Research and Development Journal Of Education
Vol. 12, No. 2, Oktober 2026, Pp: 993 – 1000 p–ISSN 2406-9744
DOI: https://doi.org/10.30998 e–ISSN 2657-1056
Available online at https://journal.lppmunindra.ac.id/index.php/RDJE

PROBLEM-BASED LEARNING OR CREATIVE PROBLEM SOLVING? EXPLORING STUDENTS' MATHEMATICAL PROBLEM-POSING ABILITY

Arista Lindi Alvita[1(*)], Isnandar Slamet[2], Budi Usodo[3]
Universitas Sebelas Maret, Surakarta, Indonesia[123]
aristalindi@student.uns.ac.id[1]

Received: 28 Oktober 2026
Revised: 28 Oktober 2026
Accepted: 29 Oktober 2026

Abstract

Mathematical problem posing is an essential component of mathematics education that encourages students to formulate, explore, and refine mathematical questions. However, empirical evidence regarding how different instructional models support students' mathematical problem-posing ability remains limited. This study aimed to compare students' mathematical problem-posing performance under Problem-Based Learning (PBL) and Creative Problem Solving (CPS) in vocational secondary education. A quantitative quasi-experimental approach employing a pretest–posttest design was implemented in two intact eleventh-grade classrooms at a vocational school in Indonesia. Both groups participated in instructional activities involving sequences and series, compound interest, and annuities. Mathematical problem-posing ability was assessed through constructed-response tasks focusing on fluency, flexibility, and originality, with mathematical correctness incorporated into the scoring process. The findings indicated improvements in average problem-posing performance in both classrooms, with a greater descriptive increase observed in the PBL group. Nevertheless, the baseline-adjusted comparison did not provide sufficient evidence of a statistically significant difference between the instructional conditions. Given the initial differences between groups and the limited number of classrooms, these findings should be interpreted cautiously rather than as evidence of the causal superiority of either model. The study highlights the potential contribution of contextual problem investigation and creative idea generation to mathematical problem posing and emphasizes the importance of instructional design that provides students with opportunities to formulate, evaluate, and refine mathematical questions.

Keywords: Creative Problem Solving; Mathematical Problem Posing; Problem-Based Learning; Quasi-Experimental Design

(*) Corresponding Author: Alvita, aristalindi@student.uns.ac.id

How to Cite: Alvita, A. L., Slamet, I., & Usodo, B. (2026). PROBLEM-BASED LEARNING OR CREATIVE PROBLEM SOLVING? EXPLORING STUDENTS' MATHEMATICAL PROBLEM-POSING ABILITY. *Research and Development Journal of Education*, 12(2), 993-1000.

INTRODUCTION

Contemporary mathematics education asks learners to do more than recognize a familiar exercise and reproduce a demonstrated algorithm. Students also need to interpret information, determine what can reasonably be asked, explain why relationships matter, and test whether their mathematical conclusions are defensible (Cai & Hwang, 2020; Singer et al., 2013). Mathematical problem posing provides a particularly useful lens on these activities because it offers opportunities to investigate students' mathematical reasoning, conceptual understanding, and creativity (Cai et al., 2013; Silver, 1994). In

problem posing, the learner creates a new question, modifies an existing one, or constructs a mathematical investigation from a representation or situation (Cai & Hwang, 2020; Silver, 1994). Unlike a conventional exercise with a predetermined question, the posing task places some responsibility for defining the mathematical problem on the learner (Singer et al., 2013). This responsibility may reveal connections and misunderstandings that remain hidden when students only calculate answers to teacher-supplied questions (English, 2020; Silver, 1994).

Silver (1994) described problem posing as a central mathematical activity associated with problem solving and mathematical inquiry. The connection is reciprocal rather than merely sequential. Solving a problem may generate new questions about changed conditions, additional constraints, or alternative representations; conversely, posing a meaningful question may require anticipating how it could be solved. This relationship matters in secondary mathematics because students can appear competent at routine calculations yet struggle to identify a mathematically coherent, worthwhile question. Problem posing therefore warrants examination as an outcome in its own right rather than as an incidental feature of problem-solving lessons.

A question-generation task can take several forms. Students may formulate a problem from an open situation, complete a partially specified question, or reformulate a solved example by altering its conditions. A student considering a savings account might ask for the accumulated balance after a given number of years, compare the outcomes of two interest rates, or work backward to identify the initial deposit needed to reach a target. Although these questions share a context, they do not impose identical mathematical demands. The first may require straightforward substitution; the second involves comparison and interpretation; the third may require rearranging a relationship. Such differences illustrate why counting questions alone cannot fully describe problem-posing competence. Research on middle-school students' arithmetic problem posing also shows why the mathematical quality of generated questions warrants attention (Silver & Cai, 1996).

Cai & Rott (2024) conceptualized problem posing as a process involving orientation, connection, generation, and reflection. This perspective directs attention to what happens before and after a question is written. A learner must first make sense of the available situation, identify potentially useful relationships, generate a mathematical question, and evaluate whether the question is clear and viable. The process may be iterative: reflection on an initial question can lead to revisions, greater precision, or a different mathematical focus. Instruction that supports these stages may be valuable even when its immediate activities are labeled problem solving rather than problem posing. A complementary phase model includes situation analysis, variation, generation, problem solving, and evaluation (Baumanns & Rott, 2022).

Research syntheses suggest that problem-posing interventions can support mathematical learning (Wang et al., 2022; Zhang et al., 2024). Nevertheless, the instructional setting matters. PBL organizes learning around the investigation of contextual problems, whereas CPS explicitly emphasizes understanding challenges, generating alternatives, and evaluating solutions (Isaksen & Treffinger, 2004). Both approaches could create opportunities to pose mathematical problems, but their relative effects cannot be inferred from their theoretical characteristics alone.

The present investigation focuses on problem posing as an assessed outcome in two Indonesian vocational-school classrooms studying financial applications of sequences and series. It asks: (1) What descriptive pretest–posttest patterns were observed in the PBL and CPS classrooms? (2) How did posttest scores compare after adjusting for baseline performance? This classroom comparison is exploratory, not a causal effectiveness trial.

METHODS

Design and participants

A quantitative two-group pretest–posttest design was used at SMK Al-Manaar Pemalang, Central Java, Indonesia. Two existing grade XI classrooms were allocated to PBL and CPS, respectively, with 20 students in each. Allocation occurred at the classroom level rather than at the individual level. One mathematics teacher taught both groups over four 90-minute meetings. Only one classroom represented each condition, preventing separation of classroom and instructional-model effects.

The pretest–posttest structure allowed observed scores to be summarized before and after the instructional period. It also allowed an exploratory adjusted comparison in which initial scores were included as a covariate. However, random assignment of the two existing classes to the model labels does not make the individual students randomized participants, nor does it eliminate possible differences between the classes. Classroom culture, peer interactions, schedules, prior learning experiences, and other unmeasured factors may be associated with the outcomes. These limitations were considered throughout the analysis and interpretation.

Table 1.
Instruction Classroom

Group	Pretest	Instruction	Posttest
Classroom 1	O1	PBL	O2
Classroom 2	O3	CPS	O4

Instructional procedures

Both groups studied comparable content involving sequences and series, compound interest, and annuities. PBL lessons followed orientation to a contextual problem, organization for learning, investigation, presentation, and reflection. CPS lessons followed understanding/problem finding, idea finding, and preparing/solution acceptance. Students examined financial situations and were given opportunities to formulate, solve, and discuss mathematical questions. Exact fidelity to these instructional stages should be checked against the implemented lesson plans.

Each classroom received four instructional meetings of 90 minutes, corresponding to eight 45-minute lesson periods. The total scheduled instructional time was therefore 360 minutes per classroom. Both groups worked with sequences and series and applications involving compound interest and annuities. These applications were intended to provide mathematical situations from which students could generate questions and investigate relationships. The short duration is relevant to interpretation: a four-meeting sequence can reveal immediate performance patterns but cannot establish the persistence of learning or longer-term development of problem-posing habits.

The same teacher implemented both instructional approaches. This arrangement may have supported greater consistency in subject-matter explanations and classroom expectations, although it also introduced the possibility that the teacher's familiarity or comfort with one model influenced enactment. Information about fidelity observation, attendance, or the exact order and dates of meetings should be inserted only after checking the original lesson plans and observation records. The absence of such verified information is a reporting limitation rather than grounds for inventing procedural details.

Instrument and scoring

The pretest assessed problem posing in sequences and series; the posttest used compound-interest and annuity situations. Five items were initially considered for each

assessment; four were retained following item screening. The rubric addressed fluency, flexibility, and originality, alongside mathematical correctness. Final publication requires documentation of item validity, reliability, scoring aggregation, and evidence of comparability across the two assessments. Because the contexts differ, raw score change is interpreted descriptively.

Data analysis

Means and standard deviations were summarized by classroom and occasion. Baseline differences were examined using a Welch comparison. An exploratory student-level ANCOVA modeled posttest score as a function of classroom condition and pretest score. Diagnostic checks considered residual normality, heterogeneity of variance, and pretest-by-condition interaction. HC3 heteroskedasticity-consistent standard errors were examined as a sensitivity analysis. Student-level p-values are descriptive diagnostics here: they cannot support a valid classroom-treatment causal test with only one cluster per condition.

RESULTS & DISCUSSION

Results

Descriptive performance

The two classrooms showed different descriptive patterns. The PBL group began with a mean pretest score of 22.55 (SD = 8.64) and had a posttest mean of 33.40 (SD = 7.34), an observed mean change of 10.85 points. The CPS group began with a higher mean pretest score of 30.00 (SD = 7.21) and had a posttest mean of 32.95 (SD = 12.31), an observed mean change of 2.95 points. Both groups therefore had higher average scores at the later assessment, but their starting points and score distributions were different.

The raw difference between the changes was 7.90 points in favor of the PBL classroom. This numerical contrast is not a randomized treatment effect. Students in the PBL classroom began substantially lower, leaving more room for an upward change on the observed score scale, while students in the CPS classroom started higher. The two assessments also used different mathematical contexts. These facts make a direct causal interpretation of the raw changes inappropriate. The result is better described as a pattern of observed performance over two occasions.

The CPS classroom exhibited a larger posttest standard deviation than the PBL classroom. Its posttest scores were therefore more dispersed around the mean. This descriptive observation may reflect differences among students in their responses to the assessment, but the available summary data do not identify whether the dispersion arose from fluency, flexibility, originality, correctness, or particular items. No distributional claims beyond the reported means and standard deviations are warranted without inspection of the individual scores.

Table 2.
Standard Deviation

Group	N	Pretest mean (SD)	Posttest mean (SD)	Raw mean change
PBL	20	22.55 (8.64)	33.40 (7.34)	+10.85
CPS	20	30.00 (7.21)	32.95 (12.31)	+2.95

Both classrooms showed higher mean posttest than pretest scores. The larger raw change in PBL should be interpreted alongside its lower baseline mean. The CPS classroom displayed greater posttest variability.

Baseline comparison and model diagnostics

The previously reported Welch comparison indicated a baseline difference between the classrooms, with a p-value of approximately 0.005. The CPS classroom had the higher initial mean. This difference is substantively important because the initial distribution of the outcome is related to the interpretation of posttest comparisons. It also demonstrates that assigning existing classes to instructional labels does not guarantee equivalence of the students' observed baseline scores.

The previously reported ANCOVA diagnostics did not detect a departure from residual normality using the Shapiro–Wilk test ($p = 0.933$). The reported pretest-by-group interaction was not statistically significant ($p = 0.199$), indicating that the available test did not detect different slopes. A nonsignificant diagnostic, however, is not proof that an assumption is exactly satisfied, especially in a modest sample. Visual residual checks and sensitivity to influential observations would strengthen the analysis.

Levene's test was reported as statistically significant ($p = 0.002$), indicating unequal residual variances. This is consistent with the larger posttest spread observed in the CPS classroom, although the two findings are not identical. The violation means that the conventional homoskedastic student-level standard errors may be unreliable. An HC3 sensitivity analysis was therefore considered. Neither the conventional nor robust approach addresses the more fundamental lack of classroom-level replication.

Table 3.
Diagnostic Tests of ANCOVA Assumptions for Mathematical Problem-Posing Scores

Diagnostic	Previously reported p	Reading
Residual Shapiro–Wilk	0.933	No departure detected
Levene variance homogeneity	0.002	Unequal variances indicated
Pretest × group interaction	0.199	No slope difference detected

Exploratory adjusted comparison

The previously reported ANCOVA estimated adjusted posttest means of 36.04 for PBL and 30.31 for CPS. The adjusted numerical difference was therefore 5.73 points. The conventional student-level comparison was reported as $F(1, 37) = 3.659$, $p = 0.064$, with partial eta squared of 0.090. Under a nominal 0.05 criterion, this comparison did not reach statistical significance. The adjusted means should be understood as model-based estimates at a common pretest value rather than as directly observed classroom averages.

Because residual variance homogeneity was not supported, a previously reported HC3 analysis was used as a sensitivity check. Its p-value was 0.116, and the reported 95% confidence interval for the adjusted contrast was approximately −1.49 to 12.95. The interval included zero and encompassed differences in either direction. The robust result therefore did not support a precise conclusion that one classroom had a higher conditional mean than the other. The interval also indicates that nonsignificance is not evidence of equivalence; a range of differences remained compatible with the fitted student-level model.

Partial eta squared of 0.090 is a descriptive measure associated with the fitted model and should not be translated into a claim that the instructional model explained nine percent of causal improvement. The value depends on the chosen outcome, covariate, sample, and model specification. More importantly, group and classroom are identical in this design. The analysis does not permit a separation of the effect of a teaching model from any

unmeasured classroom characteristic. This qualification applies regardless of whether the nominal p-value is above or below 0.05.

Table 4.
Diagno Exploratory ANCOVA Results and HC3 Robust Sensitivity Analysis of Mathematical Problem-Posing Performance

Diagnostic	Previously reported p
Quantity	Previously reported result
Adjusted mean, PBL	36.04
Adjusted mean, CPS	30.31
Adjusted contrast (PBL − CPS)	5.73
Conventional ANCOVA	$F(1,37) = 3.659;\ p = 0.064$
Partial eta squared	0.090
HC3 robust p	0.116

The adjusted mean was numerically higher for PBL, but neither the conventional nor robust student-level analysis reached the nominal 0.05 threshold. The confidence interval included zero. These results do not establish equivalence, nor do they justify a causal claim that one instructional model is more effective.

Discussion

Interpretation of observed patterns

The PBL classroom displayed a larger descriptive increase, but its substantially lower baseline performance complicates comparisons of raw change. Baseline-adjusted results were inconclusive. The observed difference could reflect instructional experiences, classroom characteristics, student composition, assessment conditions, or combinations of these factors. No mechanism was measured directly, so explanations remain tentative.

Instructional relevance of PBL and CPS

Contextual financial problems can invite students to vary interest rates, time periods, initial capital, or payment structures, producing multiple mathematically meaningful questions. PBL may offer such opportunities through investigation and reflection. CPS may offer related opportunities through explicit idea generation and refinement. Both align conceptually with problem-posing processes, but these connections should not be mistaken for demonstrated treatment effects. Research on problem posing highlights the importance of task characteristics, mathematical knowledge, and opportunities to evaluate generated questions (Baumanns & Rott, 2022; Cai & Rott, 2024; Leung & Silver, 1997).

Limitations and implications

The most consequential limitation is the absence of treatment replication: one classroom per condition makes classroom identity inseparable from model assignment. ANCOVA and HC3 standard errors cannot remedy this design limitation. Initial group differences, heterogeneous residual variances, distinct pretest and posttest contexts, a four-meeting intervention, and the absence of separate indicator-level analyses further restrict interpretation. Future studies should randomize multiple classrooms per condition, use parallel assessment forms, report scoring reliability, and examine fluency, flexibility, and originality separately. For practice, teachers can integrate explicit question generation and mathematical validation into contextual learning tasks without assuming either model is superior.

CONCLUSION

This exploratory study compared mathematical problem-posing performance in two intact vocational-school classrooms implementing Problem-Based Learning and Creative Problem Solving. Both classrooms had higher average posttest than pretest scores, and the PBL classroom showed a larger descriptive mean increase. However, the groups differed substantially at baseline, and the assessments used different mathematical contexts. The raw score patterns therefore cannot be interpreted as direct evidence of differential instructional effectiveness.

A previously reported student-level ANCOVA produced a numerically higher adjusted posttest mean for PBL, but neither the conventional comparison nor the heteroskedasticity-robust sensitivity analysis was statistically significant. These nonsignificant results do not establish equivalence. More fundamentally, because only one classroom represented each instructional condition, classroom and instructional-model effects could not be separated. The findings should be understood as descriptive and exploratory evidence from two particular learning settings.

The pedagogical contribution of the study lies in drawing attention to mathematical question generation within contextual inquiry and deliberate idea-development activities. Both PBL and CPS can provide opportunities for students to formulate, compare, and refine mathematical problems when instruction explicitly values these activities. Future research should use multiple classrooms per condition, stronger evidence of assessment comparability, longer interventions, and separate analyses of fluency, flexibility, originality, and mathematical correctness. These improvements would permit more informative conclusions about how instructional design relates to the development of mathematical problem-posing competence.

REFERENCES

Baumanns, L., & Rott, B. (2022). Developing a framework for characterising problem-posing activities: a review. *Research in Mathematics Education*, *24*, 28–50. https://doi.org/10.1080/14794802.2021.1897036

Cai, J., & Hwang, S. (2020). Learning to Teach Through Mathematical Problem Posing: Theoretical Considerations, Methodology, and Directions for Future Research. *International Journal of Educational Research*, *102*. https://doi.org/10.1016/j.ijer.2019.01.001

Cai, J., Moyer, J. C., Wang, N., Hwang, S., Nie, B., & Garber, T. (2013). Mathematical Problem Posing as a Measure of Curricular Effect on Students' Learning. *Educational Studies in Mathematics*, *83*, 57–69. https://doi.org/10.1007/s10649-012-9429-3

Cai, J., & Rott, B. (2024). On Understanding Mathematical Problem-Posing Processes. *ZDM - Mathematics Education*, *56*, 61–71. https://doi.org/10.1007/s11858-023-01536-w

English, L. D. (2020). Teaching and Learning Through Mathematical Problem Posing: Commentary. *International Journal of Educational Research*, *102*. https://doi.org/10.1016/j.ijer.2019.06.014

Isaksen, S. G., & Treffinger, D. J. (2004). Celebrating 50 years of reflective practice: Versions of creative problem solving. In *Journal of Creative Behavior* (Vol. 38, pp. 75–101). Creative Education Foudation. https://doi.org/10.1002/j.2162-6057.2004.tb01234.x

Leung, S. S., & Silver, E. A. (1997). The Role of Task Format, Mathematics Knowledge, and Creative Thinking on The Arithmetic Problem Posing of Prospective Elementary School Teachers. *Mathematics Education Research Journal*, *9*, 5–24. https://doi.org/10.1007/BF03217299

Silver, E. A. (1994). On Mathematical Problem Posing. *For the Learning of Mathematics*, *14(1)*, 19–28. https://www.jstor.org/stable/40248099

Silver, E. A., & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. *Journal for Research in Mathematics Education*, *27*, 521–539. https://doi.org/10.2307/749846

Singer, F. M., Ellerton, N., & Cai, J. (2013). Problem-Posing Research in Mathematics Education: New Questions and Directions. *Educational Studies in Mathematics*, *83*, 1–7. https://doi.org/10.1007/s10649-013-9478-2

Wang, M., Walkington, C., & Rouse, A. (2022). A Meta-Analysis on the Effects of Problem-Posing in Mathematics Education on Performance and Dispositions. *Investigations in Mathematics Learning*, *14*, 265–287. https://doi.org/10.1080/19477503.2022.2105104

Zhang, L., Stylianides, G. J., & Stylianides, A. J. (2024). Enhancing Mathematical Problem Posing Competence: A Meta-Analysis of Intervention Studies. In *International Journal of STEM Education* (Vol. 11). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1186/s40594-024-00507-1.