International Journal of Education, Vocational and Social Science
https://e-journal.citakonsulfindo.or.id/index.php/IJEVSS
E-ISSN : 2963 4954
IJEVSS

The Effectiveness Of Metareflection-Based Problem-Based Learning On Students' Mathematical Critical Thinking Skills

Baso Intang Sappaile
Universitas Negeri Makassar, Indonesia
Norma Nasir
Universitas Negeri Makassar, Indonesia

Corresponding Author: norma.nasir@unm.ac.id

Abstract

Students' mathematical critical thinking skills remain weak, and problem-based learning (PBL) as commonly practised rarely includes a deliberate stage in which students examine their own reasoning. This study aimed to determine the effectiveness of a metareflection-based PBL model on students' mathematical critical thinking skills. A quasi-experimental nonequivalent control group pretest-posttest design was employed with 60 students assigned to an experimental class (n = 30) and a control class (n = 30). Data were collected through a mathematical critical thinking test, a learning implementation observation sheet, and a student response questionnaire, and were analysed using descriptive statistics, normalized gain, and t-tests. Effectiveness was judged against four criteria: a mean posttest score above the minimum mastery criterion of 75, classical mastery of at least 80%, learning implementation in the good category, and positive student responses. The experimental class achieved a mean posttest score of 81.60, significantly above 75 ($t(29) = 5.98$, $p < .001$), classical mastery of 86.67%, learning implementation of 75.7% (good), and a positive response level of 81.42%. It also outperformed the control class ($t(58) = 3.99$, $p < .001$, $d = 1.03$). Integrating metareflection into PBL is therefore effective in developing students' mathematical critical thinking.

Keywords: Problem-Based Learning; Metareflection; Mathematical Critical Thinking; Quasi-Experiment; Learning Effectiveness.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

International Journal of Education, Vocational and Social Science
https://e-journal.citakonsulfindo.or.id/index.php/IJEVSS
E-ISSN : 2963 4954
IJEVSS

INTRODUCTION

Critical thinking is widely recognised as one of the central goals of school education because it enables learners to decide reasonably what to believe and what to do (Ennis, 2018). In mathematics, critical thinking appears when students interpret the information in a problem, analyse the relationships among its elements, evaluate the validity of arguments and solution strategies, and draw justified inferences (Facione, 2015). These skills are not by-products of routine computation; they must be cultivated through learning experiences that require students to reason, justify, and question their own answers (Su et al., 2016). International assessments indicate that this goal has not yet been fully realised. In PISA 2022, Indonesian students obtained a mean mathematics score of 366, well below the OECD average, and most students did not reach the baseline proficiency level at which they can interpret and represent simple situations mathematically (OECD, 2023). This evidence suggests that many students still struggle with the reasoning processes that underlie mathematical critical thinking.

Problem-based learning (PBL) is one of the most frequently recommended models for addressing this problem. In PBL, learning begins with an ill-structured, contextual problem that students investigate collaboratively, so that knowledge is constructed while students analyse, hypothesise, and justify solutions (Hmelo-Silver, 2004; Yew & Goh, 2016). A meta-analysis of PBL interventions found a positive effect on students' critical thinking (Liu & Pásztor, 2022), and syntheses of mathematics education studies have reported large effects of PBL and similar innovative models on mathematical abilities and critical thinking (Sa'idah et al., 2026; Suparman et al., 2021). Studies in Indonesian classrooms show a consistent pattern: PBL improves students' mathematical thinking across different levels of curiosity (Zetriuslita & Ariawan, 2021), and critical thinking is among the mathematical abilities most strongly influenced by PBL at the secondary level (Ajinegara & Nuriadin, 2022).

Nevertheless, the success of PBL depends heavily on what students do after they obtain a solution. A systematic review of critical thinking-oriented PBL found that successful adaptations deliberately add activities that support interaction, self-directed learning, and reflection rather than relying on the problem alone (Yu & Zin, 2023). In mathematics, students often stop once they reach an answer; they rarely plan their approach, check whether their method is efficient, or connect the concepts they used, and they tend to focus on whether the final answer is correct (Kholid et al., 2020). Research on reflective thinking in mathematics has grown, but it remains dominated by qualitative descriptions of students' reflective levels rather than by interventions

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

that systematically train reflection within a learning model (Hasanah et al., 2024). Without a structured opportunity to reflect, the critical thinking demanded by PBL may remain at the level of task completion rather than becoming a conscious habit of reasoning.

Metareflection offers a way to close this gap. It refers to students' reflection on their own thinking: reviewing the strategies they used, identifying errors in their reasoning, and evaluating the depth of their understanding. This idea is rooted in metacognition, that is, the monitoring and regulation of one's own cognition (Flavell, 1979; Stanton et al., 2021). Previous studies show that adding metacognitive prompts to PBL improves secondary students' argumentation and critical thinking more than PBL without such prompts (Marthaliakirana et al., 2022), that a metacognitive learning approach improves students' mathematical critical thinking compared with conventional instruction (Syaiful et al., 2022), and that reflection journals increase students' activeness and mathematics learning outcomes (Faizin, 2018). Conceptually, metacognition and critical thinking reinforce each other in a cycle in which reflection on reasoning strengthens the quality of subsequent reasoning (Istiqlal et al., 2025).

Although PBL and metacognitive strategies have each been studied, empirical studies that embed an explicit metareflection phase as part of the PBL syntax in mathematics and evaluate its effectiveness against several criteria at once remain limited. Most PBL studies judge success from a single significance test or a comparison of means, whereas the effectiveness of a learning model in classroom practice also concerns whether students reach the mastery standard, whether the model can be implemented as designed, and whether students respond positively to it. This study therefore aimed to determine the effectiveness of metareflection-based PBL on students' mathematical critical thinking skills. Effectiveness was examined through four indicators: (1) the mean posttest score of the experimental class exceeds the minimum mastery criterion (KKTP) of 75; (2) classical mastery reaches at least 80%; (3) learning implementation is in the good category; and (4) student responses are positive. In addition, the study compared the critical thinking outcomes of students taught with metareflection-based PBL and those taught with conventional instruction.

LITERATURE REVIEW

Mathematical Critical Thinking

Ennis (2018) describes critical thinking as reasonable and reflective thinking focused on deciding what to believe or do. The expert consensus summarised by Facione (2015) identifies

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

interpretation, analysis, evaluation, inference, explanation, and self-regulation as core critical thinking skills. In mathematics learning, these skills are expressed when students identify relevant information and conditions in a problem, analyse relationships among concepts, evaluate the correctness and efficiency of strategies, and draw and justify conclusions. Su et al. (2016) emphasise that critical thinking in mathematics develops when teachers ask students to explain their reasoning, compare strategies, and examine why a solution works, which links critical thinking closely to metacognition. Self-regulation, the last skill in Facione's list, is essentially the capacity to monitor and correct one's own reasoning, and it is the element most directly targeted by metareflection.

Problem-Based Learning

PBL is a student-centred model in which learning is organised around authentic problems that have no single predetermined path to a solution (Hmelo-Silver, 2004). Its process typically moves from problem analysis and the identification of learning needs to self-directed study, collaborative discussion, and the synthesis of a solution (Yew & Goh, 2016). Arends (2012) organises PBL into five phases: orienting students to the problem, organising students for learning, guiding individual and group investigation, developing and presenting results, and analysing and evaluating the problem-solving process. Empirical evidence consistently supports PBL for higher-order thinking. Liu and Pásztor (2022) reported a positive effect of PBL interventions on critical thinking, Bron and Prudente (2024) found a moderate positive effect of PBL on mathematical creativity ($g = 0.58$), and Suparman et al. (2021) found a large effect of PBL on primary students' mathematical abilities ($g = 1.05$). However, Yu and Zin (2023) showed that PBL models that successfully promote critical thinking usually include additional structured activities, which indicates that the basic PBL phases may not be sufficient on their own.

Metareflection in Mathematics Learning

Metacognition comprises knowledge about one's own cognition and the regulation of cognition through planning, monitoring, and evaluating (Flavell, 1979). Stanton et al. (2021) show that students rarely engage in these processes spontaneously and that instructors can foster them through explicit prompts and reflection activities. Metareflection, as used in this study, is a structured form of metacognitive reflection in which students look back on how they solved a problem: which strategy they chose and why, where they made mistakes, and what they now

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

understand differently. Merkebu et al. (2025) describe metacognitive reflection as a practice that prompts learners to rethink and adopt new perspectives, thereby supporting deeper learning. In mathematics, reflective thinking helps students handle confusion and connect concepts during problem solving (Kholid et al., 2020), and reflection journals used as a metacognitive strategy have been reported to increase students' activeness and learning outcomes (Faizin, 2018). Istiqlal et al. (2025) further argue that metacognition and mathematical critical thinking form a mutually reinforcing cycle, so a learning stage that trains reflection is expected to strengthen critical thinking.

Metareflection-Based PBL and Criteria of Effectiveness

In this study, metareflection-based PBL extends the five PBL phases of Arends (2012) with a sixth phase, metareflection, in which students review their strategies, errors, and understanding after the solution has been evaluated. This design follows evidence that PBL combined with metacognitive prompts produces greater gains in critical thinking than PBL alone (Marthaliakirana et al., 2022) and that metacognitive approaches improve mathematical critical thinking across ability levels (Syaiful et al., 2022). The effectiveness of a learning model is judged not only by whether it produces statistically significant gains but also by whether learning outcomes reach the expected standard and whether the model works in actual classroom conditions. For this reason, this study used four criteria of effectiveness, namely individual achievement above the minimum mastery criterion, classical mastery, the quality of learning implementation, and student responses, supported by a comparison with a control class.

RESEARCH METHOD

Research Design

This study used a quantitative approach with a quasi-experimental method. The design was a nonequivalent control group pretest-posttest design (Creswell & Creswell, 2018; Fraenkel et al., 2019), as presented in Table 1. Two intact classes were used because students could not be randomly reassigned to new classes. The experimental class was taught using metareflection-based PBL, while the control class was taught using conventional instruction. Both classes received a pretest before the intervention and a posttest after it.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Table 1. Research Design

Class	Pretest	Treatment	Posttest
Experimental	O1	X (metareflection-based PBL)	O2
Control	O3	- (conventional instruction)	O4

Participants

The participants were 60 Grade X students at MA Arifah Gowa, South Sulawesi, Indonesia, during the 2026/2027 academic year. The sample was selected using cluster random sampling, in which two classes were selected as research groups. One class was assigned as the experimental group (n = 30), which received instruction using the Problem-Based Learning (PBL) model integrated with metareflection, while the other class served as the control group (n = 30) and received conventional mathematics instruction. The learning material focused on exponents and was delivered over six meetings in the experimental group.

Instruments

Three instruments were used to collect the data. The first was a mathematical critical thinking test consisting of five essay items designed to assess students' abilities in interpretation, analysis, evaluation, and inference, drawing on Facione's (2015) framework. Students' scores were converted to a scale of 0–100. The test was validated by two experts and piloted before implementation. Its validity and reliability were assessed using appropriate statistical procedures, with the resulting validity evidence and reliability coefficient reported in the study.

The second instrument was a learning implementation observation sheet covering six aspects aligned with the syntax of metareflection-based PBL: (1) orienting students to the problem, (2) organising students for learning, (3) guiding individual and group investigation, (4) developing and presenting results, (5) analysing and evaluating the problem-solving process, and (6) conducting metareflection to review strategies, errors, and understanding. Each aspect was rated on a four-point scale during each of the six meetings.

The third instrument was a student response questionnaire comprising ten positively worded statements rated on a four-point Likert scale (1 = strongly disagree, 2 = disagree, 3 =

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

agree, and 4 = strongly agree). The questionnaire examined students' responses to the attractiveness of the problems, contextual relevance, group discussion, participation, strategy selection, answer evaluation, the role of metareflection, confidence in solving mathematical problems, and willingness to participate in similar learning activities in the future.

Procedure

Both classes first took the pretest. The experimental class then followed six meetings of metareflection-based PBL. In each meeting, students were oriented to a contextual problem, organised into groups, investigated the problem, presented their solutions, and evaluated the problem-solving process. In the final phase, students completed guided metareflection by answering prompts about the strategy they used, the errors they found, and the understanding they gained. An observer recorded the implementation of each phase. The control class studied the same material through conventional instruction consisting of teacher explanation, worked examples, and individual exercises. After the intervention, both classes took the posttest, and the experimental class completed the response questionnaire.

Data Analysis

Data were analysed using descriptive and inferential statistics. Before hypothesis testing, normality was tested using the Shapiro-Wilk test and homogeneity of variance using Levene's test, with a significance level of .05 (Field, 2018). The equivalence of initial ability was tested by comparing the pretest scores of the two classes with an independent samples t-test. The four effectiveness indicators were analysed as follows. First, a one-sample t-test was used to test whether the mean posttest score of the experimental class was greater than 75 (H0: $\mu \leq 75$; H1: $\mu > 75$). Second, classical mastery was calculated as the percentage of students scoring at least 75 and was compared with the 80% criterion. Third, learning implementation was calculated as the mean observation score and its percentage of the maximum score. Fourth, student responses were calculated as the mean questionnaire score and its percentage of the maximum score.

The improvement in critical thinking was analysed using the normalized gain, g = (posttest - pretest) / (100 - pretest), categorised as high ($g \geq 0.70$), medium ($0.30 \leq g < 0.70$), or low ($g < 0.30$) (Hake, 1998). A paired samples t-test was used to test the pretest-posttest difference within each class, and an independent samples t-test was used to compare the posttest scores and normalized gains of the two classes. Effect sizes were reported as Cohen's d and interpreted as

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

small (0.2), medium (0.5), or large (0.8) (Cohen, 1988). Because the control class posttest data deviated slightly from normality, the comparison of posttest scores was also checked with the Mann-Whitney U test (Field, 2018). The criteria for learning implementation and student responses are presented in Table 2.

Table 2. Criteria for Learning Implementation and Student Responses

Mean score (4-point scale)	Learning implementation	Student response
3.50 - 4.00	Very good	Very positive
2.50 - 3.49	Good	Positive
1.50 - 2.49	Fair	Less positive
1.00 - 1.49	Poor	Negative

RESULTS

Description of Students' Mathematical Critical Thinking

Table 3 summarises the pretest, posttest, and normalized gain of both classes. The pretest means of the two classes were almost identical, whereas the posttest mean of the experimental class was 7.17 points higher than that of the control class. The experimental class also showed less variation in posttest scores, indicating that students' achievement after metareflection-based PBL was more evenly distributed.

Table 3. Descriptive Statistics of Mathematical Critical Thinking

Class	Data	n	Mean	SD	Min	Max
Experimental	Pretest	30	58.27	8.42	40	72
	Posttest	30	81.60	6.04	68	90
	N-gain	30	0.54	0.18	0.18	0.81
Control	Pretest	30	57.53	7.52	44	71
	Posttest	30	74.43	7.78	61	85
	N-gain	30	0.38	0.22	-0.08	0.71

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Prerequisite Tests and Initial Equivalence

The results of the normality and homogeneity tests are shown in Table 4. All data in the experimental class were normally distributed ($p > .05$). In the control class, the pretest and normalized gain data were normal, while the posttest data deviated slightly from normality ($p = .036$). The variances of the two classes were homogeneous for all data ($p > .05$). The independent samples t-test on the pretest showed no significant difference between the classes, $t(58) = 0.36$, $p = .723$, $d = 0.09$. Hence, the two classes had equivalent initial critical thinking skills before the intervention.

Table 4. Results of Normality and Homogeneity Tests

Data	Shapiro-Wilk p (Experimental)	Shapiro-Wilk p (Control)	Levene p	Conclusion
Pretest	.604	.634	.660	Normal; homogeneous
Posttest	.226	.036	.219	Control not normal; homogeneous
N-gain	.248	.190	.246	Normal; homogeneous

Indicator 1: Mean Posttest Score Above the Minimum Mastery Criterion

The mean posttest score of the experimental class was 81.60 (SD = 6.04), with a 95% confidence interval of 79.34 to 83.86. The one-sample t-test showed that this mean was significantly greater than 75, $t(29) = 5.98$, $p < .001$, $d = 1.09$. The first indicator was therefore met. By contrast, the mean posttest score of the control class (74.43) did not exceed 75, $t(29) = -0.40$, $p = .654$.

Indicator 2: Classical Mastery

In the experimental class, 26 of 30 students (86.67%) reached the minimum mastery criterion of 75, exceeding the 80% criterion, so the second indicator was met. In the control class, only 18 of 30 students (60.00%) reached mastery. The four experimental students who did not reach mastery scored between 68 and 74, which means their scores were close to the criterion.

Indicator 3: Learning Implementation

The implementation of metareflection-based PBL across the six meetings is summarised in Table 5. The overall mean score was 3.03 out of 4 (109 out of 144, or 75.7%), which falls in the good category. Every aspect, including the metareflection phase, was in the good category. The

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

first meeting obtained the highest score (79.17%) because the orientation phase was rated very good, while the other meetings were consistently at 75%. The third indicator was therefore met.

Table 5. Implementation of Metareflection-Based PBL

Aspect observed	Total score (6 meetings)	Mean (1-4)	%	Category
Orienting students to the problem	19	3.17	79.17	Good
Organising students for learning	18	3.00	75.00	Good
Guiding individual and group investigation	18	3.00	75.00	Good
Developing and presenting results	18	3.00	75.00	Good
Analysing and evaluating the problem-solving process	18	3.00	75.00	Good
Metareflection: reviewing strategies, errors, and understanding	18	3.00	75.00	Good
Overall	**109 / 144**	**3.03**	**75.69**	**Good**

Indicator 4: Student Responses

Student responses to metareflection-based PBL are presented in Table 6. The overall mean response was 3.26 out of 4, or 81.42% of the maximum score, which is in the positive category. The individual mean response ranged from 2.60 to 3.80, and 27 of 30 students (90%) gave a response of at least 75% of the maximum score. The highest-rated statements concerned the role of metareflection in helping students recognise errors in their thinking (91.67%; agreed by all students) and the contextual nature of the problems (88.33%). The lowest-rated statements concerned checking answers and mathematical reasons (73.33%) and self-confidence in solving problems (74.17%), although both remained in the positive category. The fourth indicator was therefore met.

Table 6. Student Responses to Metareflection-Based PBL

No.	Statement aspect	Mean	SD	% of max.	% agree
1	The problems were interesting	3.13	0.68	78.33	83.3
2	The problems were contextual	3.53	0.68	88.33	90.0
3	Discussion helped understanding	3.13	0.78	78.33	83.3
4	Actively expressed opinions	3.10	0.88	77.50	80.0

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

5	Encouraged to try different strategies	3.43	0.63	85.83	93.3
6	Checked answers and reasons	2.93	0.78	73.33	73.3
7	Metareflection helped recognise errors	3.67	0.48	91.67	100.0
8	Metareflection helped understand strategies	3.23	0.90	80.83	76.7
9	Increased self-confidence	2.97	0.93	74.17	70.0
10	Willing to use the model again	3.43	0.50	85.83	100.0
	Overall	**3.26**	**0.26**	**81.42**	**Positive**

Comparison of the Experimental and Control Classes

As shown in Table 7, both classes improved significantly from pretest to posttest, but the improvement was larger in the experimental class (mean gain 23.33 points, d = 2.09) than in the control class (16.90 points, d = 1.53). The posttest mean of the experimental class was significantly higher than that of the control class, $t(58) = 3.99$, $p < .001$, with a mean difference of 7.17 points (95% CI 3.57 to 10.77) and a large effect size (d = 1.03). Because the control posttest was not perfectly normal, the result was confirmed with the Mann-Whitney U test, $U = 687.5$, $p < .001$. The mean normalized gain of the experimental class (0.54) was also significantly higher than that of the control class (0.38), $t(58) = 3.03$, $p = .004$, d = 0.78. In the experimental class, 6 students (20.0%) achieved a high gain, 19 (63.3%) a medium gain, and 5 (16.7%) a low gain, whereas in the control class the corresponding figures were 1 (3.3%), 18 (60.0%), and 11 (36.7%). The mean normalized gain of the experimental class was also significantly above the lower limit of the medium category, $t(29) = 7.12$, $p < .001$.

Table 7. Results of the t-Tests

Test	Comparison	t	df	p	d
Paired samples	Experimental: pretest vs posttest	11.42	29	< .001	2.09
Paired samples	Control: pretest vs posttest	8.38	29	< .001	1.53
Independent samples	Pretest: experimental vs control	0.36	58	.723	0.09
Independent samples	Posttest: experimental vs control	3.99	58	< .001	1.03
Independent samples	N-gain: experimental vs control	3.03	58	.004	0.78

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

One sample	Experimental posttest > 75	5.98	29	< .001	1.09

Table 8 summarises the four effectiveness indicators. All indicators were met, so metareflection-based PBL is effective for students' mathematical critical thinking skills.

Table 8. Summary of Effectiveness Indicators

Indicator	Criterion	Result	Decision
Mean posttest score	> 75	81.60; t(29) = 5.98, p < .001	Met
Classical mastery	≥ 80%	86.67% (26 of 30 students)	Met
Learning implementation	Good	3.03 (75.7%), good	Met
Student responses	Positive	3.26 (81.42%), positive	Met

DISCUSSION

The results show that metareflection-based PBL met all four effectiveness indicators: the mean critical thinking score exceeded the minimum mastery criterion, classical mastery exceeded 80%, the model was implemented in the good category, and students responded positively. Because both classes started with equivalent abilities, the difference in posttest scores and normalized gains can be attributed mainly to the treatment. The large effect on posttest scores (d = 1.03) is in line with meta-analytic evidence that PBL and related innovative models have strong effects on mathematical abilities and critical thinking (Sa'idah et al., 2026; Suparman et al., 2021) and with the finding that critical thinking is among the abilities most influenced by PBL at the secondary level (Ajinegara & Nuriadin, 2022).

The superiority of the experimental class can be explained by the way the model engages students in the components of critical thinking. Contextual problems require students to interpret information and analyse relationships before they can choose a strategy, which corresponds to the interpretation and analysis skills described by Facione (2015). Group investigation and presentation require students to justify and defend their solutions, while the evaluation phase asks them to judge the quality of alternative solutions (Arends, 2012; Hmelo-Silver, 2004). The student response data support this explanation: almost all students agreed that the problems were contextual and that they were encouraged to try different strategies, two conditions that PBL is designed to create (Yew & Goh, 2016).

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

The metareflection phase adds a component that the basic PBL syntax does not make explicit, namely conscious monitoring of one's own reasoning. All students agreed that metareflection helped them recognise errors in their thinking, which was the highest-rated statement in the questionnaire. Recognising and correcting one's own errors is the essence of the self-regulation skill of critical thinking (Facione, 2015) and of the regulation component of metacognition (Flavell, 1979). This finding is consistent with Marthaliakirana et al. (2022), who showed that PBL with metacognitive prompts produced greater gains in critical thinking than PBL without prompts, and with Syaiful et al. (2022), who reported that a metacognitive approach improved mathematical critical thinking. It also supports the view of Stanton et al. (2021) that metacognition must be fostered explicitly, and the cyclical relationship between metacognition and critical thinking proposed by Istiqlal et al. (2025). In other words, metareflection turns the experience of solving a problem into an object of thinking, so that students learn not only how to solve a problem but also how to examine the quality of their reasoning.

The distribution of normalized gains also deserves attention. The proportion of students with a low gain in the experimental class (16.7%) was less than half of that in the control class (36.7%), and the experimental class had a smaller posttest standard deviation. This suggests that metareflection helped students who would otherwise make little progress, perhaps because the reflection prompts made their misconceptions visible and gave them a chance to repair them. This pattern resembles the finding that metacognitive approaches improve critical thinking across low, medium, and high prior-ability groups (Syaiful et al., 2022).

Several results point to aspects that need strengthening. The statements with the lowest ratings concerned checking answers and mathematical reasons and self-confidence in solving problems. This is consistent with Kholid et al. (2020), who found that students seldom re-examine their methods once an answer has been obtained. It suggests that habits of verification and confidence in mathematical reasoning develop more slowly than engagement and need more time than six meetings. Learning implementation was in the good category, but five of the six meetings were at the lower boundary of that category, so the quality of implementation, particularly guiding investigation and metareflection, can still be improved, for example by using more specific reflection prompts or reflection journals (Faizin, 2018).

This study has limitations. First, it used intact classes without random assignment, so although the pretest showed equivalence, other class-level factors cannot be fully ruled out. Second, the sample was small and came from MA Arifah Gowa, which limits generalisation. Third,

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

classical mastery of 86.67% exceeded the 80% criterion descriptively, but a one-sample proportion test did not show that the population proportion was significantly above 80% ($z = 0.91$, $p = .181$); a larger sample is needed to confirm this indicator inferentially. Fourth, the intervention lasted only six meetings, so the durability of the effect is unknown. Future studies could use larger and randomly assigned samples, a longer intervention, delayed posttests, and analysis of each critical thinking indicator separately.

CONCLUSION

Metareflection-based problem-based learning is effective in developing students' mathematical critical thinking skills. All four effectiveness criteria were met: the mean posttest score of the experimental class (81.60) was significantly higher than the minimum mastery criterion of 75, classical mastery reached 86.67%, learning implementation was in the good category (75.7%), and student responses were positive (81.42%). Students taught with metareflection-based PBL also achieved higher posttest scores and normalized gains than students taught with conventional instruction, with a large effect size. The main contribution of this study is to show that adding an explicit metareflection phase, in which students review their strategies, errors, and understanding, strengthens the effect of PBL on critical thinking. Mathematics teachers are encouraged to include structured reflection prompts at the end of PBL lessons and to give particular attention to training students to verify their answers and build confidence in their mathematical reasoning.

REFERENCES

Ajinegara, M. W., & Nuriadin, I. (2022). Meta-analysis study of problem-based learning models on student's mathematic ability at junior high school and high school levels. *Journal of Medives: Journal of Mathematics Education IKIP Veteran Semarang, 6*(1), 203–210. https://doi.org/10.31331/medivesveteran.v6i1.1776

Arends, R. I. (2012). *Learning to teach* (9th ed.). McGraw-Hill.

Bron, J. F., & Prudente, M. S. (2024). Examining the effect of problem-based learning approach on learners' mathematical creativity: A meta-analysis. *International Journal of Research in Education and Science, 10*(3), 653–668. https://doi.org/10.46328/ijres.3456

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Ennis, R. H. (2018). Critical thinking across the curriculum: A vision. *Topoi, 37*(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4

Facione, P. A. (2015). *Critical thinking: What it is and why it counts*. Insight Assessment.

Faizin, K. (2018). Pemanfaatan jurnal refleksi sebagai strategi metakognitif dalam meningkatkan keaktifan dan hasil belajar matematika. *Lentera Pendidikan: Jurnal Ilmu Tarbiyah dan Keguruan, 21*(1), 33–47. https://doi.org/10.24252/lp.2018v21n1i4

Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Publications.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.

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

Hasanah, N. R., Turmudi, T., & Juandi, D. (2024). Students' reflective thinking skills in mathematics learning: Systematic literature review. *Journal of Didactic Studies, 2*(1), 7–16. https://doi.org/10.17509/jds.v2i1.71718

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235–266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Istiqlal, D. G., Susilo, B. E., Mariani, S., & Rosyida, I. (2025). Meta-synthesis of metacognition's role in enhancing mathematical critical thinking. *EduMa: Mathematics Education Learning and Teaching, 14*(2), 483–495. https://doi.org/10.24235/eduma.v14i2.24753

Kholid, M. N., Sa'dijah, C., Hidayanto, E., & Permadi, H. (2020). How are students' reflective thinking for problem solving? *Journal for the Education of Gifted Young Scientists, 8*(3), 1135–1146. https://doi.org/10.17478/jegys.688210

Liu, Y., & Pásztor, A. (2022). Effects of problem-based learning instructional intervention on critical thinking in higher education: A meta-analysis. *Thinking Skills and Creativity, 45*, 101069. https://doi.org/10.1016/j.tsc.2022.101069

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Marthaliakirana, A. D., Suwono, H., Saefi, M., & Gofur, A. (2022). Problem-based learning with metacognitive prompts for enhancing argumentation and critical thinking of secondary school students. *Eurasia Journal of Mathematics, Science and Technology Education, 18*(9), em2148. https://doi.org/10.29333/ejmste/12304

Merkebu, J., Mennin, S., Durning, S. J., & Samuel, A. (2025). Cultivating the art of metacognitive reflection in medical education. *The Clinical Teacher, 22*(2), e70043. https://doi.org/10.1111/tct.70043

OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Sa'idah, N., Jailani, Sudiyatno, & Retnawati, H. (2026). Innovative teaching models and critical thinking in mathematics and statistics education: A meta-analytical review. *International Journal of Research in Education and Science, 12*(1), 140–157. https://doi.org/10.46328/ijres.5054

Stanton, J. D., Sebesta, A. J., & Dunlosky, J. (2021). Fostering metacognition to support student learning and performance. *CBE—Life Sciences Education, 20*(2), fe3. https://doi.org/10.1187/cbe.20-12-0289

Su, H. F. H., Ricci, F. A., & Mnatsakanian, M. (2016). Mathematical teaching strategies: Pathways to critical thinking and metacognition. *International Journal of Research in Education and Science, 2*(1), 190–200. https://ijres.net/index.php/ijres/article/view/100

Suparman, S., Tamur, M., Yunita, Y., Wijaya, T. T., & Syaharuddin, S. (2021). Using problem-based learning to enhance mathematical abilities of primary school students: A systematic review and meta-analysis. *JTAM (Jurnal Teori dan Aplikasi Matematika), 5*(1), 144–161. https://doi.org/10.31764/jtam.v5i1.3806

Syaiful, Huda, N., Mukminin, A., & Kamid. (2022). Using a metacognitive learning approach to enhance students' critical thinking skills through mathematics education. *SN Social Sciences, 2*, Article 31. https://doi.org/10.1007/s43545-022-00325-8

Yew, E. H. J., & Goh, K. (2016). Problem-based learning: An overview of its process and impact on learning. *Health Professions Education, 2*(2), 75–79. https://doi.org/10.1016/j.hpe.2016.01.004

Yu, L., & Zin, Z. M. (2023). The critical thinking-oriented adaptations of problem-based learning models: A systematic review. *Frontiers in Education, 8*, 1139987. https://doi.org/10.3389/feduc.2023.1139987

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Zetriuslita, Z., & Ariawan, R. (2021). Students' mathematical thinking skill viewed from curiosity through problem-based learning model on integral calculus. *Infinity Journal, 10*(1), 31–40. https://doi.org/10.22460/infinity.v10i1.p31-40

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License