A meta-analysis of the effects of problem-based learning on students' creative thinking in mathematics

Yunita¹, Dadang Juandi¹, Maximus Tamur², Angela Merici G Adem¹, Jerito Pereira⁴

Abstract: This study examined the effect of Problem-based Learning (PBL) on students' creative mathematical thinking in Indonesia during the last eight years using meta-analysis. Data were obtained from primary studies published in national journals, conference proceedings, and master thesis from 2012 to 2020. This study analyzed 19 effect sizes from 19 primary studies that fulfill the inclusion criteria. The effect size index used the Hedges' g equation, and statistical analysis was assisted by comprehensive meta-analysis (CMA) software. The effect size variation test was conducted by analyzing the study's five characteristics, namely the year of study, level of education, sample size, length of treatment, and publication sources. Based on the analysis results, the overall effect size was 0.821, with a standard error of 0.130 at the 95% confidence interval (CI) referring to the random-effects model. The effect size is classified as moderate, meaning that PBL has a moderately positive effect on students' creative mathematical thinking. The results of the analysis of the characteristics of the study show that the application of PBL to improve students' creative mathematical thinking relates to the research year (Q = 55.683; p < 0.05), the treatment duration (Q = 5.644; p < 0.05), and publication sources (Q = 13.812; p < 0.05). It was found that the PBL was more effective in conditions of treatment duration of fewer than four meetings. There is a tendency that PBL has a significant impact when it is first implemented. The findings' implications are

¹ Department of Mathematics Education, Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi No. 229, Bandung 40154, Indonesia, yunita1996nita@upi.edu
² Universitas Katolik Santu Paulus Ruteng, Jl. Jend A Yani 10, Ruteng, Indonesia
³ Teachers College Columbia University, USA
⁴ Guangxi Normal University, China

© Author(s), licensed under CC-BY-NC
discussed to provide important information as a basic idea for further PBL research and implementation.

**Keywords**: Problem-based learning, Creative thinking, Meta-analysis

### A. Introduction

Mathematical creative thinking is an essential prerequisite for academic achievement and success in the workplace (Bochniak, 2014). The development of creative thinking is significant for preparing students for future challenges and helping them solve problems (Ismail et al., 2018; Nuha, Waluya, & Junaedi, 2018; Fauzi, Dirgeyase, & Priyatno, 2019; Schoevers et al., 2019). In this case, the mastery of creative thinking in mathematics has far-reaching implications for students.

Mathematics instructions can be improved by implementing more appropriate models (Freeman-Green et al., 2015). Indeed, the learning process involving contextual problems can develop students' creative mathematical thinking (Maskur et al., 2020; Mustofà & Hidayah, 2020; Sugiharto et al., 2019; Thorndahl & Stentoft, 2020). Problem-based learning (PBL) is a learning model that helps students acquire knowledge and improve their skills (Servant-Miklos, 2018). PBL is considered one of the models which support the development of students’ creative thinking in mathematics.

Over the years, there appears to be an increase in the implementation of PBL in education (Ceker & Ozdamli, 2016) since it has considerable effects on students' creative mathematical thinking. However, previous research examining this theoretical assumption has shown inconsistent results. On the one hand, several studies (e.g., Nugroho, Chotim & Dwijanto, 2013; Khoirir, 2013; Katminingsih & Widodo, 2015) have shown that PBL is effective in improving students' creative thinking in mathematics. On the other hand, other research found that the creative mathematical thinking of groups of students taught with PBL is no better than those who learn with conventional approaches (e.g., Ahmad, & Gunawan, 2019; Alifiani, Dwijanto, & Cahyono, 2019; Indriani, Widyasari, & Amri, 2019). Despite the different results, teachers need convincing information to decide concerning the usefulness of the PBL with particular references to improving students' creative mathematical thinking.

To meet this need, it is necessary to integrate convincing quantitative findings to provide useful information for education policy (Higgins & Katsipataki, 2015). A meta-analysis study is required to integrate and interpret the findings of both the overall study and the variables of the primary study (Tamur, Juandi, & Kusumah, 2020). The meta-analysis provides profound and convincing conclusions (Siddaway, Wood, & Hedges, 2019). When there is a need to draw accurate and convincing conclusions, it is necessary to analyze the results of various primary studies (Kulik, Kulik, & Shwalb, 1986) with meta-analysis (Tamur, Juandi, & Kusumah, 2020; Yaakub, 1998).

Preceding meta-analyses have evaluated the effectiveness of PBL in general (e.g., Mustaffa et al., 2014; Asror, 2016; Yanto et al., 2019). Mustaffa et al. (2014) synthesized eighteen individual studies and only highlighted the role of PBL in the development of students' various cognitive and affective dimensions. Also, Asror (2016) analyzed ten individual studies on the effects of PBL on students' mathematical abilities. Meanwhile, Yanto et al. (2019) focus on study design and research methods of 15 individual studies. Moreover, Anugraheni (2018) has
conducted a unique meta-analysis testing the effectiveness of PBL on students' critical thinking. However, the study analyzed individual studies examining the effects of PBL in primary schools. Today, there has no meta-analysis that explicitly questions the effectiveness of PBL on students' overall creative mathematical thinking.

This study examined the overall effect of PBL and attempted to analyze the causes of variation in outcomes by examining the relationship between study characteristics. These characteristics contribute to revealing important information about how PBL will be implemented in the future. The questions addressed in the study were: (1) does the use of the PBL produce a more significant effect size on students' creative mathematical thinking than conventional approaches? And (2) does the effect size of students' creative mathematical thinking on the implementation of PBL between study groups vary in terms of the study year, education level, research class, sample size, treatment duration, and publication source?

B. Methods

This study applied a meta-analysis. It is a set of quantitative techniques for combining evidence from several related studies (Cumming, 2012). The results of the meta-analysis are to tell us whether or not all of the simultaneous effects of a host of authentic research does work or does not do so (Cleophas & Zwinderman, 2017). The procedures are as follows:

Firstly, empirical data is obtained from online databases, including ERIC (Education Resources Information Center), SAGE, and Springer. The three databases capture journal articles related to the research objectives. Furthermore, Semantic Scholar and Google Scholar databases are used to identifying journal articles, productions, and theses related to the research problems. The keywords used were "problem-based learning" and "creative thinking skills" to reach English articles. The keywords "pembelajaran berbasis masalah" and "kemampuan berpikir kreatif" were searched to achieve articles written in Bahasa. At this stage, we found thirty studies that examined the implementation of PBL on students' creative mathematical thinking.

Secondly, determining inclusion criteria, namely the feasibility of the standards used in selecting primary studies. At this stage, we have considered suggestions by Rothstein and Hopewell (2009) about the need to include “gray literature” such as unpaired theses and articles that complement other sources and reduce research bias. Thus, articles are selected regardless of the indexing. Following the research objectives, the inclusion criteria in this study were:
1. The articles were published from 2012 to 2020
2. The research design used experimental and quasi-experimental with the control class as a comparison
3. The duration of the treatment is at least three weeks
4. The treatment group used PBL
5. The independent variable measured is creative thinking
6. It contains statistical information for calculating effect sizes, i.e., mean, standard deviation, and sample size.

Based on the criteria, nineteen articles were obtained. Information regarding the studies is presented in Table 1.

The meta-analysis instrument was in the form of coding sheets (Tamur et al., 2020). The coding process can help identify lost or improperly coded data. The coding in this study was carried out by extracting information from nineteen primary studies into numerical data. This
work involved two coders, namely the principal investigator and the co-researcher, then discussed with the supervisor. To ensure that the data entered were not errors, the collaborating research team was involved in filling out the coding form and then carrying out verification.

**Table 1.** Primary studies on the effect of PBL on students’ creative mathematical thinking

<table>
<thead>
<tr>
<th>No</th>
<th>Study Characteristics</th>
<th>Group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Study year</td>
<td>2012-2013</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2013-2014</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2014-2015</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2015-2016</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2016-2017</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2017-2018</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2019-2020</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Education level</td>
<td>Primary Schools</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Secondary Schools</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High Schools</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Higher Education</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Sample size</td>
<td>≤ 30</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥ 31</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>Treatment duration</td>
<td>3-4 meetings</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt; 4 meetings</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>Publication source</td>
<td>Theses</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proceeding</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Journal</td>
<td>14</td>
</tr>
</tbody>
</table>

Bias risk was evaluated based on random sequence generation, disclosure, blinding, blinding the interveners, the blinding of results evaluators, incomplete data results, selective data reporting, and other factors (Schuch et al., 2016). In this study, the factor that can be identified is selective reporting, namely the journals' tendency only to publish articles that are considered significant. Because significant studies were more likely to be included in the meta-analysis than their unpublished counterparts, there is concern that the meta-analyses might overestimate accurate effect sizes (Borenstein et al., 2009). For this reason, the impact of publication bias must be controlled by examining the research funnel plot as a first step (Pigott & Polanin, 2020). In a meta-analysis, in order to measure the presence or absence of publication bias, a funnel plot was used. In the absence of bias and heterogeneity, the plot should, therefore, appear funnel-shaped with the greatest variability at the bottom and least variability at the top (Langan et al., 2012). The funnel plot has two elements: (i) the distribution of the effect size (in our case, the PBL effect size) on students' creative thinking; (ii) confidence bands around the overall mean to assess whether the observed results are statistically significant at the 5% level (Abramo et al., 2016). If the distribution of effect sizes is symmetrical around the vertical line on the funnel plot, the study is free from bias. However, if not, then the fail-safe N statistic (FSN) was used to check whether there was an effect of publication bias or not. Figure 1 presents the research funnel plot.
Figure 1 shows that the distribution of the effect sizes is not all symmetrical around the perpendicular line as the overall effect size. Therefore, Rosental fail-safe N (FSN) statistics were used to check for publication bias. The formula used is N / (5k + 10) (Fragkos, Tsagris, & Frangos, 2016), where N is the FSN value, and k is the number of studies included in the meta-analysis. If the numeric value obtained is greater than 1, it means that the result is far from bias. The results of the failure-safe N (FSN) statistical calculations using the CMA software are shown in Table 2.

Table 2. Rosental's FSN statistics

<table>
<thead>
<tr>
<th>No</th>
<th>Bias Conditions</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number of observed studies</td>
<td>19.00000</td>
</tr>
<tr>
<td>2</td>
<td>Number of missing studies that would bring p-value to &gt; alpha</td>
<td>971.00000</td>
</tr>
</tbody>
</table>

Based on Table 2, the FSN value and formula (Fragkos et al., 2016) state that 971 / (5 * 19 + 10) the result is 9.25 greater than 1. The result indicates the study included in the analysis is resistant to bias. Thus, there is no tendency for selective reporting to lead to overrepresentation.

Thirdly, using statistical analysis to evaluate the effectiveness of PBL in creative mathematical thinking. In conducting the meta-analysis, statistical information is needed for effect size transformations, namely the mean, standard deviation, and sample size for each published study (Paulay et al., 2012). The unit of analysis in meta-analysis is the effect size (Glass, 2015). The effect size is an index which states the magnitude of the influence of the independent variable on the dependent variable (Ellis, 2010). In this study, the effect size index uses the Hedges’ $g$ equation. The equation uses the standard mean difference is given as:

$$ g = J \times d $$

where $J = 1 - \frac{3}{4df-1}$ with $df = n_1 + n_2 - 2$

the degrees of freedom used in estimating $S_{within}$ and $d$ is the Cohen’s $d$ calculated using the equation and the variance of the Hedge’s $g$ effect size $V_g = J^2 \times V_d$ where $V_d$ is the variance of Cohen’s $d$ (Marfo & Okyere, 2019). In addition to the calculations of the weighted mean effect size of Hedges’ $g$ with its standard errors and 95% confidence intervals around each mean, CMA was also used to test homogeneity by calculating Q, p, and $I^2$ (Hillmayr et al., 2020). The interpretation of the calculated effect size using criteria (Cohen, Manion, & Morrison, 2002) as follows:
• it is weak if between 0 and 0.20;
• it is small if between 0.21 and 0.50;
• it is moderate if between 0.51 and 1.00; and
• it is large if higher than 1

If a statistical test at a predetermined level of probability (p <0.05) is reached, the null hypothesis is rejected, and the alternative hypothesis is accepted (Dunst & Hamby, 2012). Within meta-analysis, the most important decision concerns the use of a fixed-effects model or a random-effects model (Haidich, 2010). The fixed-effect model suggests that study effects are equivalent between different studies. Therefore, there is no need to analyze the study. The random-effect model was used if the effect size was statistically heterogeneous ($Q_b > \chi^2.95; p < 0.05$). This shows that the results of our analysis are statistically significant for different types of study characteristics (Bayraktar, 2001).

C. Findings and Discussion

The calculation results regarding the overall effect size in each study with the help of CMA software are shown in Table 3.

Table 3. The effect size of each study

<table>
<thead>
<tr>
<th>No</th>
<th>Author</th>
<th>Effect Size</th>
<th>Standard Error</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Khoiri (2013)</td>
<td>1.75</td>
<td>0.28</td>
<td>1.20</td>
<td>2.31</td>
</tr>
<tr>
<td>2</td>
<td>Nugroho, et al. (2013)</td>
<td>0.94</td>
<td>0.25</td>
<td>0.46</td>
<td>1.43</td>
</tr>
<tr>
<td>3</td>
<td>Azmi, et al. (2014)</td>
<td>0.95</td>
<td>0.26</td>
<td>0.45</td>
<td>1.45</td>
</tr>
<tr>
<td>4</td>
<td>Anwar (2015)</td>
<td>0.75</td>
<td>0.26</td>
<td>0.25</td>
<td>1.26</td>
</tr>
<tr>
<td>5</td>
<td>Katminingsih and Widodo (2015)</td>
<td>0.95</td>
<td>0.22</td>
<td>0.53</td>
<td>1.37</td>
</tr>
<tr>
<td>6</td>
<td>Kurniawati (2016)</td>
<td>0.00</td>
<td>0.22</td>
<td>-0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>7</td>
<td>Nurqolbiah (2016)</td>
<td>0.24</td>
<td>0.23</td>
<td>-0.21</td>
<td>0.68</td>
</tr>
<tr>
<td>8</td>
<td>Rochani (2016)</td>
<td>0.71</td>
<td>0.25</td>
<td>0.22</td>
<td>1.20</td>
</tr>
<tr>
<td>9</td>
<td>Arhasy and Mulyani (2017)</td>
<td>1.01</td>
<td>0.24</td>
<td>0.55</td>
<td>1.48</td>
</tr>
<tr>
<td>10</td>
<td>Fahrudin (2017)</td>
<td>0.71</td>
<td>0.29</td>
<td>0.15</td>
<td>1.27</td>
</tr>
<tr>
<td>11</td>
<td>Septian and Rizkiandi (2017)</td>
<td>1.75</td>
<td>0.27</td>
<td>1.21</td>
<td>2.28</td>
</tr>
<tr>
<td>12</td>
<td>Astuti, et al. (2018)</td>
<td>3.22</td>
<td>0.49</td>
<td>2.27</td>
<td>4.18</td>
</tr>
<tr>
<td>13</td>
<td>Risnawati, et al. (2018)</td>
<td>1.11</td>
<td>0.14</td>
<td>0.82</td>
<td>1.39</td>
</tr>
<tr>
<td>14</td>
<td>Ahmad and Gunawan (2019)</td>
<td>0.08</td>
<td>0.28</td>
<td>-0.47</td>
<td>0.62</td>
</tr>
<tr>
<td>15</td>
<td>Alifiani, et al. (2019)</td>
<td>0.49</td>
<td>0.25</td>
<td>0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>16</td>
<td>Indriani, et al. (2019)</td>
<td>0.24</td>
<td>0.26</td>
<td>-0.27</td>
<td>0.75</td>
</tr>
<tr>
<td>17</td>
<td>Masitoh and Prasetiyawan (2019)</td>
<td>0.19</td>
<td>0.23</td>
<td>-0.25</td>
<td>0.63</td>
</tr>
<tr>
<td>18</td>
<td>Masitoh (2019)</td>
<td>0.92</td>
<td>0.29</td>
<td>0.36</td>
<td>1.49</td>
</tr>
<tr>
<td>19</td>
<td>Sefrinal (2019)</td>
<td>0.50</td>
<td>0.25</td>
<td>0.02</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 3 shows that the mean effect size of each study with 95% confidence limit is in the range of 0.00 to 3.22. Based on the interpretation of effect sizes (Cohen et al., 2002), there were three studies that had weak effect sizes, four studies had small effect sizes, seven studies had moderate effect sizes, and five studies had large effect sizes.

The information presented in Table 4 shows that the fixed effects model with a 95% confidence level has a lower limit value of 0.669 and an upper limit of 0.886 with the combined effect size of 0.777. The size of this effect is moderate. Referring to the results of the homogeneity test using the Q statistic, the Q value is 81.175. Furthermore, the value of the Q
table with a significance level of 5% and degrees of freedom of 17 is 28.869. In this case, $Q > Q_{table}$, so it is concluded that the effect size of the implementation of PBL on students' creative mathematical thinking is heterogeneous. Thus, the estimation uses a random-effects model. In the random-effects model, the combined effect size was 0.821 (moderate) with a 95% confidence level. In the statistical significance test to answer the first question, the $z$ test was used. Under the random-effects model, the $z$-value is 6.295, and $p$ is 0.000. Because the $p$-value <0.05, it was concluded that the use of the PBL resulted in a more significant effect size on students' creative mathematical thinking than conventional approaches.

Table 4. The results of the meta-analysis comparing the effects model

<table>
<thead>
<tr>
<th>No</th>
<th>Model</th>
<th>n</th>
<th>Z</th>
<th>P</th>
<th>$Q_b$</th>
<th>$I^2$ squared (p=0.05)</th>
<th>Effect size</th>
<th>Standard error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fixed effect</td>
<td>19</td>
<td>14.001</td>
<td>0.000</td>
<td>95.617</td>
<td>81.175</td>
<td>0.777</td>
<td>0.056</td>
<td>0.669 - 0.886</td>
</tr>
<tr>
<td>2</td>
<td>Random effects</td>
<td>19</td>
<td>6.295</td>
<td>0.000</td>
<td>0.821</td>
<td>0.130</td>
<td></td>
<td></td>
<td>0.566 - 1.077</td>
</tr>
</tbody>
</table>

This finding is in line with the results of Susanti, Juandi, and Tamur (2020), who found an effect size of 0.79 when they synthesized the results from 12 articles comparing the effectiveness of PBL on mathematical communication skills. Similar results were found by Paloloang et al. (2020), who found an effect size of 0.83 when they analyzed the results of 14 articles comparing the effectiveness of PBL on students' mathematical literacy abilities. However, this study's overall size is different from the results of Asror's (2016) study, which reported an effect size of 0.51 when he analyzed the impact of PBL on students' creative mathematical thinking. In general, this research is supported by another finding (Mustaffa et al., 2014) that PBL positively impacts mathematics learning. This shows that PBL and other cooperative learning models have a moderate effect on mathematics learning, specifically on students' creative mathematical thinking.

Of the 19 studies, each had an average effect size. The characteristics of the studies were carried out including; year of study, education level, sample size, duration of the experiment and research publication. A summary of the analysis results based on the research characteristics is presented in Table 5. The interpretations of Table 5 are presented as follows.

Firstly, the analysis of research characteristics in Table 5 shows that PBL is associated with differences in the study years ($Q = 55.68; P <0.05$). PBL implementation was most effective (combined effect size = 1.294) between 2012 and 2013 and ineffective (combined effect size = 0.261) between 2018 and 2020. It appears that the combined effect size of each study group was different. Surprisingly, the combined effect sizes of the most recent study groups were so small. According to Bayraktar (2001), this kind of phenomenon implies the Hawthorne effect that first-time PBL has a significant impact on students' creative mathematical thinking, perhaps because of its novelty. Similar results were also reported by Masitoh and Prasetyawan (2019) that PBL had a big impact when it was first implemented. The result of $Q_b$ is 55,683 > $Q_{table}$ is 12.59 with a significance level of 0.05, meaning that there are significant differences between groups in the aspects of the research year.
Secondly, based on the level of education, it consists of four groups: higher education, high schools, secondary schools, and primary schools, each of which has an effect size of 0.705; 0.475; 0.851; 0.929 (moderate), respectively. The result of \( Q_b \) is 6.806 < \( Q_{table} \) is 7.82 with a significance level of 0.05. In this case, there is no significant difference between groups at the level of education. On the other hand, it was found that PBL has a higher positive effect on secondary schools than high schools (Asror, 2016). Meanwhile, the results of other meta-analyses regarding the effects of cooperative learning on student academic achievement show that cooperative learning is more effectively applied at the high school level since it is able to act independently and according to the role in the group (Prasetiyo et al., 2014). Although the models used are different, it is necessary to investigate these differences in results further.

Thirdly, based on the sample size, it consisted of two groups, \( \leq 30 \) and \( \geq 31 \), with effect sizes of 0.675 and 0.799 (moderate), respectively. Based on the results, \( Q_b \) is 0.711 < \( Q_{table} \) is 3.84 with a significant degree of 0.05. This shows that there is no significant difference between sample size groups. However, these findings differ from Tamur et al. (2020) that PBL is more effective in small samples. However, they focus on realistic problems. PBL and realistic problem-based learning have some similarities, but they produce different results. This difference in results may be due to the large number of studies synthesized in this study. For this reason, it is necessary to carry out further research involving more individual studies to be analyzed. This will provide more comprehensive results for educators and interested parties.

Fourthly, based on the duration of the treatment consisted of two groups, the treatment duration of 3-4 meetings had an effect size of 1.116 (large), and more than four meetings had an effect size of 0.726 (moderate). From the result, \( Q_b \) is 5.644 > \( Q_{table} \) is 3.84, there were significant differences between groups in the duration of the trials. This is also in line with the
finding of Kadir, Milama, and Khairunnisa (2013) that PBL will be more effective if the treatment duration is not more than three weeks. Thus, problem-based learning is more effective for creative mathematical thinking if 3-4 meetings are held. This result could be partial because PBL requires more time to solve problems (Sanjaya, 2007). On the other hand, in implementing PBL, the students' initial ability factors, level and speed of thinking, and other heterogeneous aspects make teachers compromise with time (Tyas, 2017).

Fifthly, based on the publication sources, this study concluded that the results found in the thesis document (effect size = 1.218) were greater than those in the proceeding articles (effect size = 0.889) and the journal articles (effect size = 0.765). This variable is analyzed to check whether selective reporting factors give rise to the effect of publication bias. The heterogeneity test results showed that the combined effect sizes of the three study groups were different. However, they showed no signs of bias. The reasons why the combined effect size of the thesis is much larger than that of proceedings or journals are still unclear.

The current research shows that PBL has a positive impact on students' creative thinking in mathematics. This study's effect size is consistent with previous studies' findings (e.g., Susanti et al., 2020; Paloloang et al., 2020), but the findings of this study contribute to the literature on study characteristics that need to be considered in implementing PBL. Although it has been suggested that teachers need convincing information regarding PBL implementation, this finding is based only on the primary studies that met the inclusion criteria. Further studies are still needed to be analyzed to gather relevant information. Overall, these findings do not describe the effectiveness of PBL on students' mathematical creativity.

D. Conclusion

This meta-analysis aimed to analyze the effectiveness of using PBL on students’ creative thinking in mathematics. The overall effect size was 0.821, with a standard error of 0.130 and 95% degrees of freedom. This shows that PBL has a moderate effect on students' creative mathematical thinking. The results of the studies’ analysis show that the application of PBL to the students' creative thinking relates to the study year, research class, treatment duration, and publication sources. It was found that the PBL is more effective in conditions of treatment duration of fewer than four meetings. There is a tendency that the implementation of PBL for the first time has a significant impact on students' creative thinking. Besides, the results of the analysis also show variations in the effect sizes according to the published sources. Although this does not indicate a publication bias, the hypothesis that the results of a significant study are likely to be published needs to be tested again. For this reason, more studies need to be carried out in this area to provide a complete picture to mathematics educators about the effectiveness of PBL in mathematics classrooms.

Acknowledgment

The authors thank M. Borenstein and the CMA team from Englewood, the USA, for all their technical assistance and contributions to the research project.
References


