all 32 pupils initially classified in the low category, compared with 9 pupils (28.1%) in the high category, 16 (50.0%) in the medium category, and 7 (21.9%) remaining low at posttest. The two planned parallel tasks also produced significantly different posttest scores, with higher performance on item 1 than on item 2, indicating that similar contextual tasks may differ in difficulty or interpretive demand. These results show that the learning sequence was associated with meaningful progress in mathematical modeling and solution processes, while verification remained the area requiring the greatest additional instructional support

The main limitation of this study is the absence of a comparison group, which means that the observed improvements should be interpreted as within-class changes rather than as definitive evidence of the learning sequence's comparative effectiveness. Future research may address this limitation by involving comparison conditions across broader educational settings to examine the consistency and relative effectiveness of the approach. Further studies may also explore how teachers adapt local salt-production contexts to different mathematical topics and classroom characteristics. Within these boundaries, the study demonstrates how a familiar

coastal livelihood can be operationalized as an epistemic bridge between local knowledge and formal algebraic reasoning. It also contributes a Polya-stage assessment framework that helps identify where culturally grounded mathematics learning is most effective and where additional instructional support is needed, particularly in verification and reflection.

Acknowledgments

The authors thank the participating school, teacher, and pupils in Jenepono Regency for their cooperation in the classroom intervention and data collection underlying this study.

Author contributions

Syaharuddin contributed to conceptualization, methodology, investigation, formal analysis, data curation, writing – original draft, and project administration. Siti Rahaimah Binti Ali contributed to validation, methodological review, supervision, and writing – review & editing. Agustan Syamsuddin contributed to supervision, validation, interpretation of the findings, and writing – review & editing. Nursakiah contributed to investigation support, resources, data curation, and writing – review & editing. Randy Saputra Mahmud contributed to formal analysis, visualization, interpretation, and writing – review & editing. All authors reviewed and approved the final manuscript and agreed to be accountable for the integrity of the work.

Funding

This research received no external funding.

Data availability

The pretest and posttest scoring data analyzed in this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval

This study was conducted in accordance with established ethical principles for research involving human participants and the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board (IRB) of Institut Turatea Indonesia prior to data collection.

Consent to participate

Informed consent was obtained from students' parents or legal guardians prior to participation.

Consent to publication

Not applicable.

Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Declaration of Generative AI and AI-Assisted Technologies

During the preparation of this manuscript, the authors used Claude Sonnet 4.5 (Anthropic) to improve the language, grammatical accuracy, clarity, coherence, and overall readability of the manuscript. The tool was also used to produce the illustration presented in Figure 1 based on the authors' original draft. Following the use of this tool, the authors carefully reviewed, verified, and edited the AI-assisted content as necessary to ensure its accuracy, consistency, and scientific integrity. The authors take full responsibility for the content of the published article.

REFERENCES

- Batiibwe, M. S. K. (2025). Ethnomathematics as a pedagogical tool for mathematics education: Opportunities and challenges. *SN Social Sciences*, 5(12), Article 221. <https://doi.org/10.1007/s43545-025-01260-0>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
- D'Ambrosio, U. (2018). The program ethnomathematics: Cognitive, anthropological, historic and socio-cultural bases. *PNA*, 12(4), 229–247. <https://doi.org/10.30827/pna.v12i4.7851>
- Deák, C., & Kumar, B. (2024). A systematic review of STEAM education's role in nurturing digital competencies for sustainable innovations. *Education Sciences*, 14(3), Article 226. <https://doi.org/10.3390/educsci14030226>
- Fiskawarni, T. H., Bancong, H., & Sukmawati, S. (2025). STEAM-based authentic assessment in improving physics students' critical thinking skills: A systematic literature review. *PPSDP International Journal of Education*, 4(2), 395–409. <https://doi.org/10.59175/pijed.v4i2.733>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Jupri, A., & Drijvers, P. (2016). Student difficulties in mathematizing word problems in algebra. *Eurasia Journal of Mathematics, Science and Technology Education*, 12(9), 2481–2502. <https://doi.org/10.12973/eurasia.2016.1299a>
- Kadir, H., & Rosmini, R. (2025). Annabba as cultural pedagogy: Exploring local wisdom values in Bugis-Makassar wedding traditions. *Journal of Local Wisdom in Education*, 1(1), 33–47. <https://doi.org/10.26618/3ep9va76>
- Kemendikdasmen. (2026). *Rapor pendidikan Kabupaten Jeneponto tahun 2025* [Education report card of Jeneponto Regency, 2025]. <https://raporpendidikan.kemendikdasmen.go.id/>
- Kirabo, E., Batiibwe, M. S. K., Mbulankende, J. S., & Lie, J. (2024). Technology-supported problem-based learning in mathematics education for pre-service teachers: A systematic literature review. *SN Social Sciences*, 4, Article 173. <https://doi.org/10.1007/s43545-024-00973-y>
- Kurniati, Nurzamsinar, A., & Amalia, T. (2025). Traditional games as culturally responsive pedagogy: Exploring boy-boy game for character building in Pancasila education. *Journal of Local Wisdom in Education*, 1(1), 72–86. <https://doi.org/10.26618/pwgehq16>
- Martín-Cudero, D., Cid-Cid, A. I., & Guede-Cid, R. (2024). Analysis of mathematics education from a STEAM approach at secondary and pre-university educational levels: A systematic review. *Journal of Technology and Science Education*, 14(2), 507–528. <https://doi.org/10.3926/jotse.2349>
- Noralam, N., Akib, I., & Bancong, H. (2025). Teachers' readiness for STEAM and differentiated teaching: Its connection to students' critical thinking in mathematics. *PPSDP International Journal of Education*, 4(2), 1384–1395.

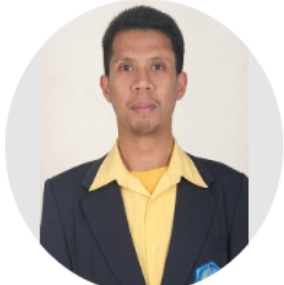
<https://doi.org/10.59175/pijed.v4i2.775>

- Nurfadila, M. Y., Nurlina, N., & Ma'ruf, M. (2023). Analysis of critical thinking skills of elementary school students through integrated problem-based learning model with mind mapping. *EDUKASIA: Jurnal Pendidikan dan Pembelajaran*, 4(2), 1373–1380. <https://doi.org/10.62775/edukasia.v4i2.445>
- OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Pathuddin, H., Kamariah, K., & Nawawi, M. I. (2021). Buginese ethnomathematics: Barongko cake explorations as mathematics learning resources. *Journal on Mathematics Education*, 12(2), 295–312. <https://jme.ejournal.unsri.ac.id/index.php/jme/article/view/3724>
- Polya, G. (1981). *Mathematical discovery: On understanding, learning, and teaching problem solving*. John Wiley & Sons.
- Prahmana, R. C. I., & D'Ambrosio, U. (2020). Learning geometry and values from patterns: Ethnomathematics on the batik patterns of Yogyakarta, Indonesia. *Journal on Mathematics Education*, 11(3), 439–456. <https://jme.ejournal.unsri.ac.id/index.php/jme/article/view/3746/279>
- Pratama, R. A., & Yanpar Yelken, T. (2024). Effectiveness of ethnomathematics-based learning on students' mathematical literacy: A meta-analysis study. *Discover Education*, 3(1), Article 202. <https://doi.org/10.1007/s44217-024-00309-1>
- Salmia, S., & Malik, A. (2025). Embedding local wisdom in joyful mathematics learning to strengthen early-grade numeracy: Evidence from the GEMBIRA approach. *Journal of Local Wisdom in Education*, 1(2), 180–194. <https://doi.org/10.65508/x2hdya18>
- Salmia, S., Nursalam, N., & Bancong, H. (2024a). Design and validation of teaching modules for the Merdeka Curriculum based on local wisdom to instil character values. *PPSDP International Journal of Education*, 3(2), 659–676. <https://doi.org/10.59175/pijed.v3i2.353>
- Salmia, S., Nursalam, N., & Bancong, H. (2024b). Effectiveness of local wisdom-based independent curriculum teaching modules in improving learning outcomes in Indonesia. *Journal of Ecohumanism*, 3(6), 1719–1726. <https://doi.org/10.62754/joe.v3i6.4131>
- Santos-Trigo, M. (2024). Problem solving in mathematics education: Tracing its foundations and current research-practice trends. *ZDM–Mathematics Education*, 56, 211–222. <https://doi.org/10.1007/s11858-024-01578-8>
- Schoenfeld, A. H. (2016). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. *Journal of Education*, 196(2), 1–38. <https://doi.org/10.1177/002205741619600202>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Sibgatullin, I. R., Korzhuev, A. V., Khairullina, E. R., Sadykova, A. R., Baturina, R. V., & Chazova, V. (2022). A systematic review on algebraic thinking in education. *Eurasia Journal of Mathematics, Science and Technology Education*, 18(1), Article em2065. <https://doi.org/10.29333/ejmste/11486>
- Syahrudin, S. (2019). Description of students' mathematics problem-solving relation in concept understanding from evaluated learning styles at Class VIII SMPN 4 Binamu

Jeneponto District. *Celebes Education Review*, 1(1), 12–18.
<https://garuda.kemdiktisaintek.go.id/documents/detail/992651>

Vessonen, T., Dahlberg, M., Hellstrand, H., Widlund, A., Korhonen, J., Aunio, P., & Laine, A. (2024). Task characteristics associated with mathematical word problem-solving performance among elementary school-aged children: A systematic review and meta-analysis. *Educational Psychology Review*, 36(4), Article 117.
<https://doi.org/10.1007/s10648-024-09954-2>

About the authors

	Syaharuddin is a lecturer in the Mathematics Education Study Program at Institut Turatea Indonesia, Jeneponto, South Sulawesi, Indonesia. He is also a doctoral student in the Doctoral Program of Educational Science at Universitas Muhammadiyah Makassar, Indonesia. His research interests include ethnomathematics, local wisdom-based learning, Problem-Based Learning integrated with STEAM, and students' mathematical problem-solving ability. His work focuses on connecting culturally familiar contexts with formal mathematics learning. His current research particularly examines how local contexts can support the stages of mathematical problem solving. Email: syahr2@gmail.com ORCID ID: https://orcid.org/0000-0001-9830-3269
	Siti Rahaimah Binti Ali is an Associate Professor at the Faculty of Human Development, Universiti Pendidikan Sultan Idris (UPSI), Tanjung Malim, Perak, Malaysia. She earned her Ph.D. in Mathematics Education from Universiti Pendidikan Sultan Idris in 2014. She previously completed a master's degree in mathematics education with Information Communication Technology at Universiti Malaya in 2006 and a bachelor's degree in science and computer with Education Mathematics at Universiti Teknologi Malaysia in 2004. Her research interests include numeracy understanding models, mathematics education values, gamification and digital learning modules, and STEM/STEAM-based instruction. Her scholarship connects mathematics education with technology-enhanced learning and interdisciplinary approaches. Email: siti.rahaimah@fpm.upsi.edu.my ORCID ID: https://orcid.org/0000-0002-3458-3756
	Agustan Syamsuddin is a professor in the Mathematics Education Study Program, Faculty of Teacher Training and Education, Universitas Muhammadiyah Makassar, South Sulawesi, Indonesia. His research interests include cognitive psychology in mathematics learning. He also works on the development of mathematics learning models. Another area of interest is the design of instructional materials for mathematics education. His scholarly work further includes STEAM-based instruction and its application to mathematics learning. Email: agustan@unismuh.ac.id ORCID ID: https://orcid.org/0000-0003-2164-852X

	Nursakiah is a lecturer in the Mathematics Education Study Program at Universitas Muhammadiyah Makassar, South Sulawesi, Indonesia. Her research interests include mathematics learning models and mathematical communication ability. She has a particular interest in the teaching and learning of systems of linear equations in two variables (SPLDV). Her work considers how instructional models can support students in expressing mathematical ideas clearly. She also contributes to research on classroom approaches that strengthen mathematical reasoning and communication. Email: nursakiah@unismuh.ac.id ORCID ID: https://orcid.org/0009-0005-5524-2984
	Randy Saputra Mahmud is a doctoral student at the University of Szeged, Hungary, supported by the Stipendium Hungaricum Scholarship. He holds a degree in Mathematics Education from Universitas Negeri Makassar, Indonesia. His research primarily focuses on students' mathematical problem-solving abilities, with particular emphasis on Pólya's four-stage problem-solving framework. His research interests also encompass students' difficulties in learning mathematics and the relationship between learning styles and mathematical problem-solving performance. Email: randy.saputra.mahmud@edu.u-szeged.hu ORCID ID: https://orcid.org/0000-0001-9407-7671