META-ANALYSIS OF COMPUTER-BASED MATHEMATICS LEARNING IN THE LAST DECADE SCOPUS DATABASE: TRENDS AND IMPLICATIONS

Maximus Tamur1*, Sabina Ndiung1, Robert Weinhandl2, Tommy Tanu Wijaya3, Emilianus Jehadus1, Eliterius Sennen1
1Universitas Katolik Indonesia Santu Paulus Ruteng, Indonesia
2Linz School of Education, STEM Education, Johannes Kepler Universität Linz, Austria
3School of Mathematical Sciences, Beijing Normal University, Beijing, China

ABSTRACT

Computer-Based Mathematics Learning (CBML) has gone global in the last decade and is making a substantial impact for educational purposes. But the fact is that in the scientific literature, it is found that studies aimed at testing these theoretical assumptions have inconsistent results. In this regard, this meta-analysis was conducted to determine the effect of CBML and to analyze categorical variables to consider the implications. Data were retrieved from the Scopus database using Publish or Perish between 2010 and 2023. This study examined 29 independent samples from 28 eligible primary studies with 1179 subjects. The population estimate was based on a random effects model, and the CMA software was used as a calculation aid. The study's results provide an overall effect size of 1.03 (large effect). This indicates that applying CBML significantly affects students' mathematical abilities. The four categorical variables considered in the study are discussed to clarify research trends. Furthermore, the research implications are outlined and contribute to future CBML implementation arrangements.

1. INTRODUCTION

Computer-based Mathematics Learning (CBML) has been widely implemented and included in the teaching system in almost every country. The use of CBML in the classroom is associated with students’ academic abilities because teachers provide exciting learning opportunities and broader content (de Mendivil et al., 2019; McLaren et al., 2017; Xin et al., 2020; Zhang & Wang, 2020). Applying CBML significantly influences students' mathematical abilities (Kim et al., 2022; Nurjanah et al., 2020; Pereira et al., 2021;
Shyshkina et al., 2018). It can be said that the use of CBML has become a trend in learning mathematics in the last decade.

The use of CBML in learning is increasingly widespread because it presents content numerically, graphically and symbolically without the additional time burden of manually calculating complex computational problems (Juandi, Kusumah, Tamur, Perbowo, Siagian, et al., 2021; Juandi, Kusumah, Tamur, Perbowo, & Wijaya, 2021; Tamur, Kusumah, Juandi, Kurnila, et al., 2021; Tamur, Kusumah, Juandi, Wijaya, et al., 2021). Integrating computer technology into learning mathematics will help students make connections in mathematics by making the learning process more realistic and effective (Zaldívar-Colado et al., 2017). CBML implementation will be more interesting, inventive, and exploratory (Aungamuthu, 2013; Foster et al., 2016; Ochkov & Bogomolova, 2015; Tamur et al., 2022). These conditions allow students to be more active and successful in learning (Das et al., 2021; Tatar et al., 2014; Timmers et al., 2013).

CBML provides useful feedback to support students' representation, visualization, and critical abilities (Granberg & Olsson, 2015). Applying CBML can improve Students' Mathematical Ability (SMA). However, in the scientific literature, studies testing these theoretical assumptions provide inconsistent findings. Several studies provide findings stating the superiority of the CBML group (e.g., Ishartono et al., 2022; Jelatu et al., 2018; Kariadinata et al., 2019; Ningsih & Paradesa, 2018; Nurjanah et al., 2020; Takači et al., 2015), whereas other studies have found that CBML implementation does not outperform the group that does not use it (e.g., del Cerro Velázquez & Morales Méndez, 2021; Hamid et al., 2020).

Various studies that analyze the same problem sometimes provide findings that vary and contradict each other. As a result, making decisions or conclusions can be subjective (Wicherts, 2020). The problem is that educators and other related parties need accurate information for making educational decisions and arrangements under what conditions the use of CBML achieves a high level of effectiveness in SMA.

The problems described above can be overcome by summarizing the various results of primary research. Quantitative research procedures can be chosen to provide accurate policy-making information (Higgins & Katsipataki, 2015). In this regard, the meta-analysis study is the choice because of its role specifically to integrate the findings of the primary study and investigate the reasons for the variation in results to be considered in its implementation in the future. Meta-analysis provides in-depth and accurate conclusions (Fadhli et al., 2020; Siddaway et al., 2019). Thus, when there is a need to draw conclusive conclusions, it is necessary to take into account the results of various individual studies using meta-analysis methods (Wicherts, 2020).

Previously, meta-analysis studies have been found in the literature regarding the effect of CBML implementation on SMA (e.g., Demir & Basol, 2014; Juandi, Kusumah, Tamur, Perbowo, & Wijaya, 2021; Tamur, Kusumah, Juandi, Kurnila, et al., 2021). However, these studies analyzed primary data from all databases, such as Google Scholar, ERIC, repository, library database, and others. In contrast to previous studies, this study analyzes data taken from the Scopus database. Apart from the fact that no studies specifically chose Scopus as the only search database, Scopus was also chosen because it applies consistent standards in selecting documents to be included in its index (Phan et al., 2022). Scopus also displays more documents than other top databases, such as the Web of Science, specifically for research reviews in education and social sciences (Hallinger & Chatpinyakoop, 2019; Hallinger & Nguyen, 2020). This study also fills a gap in the literature regarding a comparative picture of CBML implementation in Indonesia and abroad.

Previous studies have also shown inconsistent trends in the variables studied in terms of year of publication, source of publication, and level of education (e.g., Juandi, Kusumah,
Tamur, Perbowo, & Wijaya, 2021; Tamur et al., 2020; Tamur, Kusumah, Juandi, Wijaya, et al., 2021). This meta-analysis is necessary to evaluate the implementation of CBML to see the overall trend clearly and also consider the implications. Thus, this study will directly contribute to CBML science and practice in the future, especially concerning educational settings. This research answers explicitly two questions, namely; (1) whether the mean effect size representing the intervention for each study on the effect of CBML on SMA is significantly different from zero, and (2) whether there is a difference in the magnitude of the effect of CBML on SMA based on categorical variables.

2. METHOD

This study focuses on analyzing the influence of the quantitative profile of the study group from the Scopus database, which specifically investigates the effectiveness of CBML on students' mathematical abilities (SMA). This work was conducted to determine the magnitude of the overall effect of CBML and to analyze the role of categorical variables in the differences in effect sizes of each study. To achieve this goal, a meta-analysis approach was applied. Meta-analysis was developed to compare the results of various primary studies and its role is very vital for scientific development, namely providing a more objective method of drawing conclusions or decisions when various primary studies provide varying results (Cooper, 2017; Schmidt & Hunter, 2015). In general, meta-analysis research begins with formulating research problems and hypotheses, followed by a literature search, then coding variables, then statistical analysis, and ends with the interpretation of findings (Çoğaltay & Karadağ, 2015). This work also follows these stages.

2.1. Literature Search

The Scopus database was chosen as the document search location in this study. Furthermore, the Publish or Perish program is used to download a study on the application of CBML with the keyword combination used being geogebra mathematical ability; Cabri mathematical ability; maple mathematical abilities; mathematical ability algebrator, Matlab Mathematical Ability, MatchCAD, and Wingeom. As an example, Figure 1 presents the process of tracing studies related to implementing CBML on the Scopus database using the PoP application. The search results obtained 193 related studies.

Figure 1. The process of tracing data using the PoP application.
2.2. Literature Inclusion Criteria

In this study, studies that were successfully identified using the PoP application were then shared according to the following criteria:

a. Study in an English setting, and retrieved from Scopus database between 2010-2023.

b. Provide statistical information to obtain effect size values. Studies that do not meet this will be excluded from the analysis (e.g., Dockendorff & Solar, 2018; Jacinto & Carreira, 2017; Zulnaidi et al., 2020).

c. Study with a quantitative approach and must involve a control group as a comparison. Studies that used only one sample or used a qualitative approach were excluded from the analysis.

Furthermore, in this study, the data filtering process used the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses) protocol, as shown in Figure 2.

![Figure 2. Filtering data through PRISMA](image)

Figure 2 presents the process of filtering data using PRISMA. Finally, there are 28 studies, and because there are studies that investigate more than one independent sample, in the study conducted by Apriatna et al. (2020) then, 29 effect sizes were examined.

2.3. Coding

All eligible studies were coded using a detailed coding scheme based on the coding manual and coding protocol. Each preliminary study included in the analysis was coded. In this work, the research instrument is a coding sheet specially created to extract information from individual studies converted into numerical data. The coding was carried out by two
students who had previously been specifically trained according to guidelines from Cooper (2017). To determine reliability, a random sample of 5 from the eligible studies was duplicated and distributed among the coders. Each coder was provided with a copy of the article, coding form and coding protocol. To assess the reliability between coders, the percent agreement or abbreviated PA is used. The PA method is the easiest and most intuitive approach to building reliability (Syed & Nelson, 2015). PA is simply the ratio of items that the two coders agree on to the total number of items assessed and is calculated using the following formula:

\[ PA = \frac{N_A}{N_A + N_P} \times 100 \]

Where \( N_A \) is the total number of agreements and \( N_P \) is the total number of disagreements. An agreement rate of 85% or greater is predetermined to be considered high. From the calculation results, the level of both PAs is 94.26%. The figure indicates there is a substantial to near-perfect match between the coders. Thus the instrument to be used in this study is reliable.

2.4. Statistical Analysis

The parameter used to estimate the population in a meta-analysis study is the effect size. In this study, the effect size is defined as the magnitude of the effect of CBML on SMA. CMA software stands for Comprehensive Meta-Analysis used to transform the effect size of each study, including the overall effect size, P-value, Q statistics, and confidence intervals. The program also draws funnel plots and research forest plots. Hedges' g metrics were applied in this study, and the interpretation of effect sizes was based on the classification of Cohen et al. (2017) i.e., less than 0.2 (negligible), 0.2 to 0.5 (small effect), 0.5 to 0.8 (medium effect), 0.8 to 1.3 (large effect), and more than 1.3 (very large effect). The random effects model was chosen as the estimation method because it does not assume that all the studies analyzed have the same true effect (Pigott & Polanin, 2020). In this study, the random effect model was determined after fulfilling the heterogeneity test. This test is carried out by observing the p-value. The null hypothesis, which states that all studies are the same (homogeneous), is rejected if the p-value <0.05. Rejecting the null hypothesis indicates that effect sizes between studies or study groups may not measure the same population parameters (Çoğaltay & Karadağ, 2015). Conditions show that differences in study categories affect study effect sizes.

An analysis of publication bias was performed to prevent misrepresentation of the findings. Publication bias reflects the fact that it is more likely that articles deemed statistically significant are published. In addition to the scientific facts that about 6% of researchers rarely try to publish research that is not significant (Cooper, 2017). As a result, the aggregate effect size obtained may be exaggerated (Juandi et al., 2022; Park & Hong, 2016). To anticipate this, funnel plots were examined to assess the possible amount of bias, and Rosenthal's FSN statistic was used to assess the impact of bias (Çoğaltay & Karadağ, 2015). In this study, it is said to be resistant to publication bias if the distribution of effect sizes is symmetrical around the vertical line. However, if the fact is that the effect size scatter is not completely symmetrical, then a trim and fill procedure is used. Then, if it is found that the observed effect and the virtual effect created according to the random effects model are the same then the study is resistant to the influence of publication bias.
3. RESULT AND DISCUSSION

3.1. Results

First, the results of research aimed at answering question one are presented. After screening the primary studies, 28 articles from the Scopus database and 29 independent samples were analyzed in this study. Figure 3 presents the research forest plot consisting of the study name, ES (effect size), standard error, variance per the study, confidence intervals, z values, and p values.

![Plot forest research](image)

Figure 3. Research forest plot

Figure 3 shows the broad confidence levels and inconsistent response rates. Descriptively this shows the effect size of each study is heterogeneous. Statistically, this needs to be checked so that the estimation method is in accordance with the initial assumptions. Table 1 presents a summary of the overall analysis to answer question 1 as well as to determine the estimation method.

<table>
<thead>
<tr>
<th>Study name</th>
<th>Hedges's g</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priatna, 2017a</td>
<td>0.79</td>
<td>0.26</td>
<td>0.07</td>
<td>0.27</td>
<td>1.31</td>
<td>2.97</td>
<td>0.00</td>
</tr>
<tr>
<td>Priatna, Martadiputra &amp; Wibranjo, 2018</td>
<td>0.92</td>
<td>0.27</td>
<td>0.07</td>
<td>0.40</td>
<td>1.45</td>
<td>3.44</td>
<td>0.00</td>
</tr>
<tr>
<td>Munandar &amp; Sanrini, 2020</td>
<td>1.97</td>
<td>0.32</td>
<td>0.10</td>
<td>1.34</td>
<td>2.61</td>
<td>6.11</td>
<td>0.00</td>
</tr>
<tr>
<td>Juandi &amp; Priatna, 2018</td>
<td>0.36</td>
<td>0.25</td>
<td>0.06</td>
<td>0.13</td>
<td>0.86</td>
<td>1.43</td>
<td>0.15</td>
</tr>
<tr>
<td>Narah &amp; Ngurum, 2021</td>
<td>0.50</td>
<td>0.25</td>
<td>0.06</td>
<td>0.12</td>
<td>1.00</td>
<td>1.95</td>
<td>0.05</td>
</tr>
<tr>
<td>Takisu, Starkey, &amp; Milanovic, 2015</td>
<td>0.68</td>
<td>0.11</td>
<td>0.01</td>
<td>0.46</td>
<td>0.69</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td>Septian, Deth, &amp; Prabawono, 2020</td>
<td>2.23</td>
<td>0.38</td>
<td>0.15</td>
<td>1.58</td>
<td>3.08</td>
<td>6.09</td>
<td>0.00</td>
</tr>
<tr>
<td>Fharnawati, 2019</td>
<td>3.22</td>
<td>0.37</td>
<td>0.13</td>
<td>1.60</td>
<td>3.04</td>
<td>6.32</td>
<td>0.00</td>
</tr>
<tr>
<td>Saha, Ayu, &amp; Tarmizi, 2010</td>
<td>0.60</td>
<td>0.28</td>
<td>0.08</td>
<td>0.05</td>
<td>1.14</td>
<td>2.13</td>
<td>0.05</td>
</tr>
<tr>
<td>Mayrano, Rediah, &amp; Sap, 2020</td>
<td>1.10</td>
<td>0.27</td>
<td>0.07</td>
<td>0.52</td>
<td>1.63</td>
<td>4.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Pidha, Priamirah, &amp; Widjajani, 2019</td>
<td>0.73</td>
<td>0.30</td>
<td>0.09</td>
<td>0.14</td>
<td>1.32</td>
<td>2.44</td>
<td>0.01</td>
</tr>
<tr>
<td>Nagara, Wahyudin, Naratiah, &amp; Herman, 2022</td>
<td>0.72</td>
<td>0.24</td>
<td>0.06</td>
<td>0.26</td>
<td>1.19</td>
<td>3.05</td>
<td>0.00</td>
</tr>
<tr>
<td>Mumi, Sariyasa, &amp; Astana, 2017</td>
<td>3.77</td>
<td>0.43</td>
<td>0.18</td>
<td>2.81</td>
<td>4.51</td>
<td>8.81</td>
<td>0.00</td>
</tr>
<tr>
<td>Khali, Fanoo, Cai, &amp; Khal, 2018</td>
<td>0.82</td>
<td>0.32</td>
<td>0.10</td>
<td>0.19</td>
<td>1.45</td>
<td>2.53</td>
<td>0.01</td>
</tr>
<tr>
<td>Zeliqalitha, Nohiyanti, &amp; Fakheromah, 2019</td>
<td>0.01</td>
<td>0.31</td>
<td>0.10</td>
<td>0.59</td>
<td>0.62</td>
<td>0.04</td>
<td>0.97</td>
</tr>
<tr>
<td>Zulhaidi, &amp; Zart, 2017</td>
<td>0.72</td>
<td>0.51</td>
<td>0.31</td>
<td>0.31</td>
<td>0.74</td>
<td>0.74</td>
<td>0.00</td>
</tr>
<tr>
<td>Nasrah, Latif, Yuliani, &amp; Tamur, 2020</td>
<td>0.43</td>
<td>0.24</td>
<td>0.06</td>
<td>0.03</td>
<td>0.90</td>
<td>1.62</td>
<td>0.07</td>
</tr>
<tr>
<td>Kusumah, Kuswati, &amp; Herman, 2020</td>
<td>0.79</td>
<td>0.23</td>
<td>0.05</td>
<td>0.34</td>
<td>1.23</td>
<td>3.47</td>
<td>0.00</td>
</tr>
<tr>
<td>Jalakhi, Sariyasa, &amp; Astana, 2018</td>
<td>1.12</td>
<td>0.27</td>
<td>0.08</td>
<td>0.59</td>
<td>1.66</td>
<td>4.08</td>
<td>0.00</td>
</tr>
<tr>
<td>Ningiar, &amp; Palekade, 2018</td>
<td>1.53</td>
<td>0.29</td>
<td>0.08</td>
<td>0.97</td>
<td>2.10</td>
<td>5.32</td>
<td>0.00</td>
</tr>
<tr>
<td>Aprilatna, Budiyono, &amp; Indrati, 2000a</td>
<td>0.60</td>
<td>0.26</td>
<td>0.07</td>
<td>0.10</td>
<td>1.11</td>
<td>2.35</td>
<td>0.02</td>
</tr>
<tr>
<td>Aprilatna, Budiyono, &amp; Indrati, 2000b</td>
<td>0.63</td>
<td>0.26</td>
<td>0.07</td>
<td>0.13</td>
<td>1.14</td>
<td>2.45</td>
<td>0.01</td>
</tr>
<tr>
<td>Karakindani et al., 2019</td>
<td>1.35</td>
<td>0.31</td>
<td>0.10</td>
<td>0.70</td>
<td>1.96</td>
<td>4.38</td>
<td>0.00</td>
</tr>
<tr>
<td>Muntasimin, &amp; Mutani, 2018</td>
<td>0.97</td>
<td>0.32</td>
<td>0.10</td>
<td>0.34</td>
<td>1.60</td>
<td>3.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Ishalotto et al., 2012</td>
<td>0.15</td>
<td>0.26</td>
<td>0.07</td>
<td>0.15</td>
<td>0.65</td>
<td>0.39</td>
<td>0.05</td>
</tr>
<tr>
<td>Velezquez, &amp; Mendez, 2021</td>
<td>0.32</td>
<td>0.22</td>
<td>0.05</td>
<td>0.12</td>
<td>0.75</td>
<td>1.44</td>
<td>0.15</td>
</tr>
<tr>
<td>Priatna, 2017b</td>
<td>2.85</td>
<td>0.33</td>
<td>0.11</td>
<td>2.19</td>
<td>3.50</td>
<td>8.56</td>
<td>0.00</td>
</tr>
<tr>
<td>Zulnadi et al., 2019</td>
<td>0.48</td>
<td>0.22</td>
<td>0.05</td>
<td>0.04</td>
<td>0.92</td>
<td>2.16</td>
<td>0.03</td>
</tr>
<tr>
<td>Septian et al., 2020</td>
<td>1.57</td>
<td>0.34</td>
<td>0.12</td>
<td>0.91</td>
<td>2.34</td>
<td>4.63</td>
<td>0.00</td>
</tr>
<tr>
<td>1.01</td>
<td>0.12</td>
<td>0.01</td>
<td>0.78</td>
<td>1.25</td>
<td>8.58</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

![Table 1](image)

Table 1. The meta-analysis results according to the estimation model

<table>
<thead>
<tr>
<th>Model</th>
<th>N</th>
<th>Hedges's g</th>
<th>Standard error</th>
<th>Test of null</th>
<th>Z-value</th>
<th>P-value</th>
<th>Q</th>
<th>P</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed-effects</td>
<td>29</td>
<td>0.81</td>
<td>0.04</td>
<td>18.38</td>
<td>0.00</td>
<td>186.87</td>
<td>0.00</td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>Random-effects</td>
<td>29</td>
<td>1.03</td>
<td>0.01</td>
<td>8.57</td>
<td>0.00</td>
<td>186.87</td>
<td>0.00</td>
<td>Accept</td>
<td></td>
</tr>
</tbody>
</table>

When Table 1 is observed, it appears that the P value <0, which means that the distribution of effect sizes for each study is heterogeneous. This means that the estimation method for the population in this study is in accordance with the random effect model. Table
Table 1 is observed it appears that the P value for the test of null is less than 0 based on the random effects model. This means that the mean effect size representing the intervention for each study of the effect of CBML on SMA is significantly different from zero. In other words, the overall results of the study clarify the superiority of the CBML group. The overall effect size based on the random effects model is 1.03 which is classified as a large effect according to Cohen et al. (2017). It is then necessary to calculate whether this overall effect size is related to publication bias then the funnel plot of the study is observed. Figure 4 presents the funnel plot of the study.

![Funnel Plot of Standard Error by Hedges's g](image)

**Figure 4.** Funnel plot of 29 independent samples

When it is observed in Figure 4, the study ESs are spread not entirely symmetrically in the center of the funnel plot. Therefore, it is necessary to observe whether these results resist publication bias's influence. For this reason, the Trim and Fill testing procedure is carried out as illustrated in Figure 5.

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Random Effects</th>
<th>G Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies Trimmed</td>
<td>Point Estimate</td>
<td>Lower Limit</td>
</tr>
<tr>
<td>Observed values</td>
<td>0.000230</td>
<td>0.72224</td>
</tr>
<tr>
<td>Adjusted values</td>
<td>0.000230</td>
<td>0.72224</td>
</tr>
</tbody>
</table>

**Figure 5.** Trim and fill test

The trim and fill test results as illustrated in Figure 5 show that there is no difference between the observed effect sizes and the virtual effects created according to the random-effects model. Thus, there was no publication bias in this study or no studies were pruned or added due to publication bias. So the overall effect size calculated as 1.03 and categorized as large effect is not associated with publication bias. This value is not exaggerated.
Second, the research results are presented to answer question two. It has previously been shown that the estimation method fits the random effects model. This shows that the effect size of each study is heterogeneous so that categorical variables that affect the relationship between CBML and SMA must be investigated (Arik & Yilmaz, 2020). The summary of the analysis related to these categorical variables is illustrated in Table 2.

### Table 2. Summary of category variable analysis

<table>
<thead>
<tr>
<th>Category Variables</th>
<th>Category</th>
<th>N</th>
<th>Hedge's g</th>
<th>Heterogeneity (Qb)</th>
<th>df(Q)</th>
<th>P</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>30 or less</td>
<td>18</td>
<td>0.92</td>
<td>4.32</td>
<td>1</td>
<td>0.03</td>
<td>Reject H0</td>
</tr>
<tr>
<td></td>
<td>31 or over</td>
<td>11</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational stage</td>
<td>Junior high school (JHS)</td>
<td>17</td>
<td>0.87</td>
<td>6.99</td>
<td>2</td>
<td>0.03</td>
<td>Reject H0</td>
</tr>
<tr>
<td></td>
<td>Senior high school (SHS)</td>
<td>5</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>College</td>
<td>7</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Sources</td>
<td>Journal</td>
<td>12</td>
<td>0.57</td>
<td>48.6</td>
<td>2</td>
<td>0.00</td>
<td>Reject H0</td>
</tr>
<tr>
<td></td>
<td>Proceeding</td>
<td>17</td>
<td>1.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Indonesia</td>
<td>23</td>
<td>0.99</td>
<td>22.75</td>
<td>1</td>
<td>0.00</td>
<td>Reject H0</td>
</tr>
<tr>
<td></td>
<td>Foreign</td>
<td>6</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.2. Discussion

The first objective of this research is to determine the magnitude of the overall effect of applying CBML on SMA. The analysis results gave the overall effect size of 1.03 and were accepted as a large effect. These results are supported by a previous meta-analysis conducted by Tamur et al. (2020) regarding the effect of implementing mathematical software in Indonesia. Although the database and population used differ from this study, the overall effect size is almost the same, namely 1.16. In addition, the results of this study are almost the same as the findings of studies conducted by Juandi, Kusumah, Tamur, Perbowo and Wijaya (2021) and Juandi, Kusumah, Tamur, Perbowo, Siagian, et al. (2021) where the overall study effect sizes were 1.07 and 0.96, respectively. This shows a nearly similar overall trend with respect to CBML adoption.

In this study, the number of subjects was 1127 people, and the average sample size was 44 people. Related to that, the effect size of 1.03 can be interpreted as students who are ranked 22 in the experimental class are equivalent to students who are ranked 8 in the control class. This illustration illustrates the magnitude of the influence of CBML on SMA. This condition is explained in the scientific literature that the application of CBML helps increase students' knowledge of mathematical concepts and procedures (Zulnaidi & Zamri, 2017). In CBML, students are motivated to learn because of technology's wide and interesting content (Zulnaidi et al., 2020). CBML implementation produces concept organizing principles which are then used to think about various possible solutions (Santos-Trigo & Reyes-Rodriguez, 2016).

Furthermore, from the analysis of categorical variables, as illustrated in Table 2, it can be seen that there are categorical variables in this study that clarify the effect size of the study. First, the analysis results show that the difference in sample size affects the study's effect size. It appears that a small sample size (30 or less) gives a larger effect size than the other categories. These results directly contribute to future classroom arrangements. The
literature also supports this result (e.g., Juandi, Kusumah, Tamur, Perbowo, Siagian, et al., 2021; Juandi, Kusumah, Tamur, Perbowo, & Wijaya, 2021) that small sample size is defined as 30 or less great effect than in the other categories.

The analysis results also show that differences in educational level affect the effect size of CBML on SMA. The findings in this study illustrate that the implementation of CBML is more effective in JHS than in SHS and College. This result is surprising because it differs from previous studies' findings (e.g., Juandi, Kusumah, Tamur, Perbowo, & Wijaya, 2021). This difference may be due to the database used is not the same. However, further studies are needed to verify and clarify the consistency of the results regarding this variable.

Furthermore, Table 2 presents the analysis results related to the publication source variable. The results of the analysis show that differences in publication sources also explain the variation in effect sizes between studies. Statistically, the mean effect size of the study group from proceedings was greater than that of the journal. This indicates that this study is not associated with publication bias, as supported by the trim and fill test. In this study, there was no indication that journal editors chose to publish only significant articles (Cooper, 2017) due to the fact that the study group originating from proceedings was even more significant.

This study also considers country differences as a categorical variable. In the literature, six studies were found from abroad that examined CBML for SMA, and the rest were from Indonesia. There is a significant difference in the mean effect size between categories. This result is not surprising because the implementation of CBML abroad has been going on for a long time, as Sugandi and Delice (2014) reported. These results were influenced by the Hawthorne effect, as investigated by Tamur, Kusumah, Juandi, Wijaya, et al. (2021) that students feel enthusiastic about learning because of the novelty of the treatment. If the same treatment is given continuously, the effect will decrease. However, this difference in results is interesting for further investigation through further research.

Finally, given the importance of implementing CBML, teaching teachers how to manage technology-enabled learning (Dockendorff & Solar, 2018). Continuous training for teachers is needed to integrate technology effectively (Kartal & Çınar, 2022). These results have implications for the importance of mathematics teacher training and its impact on developing technological pedagogical content knowledge.

In this study, many studies were excluded because the categories were not measured. Hartatiana et al. (2017) for example conducted an experiment regarding the effectiveness of CBML by only considering gender differences. Although there is statistical information that supports the effect size transformation, it does not lead to the purpose of this study. Thus, further studies that specifically consider gender differences as a categorical variable need to be carried out.

Furthermore, of the 193 identified studies, only 28 met the inclusion requirements. Of the number that do not meet the analysis requirements, it examines more about analyzing student abilities related to CBML implementation (e.g., Mendrek et al., 2018), development studies (e.g., Khalil et al., 2018) and review research (e.g., Juandi, Kusumah, Tamur, Perbowo, Siagian, et al., 2021; Tamur et al., 2020; Wang & Tahir, 2020). From the distribution of data extracted, it was identified that several studies specifically investigated differences in initial ability as a categorical variable (e.g., Kariadinata et al., 2019; Muntazhimah & Miatun, 2018), whereas more studies did not measure it. The direct implication of this condition is that in the future, it is necessary to investigate whether these categorical variables influence the effectiveness of CBML on students' mathematical performance.
4. CONCLUSION

This meta-analysis analyzed 29 independent samples from 28 primary studies. Based on the random effects model, it was found that the application of CBML had a large effect on SMA. The variation in the effect size of each primary study was moderated by the categorical variables observed in this study. However, these findings are only based on the Scopus database searched through the PoP application. The fact is that there are still many related studies that cannot be downloaded, especially publications from the IIEEE, which must go through an affiliate and be paid specifically. Further cooperation procedures are needed to obtain more related studies.

ACKNOWLEDGEMENTS

The authors would like to appreciate the technical assistance from two students of the mathematics education study program at the Catholic University of Indonesia, Santu Paulus Ruteng, who were involved as variable coders.

REFERENCES


