International Journal of Research in Education and Science

www.ijres.net

Innovative Teaching Models and Critical Thinking in Mathematics and Statistics Education: A Meta-Analytical Review

Nusrotus Sa'idah [1] *, Jailani [2], Sudiyatno [3], Heri Retnawati [4]

[1] Universitas Negeri Yogyakarta, Catur Tunggal Depok Sleman Yogyakarta, Indonesia, 0000-0002-2234-8476
[2] Universitas Negeri Yogyakarta, Catur Tunggal Depok Sleman Yogyakarta, Indonesia, 0000-0002-4162-7235
[3] Universitas Negeri Yogyakarta, Catur Tunggal Depok Sleman Yogyakarta, Indonesia, 0000-0003-3953-9356
[4] Universitas Negeri Yogyakarta, Catur Tunggal Depok Sleman Yogyakarta, Indonesia, 0000-0002-1792-5873
* Corresponding author: Nusrotus Sa'idah (nusrotussaidah.2022@student.uny.ac.id)

Article Info

Article History

Received:
11 June 2025

Revised:
26 October 2025

Accepted:
4 December 2025

Published:
1 January 2026

Keywords

Meta-analysis
Critical thinking
Mathematics Education
Statistics education
Publication bias

Abstract

This meta-analysis investigates the impact of innovative learning models—such as problem-based learning, blended learning, and metaverse-based instruction—on students' critical thinking skills in mathematics and statistics education. Based on data from 19 studies, the analysis revealed a large overall effect (Hedges' g = 1.03) in favor of innovative methods compared to conventional approaches. Substantial heterogeneity (I^2 = 91.37%) indicates potential moderating factors influencing the outcomes. Results from publication bias analyses—including funnel plots, Egger's test, Trim and Fill, and Rosenthal's Fail-Safe N (N = 1559) showed no significant bias, supporting the validity of the findings. These results underscore the transformative potential of innovative teaching practices in mathematics and statistics classrooms. However, the limited representation of statistical critical thinking studies suggests a need for further research across diverse contexts and long-term implementations. Overall, the findings highlight the importance of adopting student-centered learning strategies to meet the demands of 21st-century education.

Citation: 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 (IJRES), 12*(1), 140-157. https://doi.org/10.46328/ijres.5054

Introduction

Critical thinking is widely recognized as one of the core competencies required for success in the 21st century, enabling individuals to navigate the complexities of modern life (Voogt & Roblin, 2012). It has been shown to be a strong predictor of an individual's future success (Haynes et al., 2016). In the field of education, particularly in mathematics and statistics, critical thinking equips students with the ability to analyze, evaluate, and solve problems logically and systematically (Çelik & Özdemir, 2020; Romero Ariza et al., 2024). This cognitive skill also demands that students articulate their reasoning processes and justify their answers, highlighting the centrality of argumentation in mathematics learning (Dogruer & Akyuz, 2020). Reynders et al (2020) define critical thinking through components such as analyzing, synthesizing, formulating arguments, and evaluating. In a similar vein, Cortázar et al. (2021) propose six key indicators: interpretation, analysis, inference, argument, evaluation, and metacognition. The development of these skills enhances students' mathematical reasoning and problem-solving abilities, fostering curiosity, perseverance, and self-confidence (Apriliana et al., 2019; Gaspersz & Salamor, 2021).

However, students often face challenges in understanding mathematical concepts (Marito & Riani, 2022; Yusriani et al., 2020), with 40% of college graduates still lacking critical thinking skills (Belkin, 2015). This is evident in the statistics learning outcomes of college students, which average at 70 (Marito & Riani, 2022), likely due to predominantly teacher-centered instructional methods. This issue is not exclusive to higher education; approximately 9% of elementary school students in Indonesia also exhibit underdeveloped critical thinking skills (Lestari et al., 2021). Moreover, numerous studies have shown that traditional teaching methods are often less effective in optimally enhancing students' mathematics learning outcomes (Yusriani et al., 2020).

One proven approach to improving mathematics learning outcomes is cooperative learning models. These models create opportunities for students to collaborate in groups, share ideas, and assist one another in solving problems. Various techniques within cooperative learning, such as Problem-Based Learning (Apriliana et al., 2019; Happy & Widjajanti, 2014; Yusriani et al., 2020), metaverse technology applications Rachmadtullah et al., 2023)), and Group Investigation (Gaspersz & Salamor, 2021), are designed to actively engage students in the learning process.

In Indonesia, the application of cooperative learning models has been increasingly adopted across educational levels, from elementary schools to higher education institutions. These levels have unique characteristics, as students are expected not only to master theoretical knowledge but also practical skills relevant to the workforce. Despite the abundance of individual studies on cooperative learning models, their results often vary due to differences in context, methodology, and sample size. A meta-analysis approach is therefore needed to integrate findings from various studies to obtain a clearer and more comprehensive picture of the effectiveness of innovative learning models on mathematics learning outcomes (Borenstein, 2023). The dataset in this study consists of journal publications categorized based on measured variables, namely critical thinking skills in mathematics and statistics, from elementary to higher education levels, with experimental groups employing innovative teaching methods and control groups using conventional instruction. A total of 19 studies were identified, revealing differences in critical thinking skills for mathematics and statistics compared to conventional learning methods.

The numerical data utilized include standard deviation and sample size from both experimental and control groups. This information serves as the initial basis for calculating the effect sizes of each study.

There is still limited research using a meta-analysis approach to identify publication bias. Several studies, such as those conducted by (Ridwan et al., 2022), indicate no publication bias in examining the effectiveness of cooperative learning in improving mathematics learning outcomes for vocational school students. This study used the trim and fill method to detect publication bias. Furthermore, research by (A. D. Putri et al., 2024) also employed a random-effects model with the trim and fill method, demonstrating no publication bias in identifying the impact of RME on mathematical literacy in the Indonesian context. Another relevant study by (Nurhayati et al., 2023) identified the impact of innovative learning models on improving student achievement in Indonesia, also using the trim and fill method to detect publication bias. Meta-analysis also allows researchers to evaluate potential publication bias and heterogeneity across studies. Analyses such as funnel plots, Egger's test, and the Fail-Safe N method are utilized to ensure the validity of the results obtained (Ruppar, 2020). Thus, meta-analysis is a crucial tool for providing evidence-based recommendations to improve mathematics and statistics education (Retnawati et al., 2018).

This study aims to identify the presence of publication bias in the effectiveness of innovative learning models on mathematics and statistics learning outcomes for students ranging from elementary to higher education. The calculation process in this meta-analysis includes several stages: calculating the effect size, conducting heterogeneity tests, analyzing results using a forest plot, and identifying publication bias through funnel plot analysis using the trim and fill method. The findings of this study are expected to provide insights for educators and policymakers in optimizing mathematics education through innovative teaching methods.

Method

Research Strategy

This study is a quantitative research utilizing a meta-analysis approach. Meta-analysis is a quantitative method used to systematically combine or evaluate other research studies to draw conclusions (Retnawati et al., 2018), It serves as a flexible tool for various purposes, not only combining evidence on intervention effectiveness but also identifying trends or research gaps (Borenstein, 2009). The stages of research using a meta-analysis approach include: (1) formulating the problem, (2) conducting a literature search, (3) collecting information, (4) evaluating the quality of studies, and (5) analyzing and interpreting study results (Cooper, 2017). According to (Retnawati et al., 2018), the stages involve determining the research questions, identifying relevant studies, tracking and collecting studies, coding data, and calculating effect size. The meta-analysis in this study utilized data from research studies related to the application of innovative learning methods to enhance students' critical thinking skills, ranging from elementary to higher education levels.

Collecting Data

The research data were obtained from several literature reviews by collecting and analyzing numerical information that included measurable variables. The criteria for studies identified in this research to be included in the meta-analysis are as follows: (1) Studies with the dependent variable of critical thinking skills in either statistics or mathematics, (2) Studies involving two distinct groups, namely an experimental group and a control group, (3) Data coding based on sample size, mean, and standard deviation for both groups, (4) The independent variable in the studies being the implementation of innovative learning models, (5) Educational levels covering elementary, secondary, and higher education, (6) Publications ranging from 2017 to 2024.The studies were retrieved from online databases such as ERIC, Scopus, and Google Scholar (see Figure 1). Keywords used included {statistics}, {critical thinking in statistics}, {mathematical critical thinking}, {effectiveness of learning models}. The selection of data adheres to systematic review and meta-analysis standards (PRISMA), which are reported comprehensively and consistently, aiding in quality assessment to make evidence-based decisions (Pigott & Polanin, 2020).

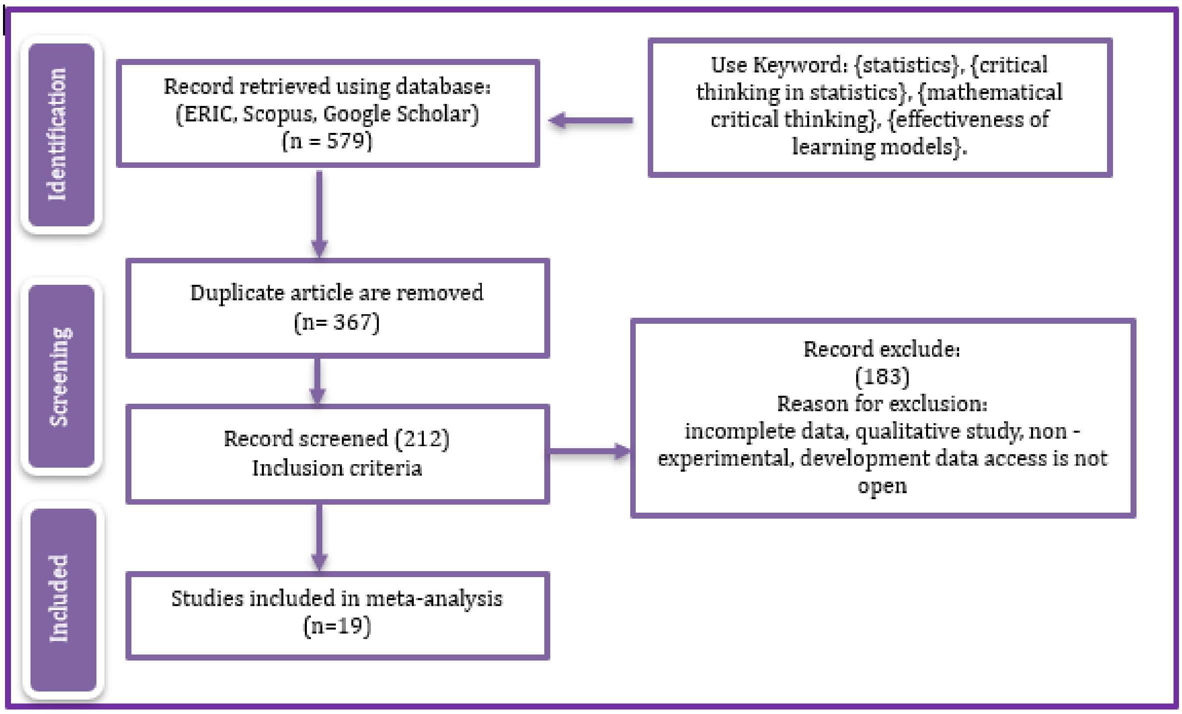

Figure 1. Flowchart Illustrating the Use of PRISMA to Conduct a Meta-Analysis Investigating Critical Thinking Success in Statistics and Mathematics

Data Analyse

The data analysis in this meta-analysis research aims to determine the effect size (Schmidt et al., 2009) to assess the effectiveness of innovative learning models on mathematical critical thinking skills. The data analysis process was conducted using the JASP software. The effect size was calculated using Hedges' g (Borenstein, 2009). which measures the difference between the control and experimental groups (D'Angelo et al., 2013). Effect sizes were classified according to (Cohen, 1988) as large ($g \geq 0.80$), moderate ($0.50 \leq g \leq 0.79$), small ($0.20 \leq g \leq 0.49$), and negligible ($g < 0.19$).

Additionally, the trim and fill method (Duval & Tweedie, 2000b) was used to analyze publication bias, which examines the relationship between the probability of publication and the statistical significance of study results, potentially leading to an asymmetric funnel plot. This method also estimates missing studies due to extreme funnel plot points in meta-analysis (Duval & Tweedie, 2000a). However, publication bias is only one of the possible causes of an asymmetric funnel plot Sterne & Harbord, 2004). Funnel plots and Fail-safe N (FSN) statistics were employed to assess the potential for bias. If the funnel plot is symmetric, FSN statistics are used. An FSN value greater than 5K + 10 (K = the number of individual studies) indicates no publication bias in the meta-analysis (Fragkos et al., 2014).

Heterogeneity measurement was further conducted using Q and I^2 statistics to evaluate the extent of variation among study results. An I^2 value above 50% indicates significant heterogeneity (Borenstein, 2023), meaning that the differences among study results are greater than what would be expected by chance. This test uses a 95% confidence interval and a p-value < 0.05, which is considered statistically significant. Several factors influence heterogeneity in meta-analysis: (1) the number of studies included in the meta-analysis, where a larger number of studies results in more accurate heterogeneity estimation, (2) between-study variance in effect sizes, and (3) within-study variance in observed effect sizes(Ruppar, 2020).

Results

Data Encoding

This data includes a collection of studies evaluating the effectiveness of various innovative teaching methods on critical thinking skills, specifically mathematical and statistical critical thinking, comprising a total of 19 studies. The data contains information about experimental and control groups, including the number of participants (N), mean values, and standard deviations (SD) for each group. Table 1 summarizes the findings from the data coding process.

Table 1. Result of Data Encoding

No.	Name of Study	Experiment Group			Control Group			Dependent	Independent	Grade
		N	Mean	*SD*	*N*	Mean	*SD*			
1	(Gaspersz & Salamor, 2021)	25	13.2	28,13	25	9,28	4,077	Statistics Critical Thinking	group investigation	Collage
2	(Marito & Riani, 2022)	25	76.125	7,37	25	58,5	11,36	Statistics Critical Thinking	Blended Learning	Collage
3	(Arisoy & Aybek, 2021)	31	37.03	5,06	31	29,00	3,75	Mathematical Critical Thinking	Subject-Based Critical Thinking	Elementary
4	(Bates et al., 2024)	52	0.75	3,48	11	0,80	2,51	Statistics Critical Thinking	Integrated Cogency Method/ICM	Collage

No.	Name of Study	Experiment Group			Control Group			Dependent	Independent	Grade
		N	Mean	*SD*	*N*	Mean	*SD*			
5	(Zetriuslita et al., 2017)	25	0.87	0,2	26	0,55	0,22	Mathematical Critical Thinking	Problem Based Learning	Collage
6	(Idris & Khaulah, 2020)	32	36.62	10,32	32	34,21	9,32	Mathematical Critical Thinking	AMORA Learning	Secondary
7	(Yusriani et al., 2020)	30	0.5873	0,092	30	0,416	0,136	Mathematical Critical Thinking	Problem Based Learning	Elementary
8	(Helma Lia Sapitri et al., 2024)	24	84.63	7,728	25	72,42	6,198	Mathematical Critical Thinking	Contextual Teaching	Secondary
9	(Suryawan et al., 2023)	52	64.97	9,72	48	54,43	8,17	Mathematical Critical Thinking	Problem-Based Multimodal Approach	Secondary
10	(Ashidiq et al., 2024)	30	80.4	8,43	30	76,38	9,07	Mathematical Critical Thinking	Hand-Made Projector Activity	Secondary
11	(Lestari et al., 2021)	23	79.65	14,28	25	68,28	16,2	Mathematical Critical Thinking	Mathematics Comic through Contextual Teaching	Elementary
12	(Apriliana et al., 2019)	35	10.4	5,25	35	7,11	3,7	Mathematical Critical Thinking	Problem Centered Learning	Secondary
13	(Rachmadtullah et al., 2023)	31	85.77	7,112	31	79,52	7,04	Mathematical Critical Thinking	Metaverse technology applications	Elementary
14	(Zulkarnain et al., 2023)	32	34.84	7,02	31	29,84	7,26	Mathematical Critical Thinking	Problem Based Learning	Secondary
15	(S. A. Putri et al., 2019)	32	66.06	14,73	31	51,13	15,8	Mathematical Critical Thinking	Problem Based Learning	Secondary
16	(Sari et al., 2019)	33	12.97	1,468	33	12,61	2,061	Mathematical Critical Thinking	Problem Based Learning	Secondary

No.	Name of Study	Experiment Group			Control Group			Dependent	Independent	Grade
		N	Mean	*SD*	*N*	Mean	*SD*			
17	(Steven et al., 2019)	34	59.56	19,95	33	51,52	14,91	Mathematical Critical Thinking	Problem Based Learning	Secondary
18	(Anwar & Setyaningrum, 2021)	34	79.2	9,61	34	72,9	9,76	Mathematical Critical Thinking	Blended Learning	Secondary
19	(Setiyawan et al., 2024)	34	88.32	4,656	34	68,59	3,791	Mathematical Critical Thinking	flipped classroom models	Tertiary

Table 1. presents studies encompassing various levels of education, ranging from elementary to higher education. For the independent variable, various teaching methods were employed as interventions to influence mathematical and statistical critical thinking, such as group investigation, blended learning, problem-based learning, contextual teaching, and metaverse technology. The calculation of effect size was conducted using the standardized mean difference method, dividing the difference between the experimental and control groups by the standard deviation (Ridwan et al., 2021).

Effect Size Calculation

Effect size is a statistical measure used to indicate the magnitude of the impact or effect of an intervention or independent variable on the dependent variable in a study (Borenstein, 2009; Retnawati et al., 2018). This measure helps explain the extent of the impact of an intervention, independent of the study's sample size. Unlike the p-value (statistical significance), which only indicates the presence or absence of an effect, effect size provides a quantitative representation of the magnitude of the effect. The effect size calculations are presented in Table 2.

Table 2. Results of Effect Size Calculation

No	Researcher and Year of Research	g	Seg	Classification
1	(Gaspersz & Salamor, 2021)	0.192	0.281	Less strong
2	(Marito & Riani, 2022)	1.812	0.335	Very strong
3	(Arisoy & Aybek, 2021)	1.780	0.299	Very strong
4	(Bates et al., 2024)	-0.015	0.330	No effect
5	(Zetriuslita et al., 2017)	1.497	0.316	Very strong
6	(Idris & Khaulah, 2020)	0.242	0.249	Less strong
7	(Yusriani et al., 2020)	1.456	0.289	Very strong
8	(Helma Lia Sapitri et al., 2024)	1.719	0.333	Very strong
9	(Suryawan et al., 2023)	1.161	0.216	Very strong

No	Researcher and Year of Research	g	Seg	Classification
10	(Ashidiq et al., 2024)	0.453	0.260	Strong
11	(Lestari et al., 2021)	0.730	0.296	Strong
12	(Apriliana et al., 2019)	0.716	0.245	Strong
13	(Rachmadtullah et al., 2023)	0.872	0.264	Very Strong
14	(Zulkarnain et al., 2023)	0.692	0.258	Strong
15	(S. A. Putri et al., 2019)	0.966	0.256	Very Strong
16	(Sari et al., 2019)	0.199	0.245	less strong
17	(Steven et al., 2019)	0.450	0.246	less strong
18	(Anwar & Setyaningrum, 2021)	0.643	0.247	Strong
19	(Setiyawan et al., 2024)	4.594	0.464	Very Strong

Table 2. presents the effect size (g) and associated standard error (Seg) from various studies. The calculation of effect size (g) was performed by dividing the mean difference between the two groups by the pooled standard deviation of both groups. According to (Hedges, 1981) the estimation of sample mean differences tends to yield more significant values compared to the actual population parameter. This indicates the presence of bias in the effect size estimation, particularly when using measures like Cohen's d, where d = $\frac{\overline{x_1}-\overline{x_2}}{S_{within}}$, This bias can lead to overestimation, meaning the calculated d values often exaggerate the actual effect size in the population. To minimize this bias, a conversion of d to a more accurate effect size, that is, the value of g, is required using a correction factor, J. The J correction is a mathematical factor that helps eliminate bias caused by small sample sizes (Ridwan et al., 2021). It is calculated using the formula 1- $\frac{3}{4df-1}$ di mana df = n_1+n_2-2. By multiplying d by the J correction factor, a corrected g value is obtained, ensuring more accurate and reliable estimates that represent the population parameters. A positive (g) value indicates a beneficial effect of the intervention, while a negative (g) value indicates a detrimental effect or no effect at all. The classification of effect size values according to Cohen includes 21% categorized as less strong, 26.3% as strong, and 47.4% as very strong. Only one study reported no effect on the improvement of critical thinking skills, as indicated by the negative value (-0.015).

Heterogeneity Test, Q and I^2 Statistics

The researchers assumed that the distribution of studies in this meta-analysis indicates heterogeneity; therefore, heterogeneity was tested using Q and I^2 statistics (Cooper et al., 2009).

Table 3. Heterogenity Test

Q Value	Df (Q)	P	I^2
132.358	18	p < .001	91.370

Table 3. presents the statistical test results from the meta-analysis of 19 studies, where the Q value is 22.797 with

a degree of freedom (df) of 1 and a p-value < .001, indicating that the model coefficients are statistically significant and contribute to the overall model. This provides strong evidence to reject the null hypothesis that all coefficients are zero. The residual heterogeneity value (Q = 132.358; df = 18; p < .001) indicates significant heterogeneity. Therefore, the Random Effects model is more appropriate for estimating the average effect size from the 19 studies analyzed. The results also suggest potential for exploring moderator variables that influence the application of innovative teaching methods to enhance mathematical and statistical critical thinking. Unlike Q statistics, I^2 statistics are not affected by the number of studies. The interpretation criteria for I^2 are as follows: 25% indicates low heterogeneity, 50% indicates moderate heterogeneity, and 75% indicates high heterogeneity (Cooper et al., 2009). The heterogeneity level in this meta-analysis is very high, with an I^2 value of 91.37% (see Table 3).

Table 4. Summary Effect/Mean Effect Size

Coefficients						
					95% Confidence Interval	
	Estimate	**Standard Error**	**z**	**p**	**Lower**	**Upper**
intercept	1.030	0.216	4.775	< .001	0.607	1.453

Note. Wald test.

The analysis using the Random Effects model indicates a significant impact of innovative learning models on improving mathematical and statistical critical thinking (z = 4.775; p < .001; 95% CI [0.607; 1.453]). The overall impact of innovative teaching models is very high (random effect = 1.030). According to Cohen's criteria (1988), r=0.1 (low), r=0.3 (moderate), and r=0.5 (high). The analysis results, as shown in Figure 2, reveal that the effect sizes of the analyzed studies vary greatly, ranging from -0.01 to 1.78.

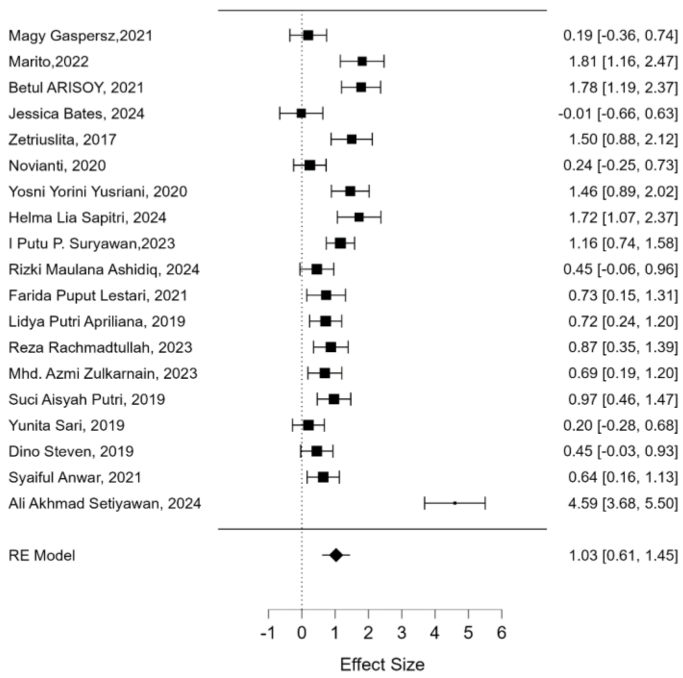

Figure 2. Forest Plot Summary Effect Model Random Effects

The size of the black boxes in the forest plot represents the relative weight of each study in the combined effect size. Studies with larger sample sizes or smaller variance have larger black boxes, indicating a greater contribution to the overall results. Conversely, studies with smaller weights have smaller boxes, indicating a lower influence on the combined effect size.

Biased Publication Analysis

To assess the presence of publication bias in a meta-analysis study on the application of learning models to enhance mathematical and statistical critical thinking, a funnel plot diagram was used. This diagram helps examine the reliability and validity of publication bias by plotting standard error values against the effect sizes of the included studies (Riley et al., 2021). Three methods are utilized to detect publication bias: (1) Funnel Plot Trim and Fill method, (2) Egger's Test, and (3) Fail-Safe N. It is challenging to justify whether the plot is symmetrical due to the irregular distribution of points, requiring statistical approaches to test asymmetry in the plot.

The funnel plot represents the distribution of effect sizes from the analyzed studies against standard error. The standard error increases as it moves downward and decreases as it moves upward. A funnel plot is considered symmetrical when it does not indicate publication bias, with data points typically distributed symmetrically around a central vertical line representing the average effect size. However, in this plot, most points are concentrated on the left side with fewer points on the right, indicating a potential imbalance or bias.

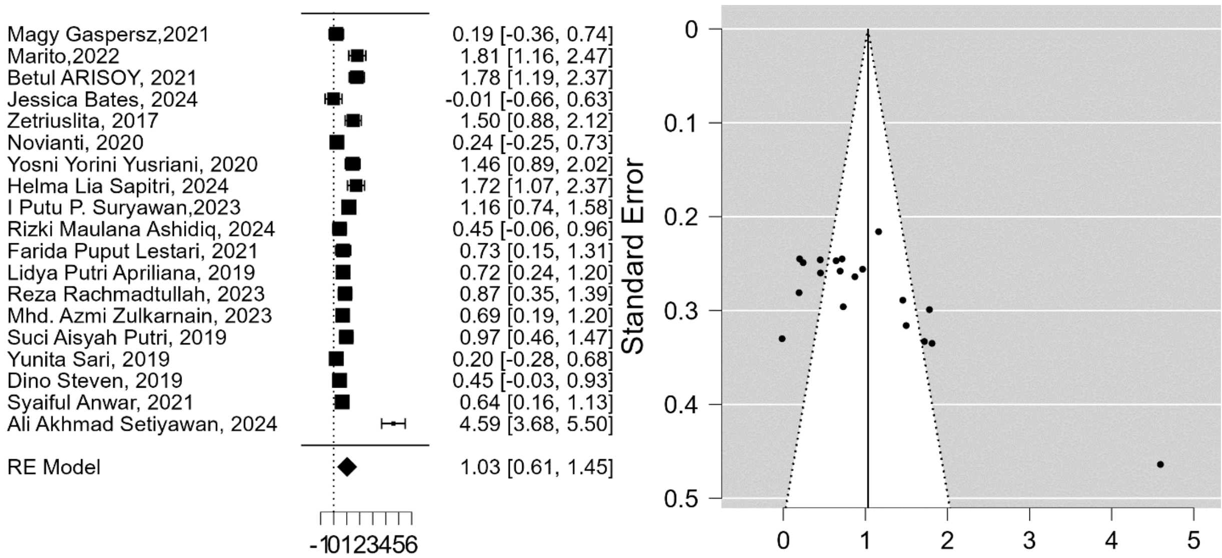

Figure 3. Funnel Plot Funnel Plot Trim and Fill Method

A meta-analysis study is considered robust if it is free from publication bias. Publication bias occurs when not all studies on a particular subject are included in the analysis ((Kutahya Dumlupinar University, Turkey & Turkuresin, 2021). The funnel plot diagram for the meta-analysis is presented in Figure 3. The X-axis in Figure 3 represents the effect size, specifically Hedges's g effect, from the included studies on the application of innovative learning models to enhance critical thinking in mathematics and statistics. The Y-axis represents the standard error of the studies included in the meta-analysis. Larger studies appear at the top and cluster around the main effect size, while smaller studies are positioned at the bottom with a wider spread due to greater sample error variation.

Publication bias can be detected using the Trim and Fill method. This method employs an iterative procedure to remove the most extreme small studies on the positive side of the funnel plot and recalculates the effect size at each iteration until the funnel plot achieves symmetry (Nurhayati et al., 2023; Ridwan et al., 2021).The results of the funnel plot in Figure 3 show closed circles forming a symmetrical pattern. This indicates no publication bias, meaning no studies are missing or unpublished. There is a difference in the effectiveness of innovative and conventional learning models on mathematical and statistical critical thinking, free from potential publication bias. Additionally, the symmetry of the funnel plot can be further analyzed using statistical methods such as Egger's Test.

Table 5. Result of Egger Test

Regression test for Funnel plot asymmetry ("Egger's test")		
	z	**p**
sei	5.000	< .001

Table 5 presents the results of Egger's Test for detecting publication bias. Egger's Test uses a criterion of p>0.05, which confirms that the funnel plot is symmetrical. However, the results of Egger's Test in Table 5 show a p<0.001, indicating that the funnel plot is asymmetrical, meaning there is publication bias in this study. Another method to assess the reliability and validity of publication bias is by examining the Fail-Safe N value. Fail-Safe N is used to detect the file-drawer effect, a term referring to studies that remain unpublished because their results are deemed insignificant.

Table 6. Rosenthal's Fail-Safe N Results

	Fail-safe N	**Target Significance**	**Observed Significance**
Rosenthal	1.559.000	0.050	< .001

Table 6 shows that the Fail-Safe N value is 1559. There is a tentative hypothesis that approximately 1559 studies with biased results have not been published. The Fail-Safe N value is then compared to the value of 5K+105, where K=19, resulting in 5(19) + 10=1055. Since the Fail-Safe N value is 1559 with a target significance level of 0.05 and p<0.001, this indicates that the Fail-Safe N value is greater than 5K+10. The findings from this meta-analysis investigation suggest there is no publication bias. Combining funnel plot interpretations with various statistical tests demonstrates that each method for detecting publication bias has its own strengths and limitations. For example, asymmetry tests like Egger's test can identify early indications of publication bias but are sensitive to small sample sizes and variability between studies (Egger et al., 1997). On the other hand, Fail-Safe N provides an estimate of how many additional studies would be required to render the results nonsignificant but does not account for variations in effect sizes across studies (Rosenthal, 1979). By using a combination of these methods, researchers gain a more comprehensive and accurate understanding of the presence of publication bias in meta-analyses.

Discussion

Critical thinking skills are essential to navigate the complexities of changes in the modern world. Education plays a vital role in the 21st century, including in mathematics and statistics, where critical thinking is crucial for helping students analyze, evaluate, and interpret information to draw logical conclusions. Critical thinking in learning can be supported by the implementation of innovative teaching methods. Innovative learning models such as problem-based learning (Bron & Prudente, 2024; Suryawan et al., 2023; Zetriuslita et al., 2017), flipped classrooms, blended learning (Anwar & Setyaningrum, 2021), and the use of technology such as metaverse applications (Rachmadtullah et al., 2023) have been proven to actively engage students and support the development of critical thinking skills.

Many studies have discussed the meta-analysis of innovative learning models applied to mathematics and statistics education; however, their results often vary due to differences in research context, methodology, or the focus of discussions. For example, a meta-analysis study on the effect of problem-based learning approaches on mathematical creativity (Bron & Prudente, 2024; Katz & Stupel, 2015) reported varying results. Similarly, studies by (Kutahya Dumlupinar University, Turkey & Turkuresin, 2021) identified the impact of cooperative learning (Ridwan et al., 2022) and innovative learning (Nurhayati et al., 2023) in improving mathematics achievement in Indonesia.

This study utilizes data from 19 studies that include mean values, standard deviations, and sample sizes from both experimental and control groups. The results encompass effect size values, heterogeneity tests, summary effect values, forest plots, and publication bias analysis using funnel plots with Fail-Safe N and trim and fill methods. The calculations, assisted by JASP software, revealed an effect size of 1.03, with a lower limit of 0.61 and an upper limit of 1.45. This indicates that mathematical and statistical critical thinking skills influenced by innovative teaching improved by 103% compared to conventional teaching. For the dependent variable of statistical critical thinking, only two studies were included: (Gaspersz & Salamor, 2021) which reported an effect size of 0.192 (less strong), indicating a weak impact of Group Investigation assisted by SPSS in statistics learning; and (Marito & Riani, 2022) which reported an effect size of 1.812 (very strong), indicating a significant impact of blended learning in statistics education. The limited research on statistical critical thinking underscores the need for further studies.

These findings align with those of (Ridwan et al., 2021), which showed that problem-solving abilities in mathematics for middle school students improved by nearly 100% (effect size = 95%) with innovative teaching methods. Similarly, research by (A. D. Putri et al., 2024) identified the use of Realistic Mathematics Education (RME) in a meta-analysis, yielding an effect size of 1.03, indicating a 103% significant impact of RME on mathematics education from primary to high school. The heterogeneity analysis, as shown by the I^2 statistic (Table 3), indicated a value of 91.37%, and the residual heterogeneity was significant ($Q = 132.358$; $df = 18$; $p < .001$). This suggests variability among the analyzed studies, influenced by factors such as study period, education level, sample size, and the use of innovative teaching methods. Therefore, conducting moderator analyses in meta-analysis is crucial (Higgins & Thompson, 2002).

Further analysis detected publication bias in the included studies. Using the funnel plot and Egger's test, the results indicated $p<0.001$, suggesting that the funnel plot was asymmetrical. However, regression asymmetry analysis using Egger's test showed that the asymmetry was not statistically significant (Permatasari et al., 2024). Comparing these results with the Trim and Fill method, the funnel plot in Figure 2 displayed closed circles forming symmetry, indicating that no studies were missing or unpublished. The Trim and Fill method is used to evaluate and correct potential publication bias in meta-analysis by theoretically imputing studies into the funnel plot. If the changes in effect size are minimal, it indicates that the meta-analysis results remain valid even if there are potential unpublished studies (Retnawati et al., 2018).Publication bias detection can also be conducted using the Fail-Safe N (FSN) method. The calculated value of 5K+10, where K=19, is 5(19)+10=1055. The Fail-Safe N value was 1559 (Table 6) with a target significance level of 0.05 and $p<0.001$, indicating that the Fail-Safe N value exceeds 5K+10. These findings suggest that there is no publication bias in this meta-analysis.

A meta-analysis review by (Ridwan et al., 2022) detected no publication bias using the funnel plot, rank correlation method, and FSN, related to the application of cooperative learning methods in vocational high school mathematics education. Similarly, studies by (Nurhayati et al., 2023; A. D. Putri et al., 2024) using the Trim and Fill method identified no publication bias in the impact of innovative learning and RME on mathematics education.

Conclusion

The results of this study indicate that the application of innovative learning models, such as problem-based learning, blended learning, and metaverse applications, has a significant impact on enhancing mathematical and statistical critical thinking skills. Based on a meta-analysis of 19 studies across various educational levels, the overall effect size was found to be 1.03, indicating that innovative learning increases critical thinking skills by 103% compared to conventional teaching methods. Heterogeneity testing showed a high degree of variability ($I^2 = 91.37\%$), highlighting the need for moderator analysis to understand the factors influencing variations in study outcomes. Publication bias analysis using funnel plots, Egger's test, and Fail-Safe N confirmed that there is no significant publication bias in this research, making the meta-analysis results valid and reliable.

This study provides valuable insights for educators and policymakers to adopt innovative learning models to enhance the quality of mathematics and statistics education. However, the research is limited by the small number of studies, particularly those focusing on statistical critical thinking skills. Therefore, further research with a broader scope is needed, including the exploration of various innovative teaching techniques and their applications in diverse educational contexts. These findings emphasize the importance of pedagogical transformation toward learning that enhances students' critical thinking skills.

Recommendations

Based on the findings of this meta-analysis, several recommendations can be proposed to enhance educational practices, inform policymaking, and guide future research. First, educators at all levels, particularly in mathematics and statistics, are encouraged to adopt innovative learning models such as problem-based learning, blended learning, and metaverse-based instruction. These approaches have demonstrated a significant positive impact on students' critical thinking skills, especially when applied through student-centered and inquiry-based strategies. Professional development programs should be designed to equip teachers with the necessary pedagogical and technological competencies to implement these models effectively.

Second, policymakers are advised to support the integration of innovative pedagogies by revising curriculum guidelines and instructional frameworks to emphasize the cultivation of higher-order thinking skills. Investments in digital infrastructure are essential to facilitate the use of emerging technologies, such as virtual learning environments and immersive platforms, ensuring equitable access across educational institutions.

Lastly, future research should address the considerable heterogeneity identified in this study by exploring potential moderate variables such as educational level, subject focus, duration of intervention, and cultural context. There is also a notable gap in studies focusing specifically on statistical critical thinking, which warrants further investigation. Researchers are encouraged to employ rigorous designs and provide comprehensive data reporting to enhance the reliability and replicability of meta-analytic findings. These recommendations aim to support the transformation of mathematics and statistics education toward more innovative, inclusive, and effective teaching practices.

Acknowledgements

This article is part of a doctoral study on Educational Research and Evaluation pursued by the first author (Nusrotus Sa'idah) at Yogyakarta State University, Indonesia. We express our deepest gratitude to The Ministry of Education, Culture, Research, and Technology (KEMENDIKBUDRISTEK), The Center for Education Financial Services (PUSLAPDIK), and The Indonesia Endowment Funds for Education (LPDP) of The Republic of Indonesia for providing the Indonesia Educational Scholarships (BPI) so that the first author (Nusrotus Sa'idah) is able to pursue this academic degree.

References

Anwar, S., & Setyaningrum, W. (2021). Can Blended Learning Help Improve Students' Critical Thinking Skills? *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *10*(2), 721. https://doi.org/10.24127/ajpm.v10i2.3455

Apriliana, L. P., Handayani, I., & Awalludin, S. A. (2019). The Effect of a Problem Centered Learning on Student's Mathematical Critical Thinking. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *4*(2), 124–133. https://doi.org/10.23917/jramathedu.v4i2.8386

Arisoy, B., & Aybek, B. (2021). The Effects of Subject-Based Critical Thinking Education in Mathematics on Students' Critical Thinking Skills and Virtues. *Eurasian Journal of Educational Research*, *21*(92). https://doi.org/10.14689/ejer.2021.92.6

Ashidiq, R. M., Winarno, N., Prima, E. C., Widodo, A., & Chang, C.-Y. (2024). Investigating the Impact of STEM

Learning on Students' Critical Thinking Skills through Hand-Made Projector Activity. *Journal of Science Learning*, *7*(2), 187–203.

Bates, J., Cheng, S., Ferris, M., & Wang, X. (Frank). (2024). Cultivating Critical Thinking Skills: A Pedagogical Study in a Business Statistics Course. *Journal of Statistics and Data Science Education*, 1–11. https://doi.org/10.1080/26939169.2024.2394534

Belkin, D. (2015). Test Finds College Graduates Lack Skills for White Collar Jobs: Forty Percent of Students Seem Ill-Prepared to Enter Workforce; Critical Thinking Key. *The Wall Street Journal January 16*. https://www.wsj.com/articles/test-finds-many-students-ill-prepared-to-enter-work-force-1421432744

Borenstein, M. (Ed.). (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Borenstein, M. (2023). How to understand and report heterogeneity in a meta-analysis: The difference between I-squared and prediction intervals. *Integrative Medicine Research*, *12*(4), 101014. https://doi.org/10.1016/j.imr.2023.101014

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

Çelik, H. C., & Özdemir, F. (2020). Mathematical Thinking as a Predictor of Critical Thinking Dispositions of Pre-service Mathematics Teachers. *International Journal of Progressive Education*, *16*(4), 81–98. https://doi.org/10.29329/ijpe.2020.268.6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition).* NJ: Lawrence Erlbaum. https://doi.org/10.4324/9780203771587

Cooper, H. (2017). *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*. SAGE Publications, Inc. https://doi.org/10.4135/9781071878644

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The Handbook of Research Synthesis and Meta Analysis* ((2nd edition)). Russell Sage Publication.

Cortázar, C., Nussbaum, M., Harcha, J., Alvares, D., López, F., Goñi, J., & Cabezas, V. (2021). Promoting critical thinking in an online, project-based course. *Computers in Human Behavior*, *119*, 106705. https://doi.org/10.1016/j.chb.2021.106705

D'Angelo, C., Harris, C., & Rutstein, D. (2013). *Systematic Review and Meta-Analysis of STEM Simulations*. *2*.

Dogruer, S. S., & Akyuz, D. (2020). Mathematical Practices of Eighth Graders about 3D Shapes in an Argumentation, Technology, and Design-Based Classroom Environment. *International Journal of Science and Mathematics Education*, *18*(8), 1485–1505. https://doi.org/10.1007/s10763-019-10028-x

Duval, S., & Tweedie, R. (2000a). A Nonparametric "Trim and Fill" Method of Accounting for Publication Bias in Meta-Analysis. *Journal of the American Statistical Association*, *95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

Duval, S., & Tweedie, R. (2000b). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Fragkos, K. C., Tsagris, M., & Frangos, C. C. (2014). Publication Bias in Meta-Analysis: Confidence Intervals for Rosenthal's Fail-Safe Number. *International Scholarly Research Notices*, *2014*, 1–17. https://doi.org/10.1155/2014/825383

Gaspersz, M., & Salamor, R. (2021). Pembelajaran Grup Investigasi Berbantuan Spss Pada Mata Kuliah Statistik Untuk Meningkatkan Kemampuan Berpikir Kritis Dan Self Concept Matematis Mahasiswa Fkip. *Jurnal Magister Pendidikan Matematika (JUMADIKA)*, *3*(1), 26–34. https://doi.org/10.30598/jumadikavol3iss1year2021page26-34

Happy, N., & Widjajanti, D. B. (2014). KEEFEKTIFAN PBL DITINJAU DARI KEMAMPUAN BERPIKIR KRITIS DAN KREATIF MATEMATIS, SERTA SELF-ESTEEM SISWA SMP. *Jurnal Riset Pendidikan Matematika*, *1*(1), 48. https://doi.org/10.21831/jrpm.v1i1.2663

Haynes, A., Lisic, E., Goltz, M., Stein, B., & Harris, K. (2016). Moving Beyond Assessment to Improving Students' Critical Thinking Skills: A Model for Implementing Change. *Journal of the Scholarship of Teaching and Learning*, *16*(4), 44–61. https://doi.org/10.14434/josotl.v16i4.19407

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal of Educational Statistics*, *6*(2), 107. https://doi.org/10.2307/1164588

Helma Lia Sapitri, Zulhendri, Z., & Kasman Ediputra. (2024). Pengaruh Pendekatan Kontekstual terhadap Kemampuan Berpikir Kritis Matematis Peserta Didik SMP. *JURNAL PENDIDIKAN MIPA*, *14*(2), 474–480. https://doi.org/10.37630/jpm.v14i2.1572

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Idris, N., & Khaulah, S. (2020). Penerapan Model Pembelajaran Amora Terhadap Kemampuan Berpikir Kritis Siswa. *Jurnal Ilmiah Pendidikan Matematika Al Qalasadi*, *4*(2), 91–97. https://doi.org/10.32505/qalasadi.v4i2.2198

Katz, S., & Stupel, M. (2015). Promoting Creativity and Self-efficacy of Elementary Students through a Collaborative Research Task in Mathematics: A Case Study. *Journal of Curriculum and Teaching*, *4*(1), p68. https://doi.org/10.5430/jct.v4n1p68

Kutahya Dumlupinar University, Turkey, & Turkuresin, H. E. (2021). The Effect of Using Technology in Education on Academic Achievement of Students: The Case of Geographical Information Systems. *Education Quarterly Reviews*, *4*(2). https://doi.org/10.31014/aior.1993.04.02.294

Lestari, F. P., Ahmadi, F., & Rochmad, R. (2021). The Implementation of Mathematics Comic through Contextual Teaching and Learning to Improve Critical Thinking Ability and Character. *European Journal of Educational Research*, *volume–10–2021*(volume–10–issue–1–january–2021), 497–508. https://doi.org/10.12973/eu-jer.10.1.497

Marito, W., & Riani, N. (2022). Efektifitas Model Pembelajaran Blended Learning dalam Meningkatkan Kemampuan Berpikir Kritis dan Kemandirian Belajar Mahasiswa UPMI pada Mata Kuliah Statistik. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *6*(1), 223–233. https://doi.org/10.31004/cendekia.v6i1.1073

Nurhayati, R., Suranto, S., Dwiningrum, S. I., Retnawati, H., & Herwin, H. (2023). The Effect of Innovative Learning on Student Achievement in Indonesia: A Meta-Analysis. *Pegem Journal of Education and Instruction*, *13*(3). https://doi.org/10.47750/pegegog.13.03.17

Permatasari, A., Firmansyah, F., & Retnawati, H. (2024). The Effect of Test Preparation on English Proficiency

Performance of English Learners. *International Journal of Language Testing*, *Online First*. https://doi.org/10.22034/ijlt.2024.432975.1316

Pigott, T. D., & Polanin, J. R. (2020). Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review. *Review of Educational Research*, *90*(1), 24–46. https://doi.org/10.3102/0034654319877153

Putri, A. D., Juandi, D., & Turmudi, T. (2024). Realistic mathematics education and mathematical literacy: A meta-analysis conducted on studies in Indonesia. *Journal of Education and Learning (EduLearn)*, *18*(4), 1468–1476. https://doi.org/10.11591/edulearn.v18i4.21650

Putri, S. A., Fauzan, A., & Hamka, J. D. (2019). *Pengaruh Model Problem Based Learning Terhadap Kemampuan Berpikir Kritis Matematis Peserta Didik Di Kelas Viii Smp Negeri 2 Padang Panjang*. *8*(2).

Rachmadtullah, R., Setiawan, B., Wasesa, A. J. A., Wicaksono, J. W., & Rasmitadila. (2023). The utilization of metaverse technology applications based on science, technology, engineering and mathematics (Meta-STEM) to improve critical thinking skills. *Journal of Education and E-Learning Research*, *10*(4), 778–784. https://doi.org/10.20448/jeelr.v10i4.5203

Retnawati, H., Apino, E., Djidu, H., & Anazifa, R. D. (2018). *Pengantar Analisis Meta* (Edisi 1). Parama Publishing.

Reynders, G., Lantz, J., Ruder, S. M., Stanford, C. L., & Cole, R. S. (2020). Rubrics to assess critical thinking and information processing in undergraduate STEM courses. *International Journal of STEM Education*, *7*(1), 9. https://doi.org/10.1186/s40594-020-00208-5

Ridwan, M. R., Hadi, S., & Jailani. (2022). Identification of Effectiveness Measurements and Bias Publication of Literature Results Study: A Cooperative Learning Models on Mathematics Learning Outcomes of Vocational School Students in Indonesia. *Journal on Efficiency and Responsibility in Education and Science*, *15*(3), 189–200. https://doi.org/10.7160/eriesj.2022.150306

Ridwan, M. R., Retnawati, H., Hadi, S., & Jailani, J. (2021). The Effectiveness of Innovative Learning on Mathematical Problem-Solving Ability: A Meta-Analysis. *International Journal of Research in Education and Science*, *7*(3), 910–932. https://doi.org/10.46328/ijres.2287

Riley, R., Tierney, J., & Stewart, L. A. (2021). *Individual participant data meta-analysis* ((1st ed.)). Wiley.

Romero Ariza, M., Quesada Armenteros, A., & Estepa Castro, A. (2024). Promoting critical thinking through mathematics and science teacher education: The case of argumentation and graphs interpretation about climate change. *European Journal of Teacher Education*, *47*(1), 41–59. https://doi.org/10.1080/02619768.2021.1961736

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Ruppar, T. (2020). Meta-analysis: How to quantify and explain heterogeneity? *European Journal of Cardiovascular Nursing*, *19*(7), 646–652. https://doi.org/10.1177/1474515120944014

Sari, Y., Surya, E., & Asmin. (2019). The Increasing of Student's Mathematics Critical Thinking Ability Through Problem Based Learning. *Proceedings of the 4th Annual International Seminar on Transformative Education and Educational Leadership (AISTEEL 2019)*. Proceedings of the 4th Annual International Seminar on Transformative Education and Educational Leadership (AISTEEL 2019), Medan City, Indonesia. https://doi.org/10.2991/aisteel-19.2019.126

Schmidt, F. L., Oh, I., & Hayes, T. L. (2009). Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology*, *62*(1), 97–128. https://doi.org/10.1348/000711007X255327

Setiyawan, A. A., Agoestanto, A., & Isnarto, I. (2024). Enhancing self-regulated learning and critical thinking in STEM through flipped classroom models. *Jurnal Elemen*, *10*(2), 341–362. https://doi.org/10.29408/jel.v10i2.25322

Steven, D., La Ndia, L. N., & La Arapu, L. A. (2019). Pengaruh Model Problem Based Learning Terhadap Kemampuan Berpikir Kritis Matematis Peserta Didik Kelas Viii Smp Negeri 2 Kendari. *Jurnal Penelitian Pendidikan Matematika*, *7*(3), 15. https://doi.org/10.36709/jppm.v7i3.9278

Suryawan, I. I. P., Jana, P., Pujawan, I. G. N., & Putri, Putu. E. W. (2023). Improving Students' Critical Thinking Skills through an Ethnomathematically Controversial Problem-Based Multimodal Approach. *Pegem Journal of Education and Instruction*, *13*(3). https://doi.org/10.47750/pegegog.13.03.33

Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of Curriculum Studies*, *44*(3), 299–321. https://doi.org/10.1080/00220272.2012.668938

Yusriani, Y. Y., Sentryo, I., & Yasin, M. (2020). Perbedaan Kemampuan Berpikir Kritis Siswa Dalam Pembelajaran Matematika Antara Model Problem Based Learning Dan Model Pembelajaran Langsung Kelas Iv Sdn 95 Kendari. *Jurnal Ilmiah Pembelajaran Sekolah Dasar*, *2*(1), 10. https://doi.org/10.36709/jipsd.v2i1.13693

Zetriuslita, Z., Wahyudin, W., & Jarnawi, J. (2017). Mathematical Critical Thinking and Curiosity Attitude in Problem Based Learning and Cognitive Conflict Strategy: A Study in Number Theory course. *International Education Studies*, *10*(7), 65. https://doi.org/10.5539/ies.v10n7p65

Zulkarnain, M. A., Syaiful, S., & Suratno, S. (2023). Using learning models problem based learning to improve students' mathematical critical thinking skills. *Desimal: Jurnal Matematika*, *6*(2), 153. https://doi.org/10.24042/djm.v6i2.18087