The Effectiveness of Physical and Technology-based Virtual Manipulatives in Enhancing Student Engagement: A Meta-analysis

Julious E. Napoles [1], Judy Ann M. Lloveras [1], Vince Rhodel T. Mangco [1], Van Baines S. Bascon [1], Jeson V. Viñas [1,2,*], Joshua M. Ayade [1,3], and Jefry E. Aransado [1,4]

[1] School of Teacher Education, Biliran Province State University, Naval, Biliran, Philippines
[2] Faculty of Languages and Literature, Philippine Normal University, Manila, Philippines
[3] Graduate School, Leyte Normal University, Tacloban, Philippines
[4] Department of Science Education, Br. Andrew Gonzales FSC College of Education (BAGCED), De la Salle University, Manila, Philippines

Email: juliousnapoles99@gmail.com (J.E.N.); judyannlloveras50@gmail.com (J.A.M.L.); vincerhodelmangco12@gmail.com (V.R.T.M.); alcarazfamily12334@gmail.com (V.B.S.B.); jeson.vinas@bipsu.edu.ph (J.V.V.); joshua.ayade@bipsu.edu.ph (J.M.A.); jefry.aransado@bipsu.edu.ph (J.E.A.)
*Corresponding author

Manuscript received February 6, 2026; revised March 10, 2026; accepted April 15, 2026; published August 21, 2026

***Abstract*—This meta-analysis investigates the effectiveness of physical and technology-based virtual manipulative instructional materials in enhancing students' engagement across elementary, secondary, and tertiary education. Drawing on 10 empirical studies published from 2020 to 2024 that yielded 15 individual effect sizes, this study synthesized the quantitative data using Hedges' *g* under a random-effects model. The overall pooled effect size was large and statistically significant, indicating substantial improvement in student engagement after using manipulatives. The included studies addressed a variety of mathematics topics, such as trigonometry, algebra, set theory, calculus, geometry, logic reasoning, and multiplication fluency. They also used diverse manipulatives: physical, technology-based, virtual, and combined approaches. Despite considerable heterogeneity, the robustness of the findings was supported by a Fail-safe N of 1564 and significant results from the equivalence test. Potential publication bias was suggested through Egger's regression and Kendall's tau tests, but the observed effect remained highly reliable. The results affirm the pedagogical value of manipulatives in promoting behavioral, cognitive, and emotional aspects in mathematics learning. The study recommends that educators and curriculum developers strategically incorporate both physical and technology-based virtual manipulatives into instructional design and encourages further research to explore moderating variables such as educational level, manipulative type, mathematics content, and intervention duration.**

***Keywords*—meta-analysis, technology-based virtual manipulatives, student engagement, mathematics education, educational technology**

I. Introduction

Student engagement in mathematics has long been a central concern in education, as many students struggle to sustain attention, actively participate, and develop meaningful connections with mathematical concepts. Prior research consistently demonstrates that engagement, encompassing behavioral, cognitive, and emotional dimensions, is strongly linked to mathematics achievement, persistence, and positive learning outcomes [1–3] and remains a critical factor in technology-enhanced learning environments [4]. Despite growing evidence that manipulative instructional materials, both physical and technology-based virtual materials, can promote student engagement by making abstract mathematical ideas more concrete and interactive [5, 6], findings across studies remain fragmented and inconsistent. Accordingly, the present study aims to determine the overall effectiveness of manipulative materials in enhancing students' engagement in mathematics. This aim is addressed through a meta-analytic approach that quantitatively synthesizes effect sizes from empirical studies to provide a more comprehensive and reliable estimate of their impact.

Manipulative instructional materials, including physical tools such as counters, algebra tiles, and base-ten blocks, as well as technology-based virtual manipulatives such as digital simulations and interactive representations, are grounded in constructivist theories of learning. These theories emphasize that learners actively construct knowledge through interaction with materials and experiences rather than through passive reception [7, 8]. By enabling students to explore mathematical relationships through tangible or visual representations, manipulatives support sense-making, experimentation, and active participation in learning tasks, aligning with recent findings on technology-supported student engagement.

Empirical research has demonstrated that manipulatives can positively influence students' conceptual understanding, motivation, and engagement in mathematics classrooms [5, 9], with recent evidence supporting the effectiveness of blended physical and virtual manipulatives in enhancing learning outcomes [10]. Technology-based virtual manipulatives, in particular, offer additional affordances such as immediate feedback, dynamic representations, and accessibility across learning environments [6], highlighting their role in enhancing interactivity and engagement in digital learning environments [11]. However, individual studies differ substantially in terms of instructional design, participant characteristics, engagement measures, and reported outcomes, making it difficult to draw firm conclusions about their overall effectiveness. Meta-analysis provides a systematic and rigorous method for synthesizing findings across independent studies, allowing researchers to estimate

doi: 10.18178/ijiet.2026.16.8.2681

an overall effect size and assess the consistency of results across contexts [12, 13]. By statistically integrating results from multiple studies, meta-analysis provides stronger empirical evidence than individual studies and helps clarify whether the observed effects are robust and generalizable.

The present meta-analysis aims to examine the overall effectiveness of manipulative instructional materials in enhancing students' engagement in mathematics. Specifically, this study synthesizes empirical research to estimate the pooled effect size of manipulative-based instruction on student engagement outcomes. By focusing on student engagement rather than achievement alone, this study contributes to a deeper understanding of how instructional materials influence students' active involvement in mathematics learning. Specifically, this study sought to answer the following research questions:

(1) What mathematics topics or concepts are commonly addressed using manipulative instructional materials in the included studies?
(2) What types of manipulative materials are used to enhance students' engagement in mathematics?
(3) What specific manipulative materials are most frequently employed in mathematics instruction to promote student engagement?
(4) How effective are manipulative instructional materials in improving students' engagement in mathematics?

II. Materials and Methods

A. Research Design

This study employed a meta-analysis research design to examine the effectiveness of manipulative instructional materials on students' engagement in mathematics education. Meta-analysis is a statistical method that systematically synthesizes the quantitative findings of independent studies to determine the overall effect of an intervention [12]. It provides a comprehensive understanding of existing research by integrating results and accounting for variations in sample size, research design, and outcomes.

To ensure a thorough and transparent process, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework was followed in identifying, screening, and selecting relevant studies. The systematic review component critically synthesized findings from empirical studies addressing a clearly formulated research question [14], focusing on the impact of manipulative materials, both physical and technology-based virtual, on students' engagement, including behavioral, emotional, and cognitive dimensions. Relevant studies were retrieved using the Publish or Perish software, focusing on mathematics education research that examined student engagement in relation to the use of manipulative materials.

The meta-analysis followed five standard steps: (1) defining the research problem or hypothesis; (2) searching for relevant literature; (3) extracting and coding data; (4) applying statistical methods to compute effect sizes; and (5) presenting and interpreting the results [12, 15]. This approach allowed for a logical, transparent, and analytical synthesis of the quantitative findings from comparable studies [16, 17].

The goal of this study was to determine the overall effectiveness of technology-based virtual manipulative materials on mathematics student engagement, explore potential moderators such as educational level and type of manipulative material, and provide insights and recommendations for future research and instructional practice [18, 19].

B. Search Protocol

The identification of studies in this report was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol. Before conducting the search, the inclusion and exclusion criteria were clearly established. Electronic database searches were performed to ensure a comprehensive collection of relevant literature.

A series of searches was conducted using Harzing's Publish or Perish software and three meta-search engines: Google Scholar, Scopus, and Semantic Scholar. The search was limited to studies published from 2020 to 2024 to capture recent empirical evidence on the use of manipulative instructional materials in mathematics education, particularly within contemporary learning environments characterized by the increased integration of digital tools and evolving instructional practices. This timeframe also reflects the post-pandemic shift toward flexible and technology-supported learning environments, which has influenced the design and implementation of instructional strategies. Methodologically, restricting the timeframe helps ensure greater comparability of studies by focusing on research conducted under relatively similar educational contexts and conditions.

The search strategy employed a combination of keywords and Boolean operators applied to titles, abstracts, and keywords of published studies. The following search string was used:

("manipulative materials" or manipulatives or "instructional manipulatives" or "hands-on materials" or "concrete materials" or "learning manipulatives") and ("student engagement" or "learner engagement" or "classroom engagement" or "academic engagement") and (mathematics or math).

This systematic search process ensured a comprehensive and reproducible identification of relevant empirical studies aligned with the meta-analysis's objectives.

C. Inclusion and Exclusion Criteria

A meta-analysis of published studies was conducted to examine the effect of manipulative instructional materials on students' engagement in mathematics education. The titles, abstracts, and contents of the included studies were evaluated based on the 8 criteria listed below.

(1) The study explicitly utilizes manipulative instructional materials, either physical or technology-based virtual, as a primary intervention or teaching approach in a mathematics learning context.
(2) The study measures students' engagement in mathematics learning, which may include one or more of the following dimensions: behavioral, emotional, or cognitive.
(3) The study involves elementary, secondary, or tertiary students as participants.
(4) The study was published from 2020 to 2024.
(5) The study is written in English.

(6) The study employs a pretest-posttest, in which students' engagement in mathematics is measured before and after the implementation of manipulative instructional materials.
(7) The study utilizes a quantitative or mixed-methods research design, provided that the quantitative component reports sufficient statistical information (e.g., mean, standard deviation, and sample size) necessary for computing effect sizes.
(8) The full text of the article or thesis is available online.

Studies that did not meet the inclusion criteria were excluded from the analysis. To enhance methodological transparency, borderline cases (e.g., studies with unclear measures of student engagement, partial reporting of statistical data, or mixed instructional interventions) underwent a further full-text review. When necessary, decisions were based on whether the study provided sufficient quantitative data to compute effect sizes and whether manipulative materials were a central component of the intervention.

All inclusion and exclusion decisions were conducted collaboratively by the researchers, and any uncertainties or disagreements were resolved through discussion and consensus to ensure consistency in study selection. The researchers carefully screened and selected relevant articles using the Publish or Perish tool and the PRISMA search strategy, based on the set parameters.

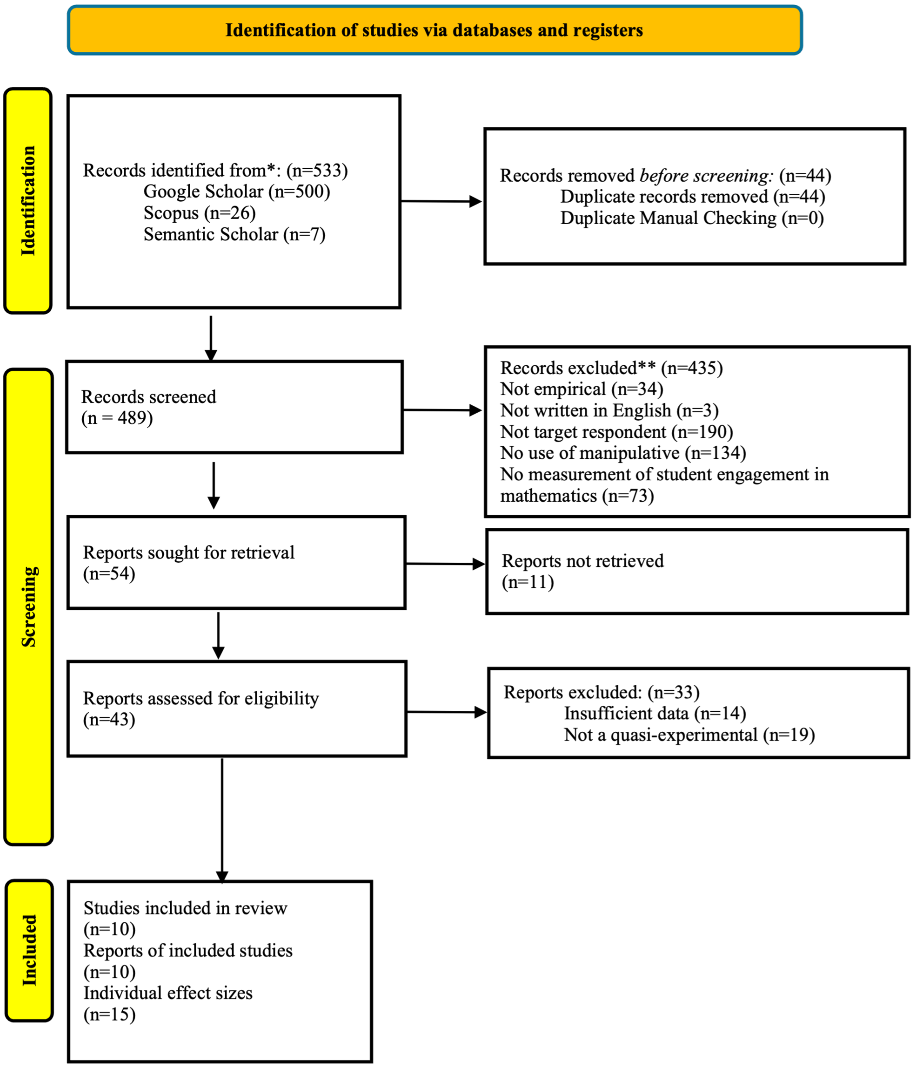

Fig. 1. Flow diagram of the literature search using PRISMA protocol.

Fig. 1 presents the PRISMA flow diagram that summarizes the systematic process used to identify, screen, assess eligibility, and include studies in the meta-analysis examining the effectiveness of manipulative materials in enhancing students' engagement in mathematics. The identification phase began with a comprehensive literature search across multiple academic databases and registers. A total of 533 records were initially identified, with the majority retrieved from Google Scholar (n = 500), followed by Scopus (n = 26) and Semantic Scholar (n = 7). To ensure the accuracy and integrity of the dataset, all records underwent duplicate screening before formal screening. As a result, 44 duplicate records were removed, yielding 489 unique records for further evaluation.

During the screening phase, the titles and abstracts of the 489 records were carefully reviewed based on predefined inclusion and exclusion criteria. This process resulted in the exclusion of 435 records that did not align with the scope or objectives of the meta-analysis. Specifically, records were excluded because they were not empirical in nature (n = 34), were not written in English (n = 3), did not involve the target population relevant to the study (n = 190), did not include the use of manipulative materials as an instructional intervention (n = 134), or failed to measure student engagement in mathematics as an outcome variable (n = 73). Following this initial screening, 55 reports were deemed potentially relevant and were sought for full-text retrieval.

In the eligibility assessment phase, full-text versions of the 54 identified reports were requested and reviewed. However, 11 reports could not be retrieved despite reasonable efforts, leaving 43 full-text articles available for detailed assessment. These 43 studies were rigorously evaluated against the meta-analysis eligibility criteria, which required the use of experimental or quasi-experimental research designs and the reporting of sufficient quantitative data necessary for effect size computation. As a result of this assessment, 33 studies were excluded. Of these, 13 were excluded for insufficient statistical data, and 19 were excluded for not employing experimental or quasi-experimental methodologies.

The final inclusion phase resulted in 10 studies that met all inclusion criteria and were included in the meta-analysis. These studies collectively provided the quantitative evidence necessary to examine the impact of manipulative materials on students' engagement in mathematics. Several of these studies reported multiple independent outcomes or participant groups, yielding a total of 15 individual effect sizes for quantitative synthesis. These multiple effect sizes were derived from distinct and non-overlapping participant groups (e.g., different schools, class sections, or achievement levels). Because the participant samples were independent, each effect size was treated as an independent estimate in the meta-analysis.

D. Coding Procedure

The 10 included studies were critically analyzed, and relevant information was systematically extracted and coded. Several studies reported results for multiple independent participant groups, resulting in a total of fifteen individual effect sizes used in the meta-analysis. These multiple effect sizes were derived from studies that reported separate outcomes for different schools, class sections, or achievement levels.

The coding process included the following variables: (a) authors and year of publication; (b) outcome measure characteristics, including pretest and posttest means, standard deviations, and sample sizes; (c) specific mathematical topics or concepts addressed; (d) type of manipulative material used physical or technology-based virtual; and (e) the specific manipulative materials utilized.

The data extraction and coding procedures were conducted collaboratively by the research team using a predefined coding template. Each member participated in extracting relevant data from the included studies, while other members independently cross-checked the extracted information to ensure accuracy and consistency. Any discrepancies identified during the cross-checking process were discussed among the researchers and resolved through consensus. All extracted data were compiled and organized using Microsoft Excel to facilitate data management and subsequent meta-analytic computations.

E. Effect Size Calculation

To determine the effectiveness of manipulative instructional materials on students' engagement in mathematics, effect sizes were computed using Hedges' g [20]. Effect sizes quantify the magnitude of an intervention's impact and provide a standardized measure that allows results from different studies to be compared and synthesized [12].

Given that the majority of the included studies employed single-group pretest-posttest designs, Hedges' g [20] was calculated as the standardized mean difference between posttest and pretest engagement scores, with a correction for small sample bias [21]. Because most primary studies did not report correlations between pretest and posttest scores, the computation relied on the available summary statistics (sample sizes, means, and standard deviations). In the absence of reported correlations, the estimation of the sampling variance assumes independence between measurements, which may slightly influence the precision of the variance estimate but generally has minimal impact on the pooled effect size in meta-analytic synthesis [21].

The computation of individual effect sizes (Hedges' g [20]) and their corresponding Standard Errors (SE) was first performed in Microsoft Excel using standard meta-analytic formulas based on the extracted means, standard deviations, and sample sizes of the included studies. These computed values were then imported into Jamovi (version 2.6.44) using the Meta-Analysis Jamovi R (MAJOR) Meta-Analysis module, specifically the Effect Sizes and (Sampling Variances or Standard Errors) analysis option.

For studies that included both experimental and control groups, only data from the experimental group were used to compute effect sizes to maintain consistency across studies, as complete control-group statistics were limited in availability.

A total of 10 studies were included in the meta-analysis. However, several studies reported multiple independent outcomes or participant groups, such as separate results for different schools, class sections, or achievement levels. As a result, 15 individual effect sizes were extracted and analyzed. Each effect size was treated as an independent estimate because the participant groups were distinct and non-overlapping.

Hedges' g [20] was selected over Cohen's d [16] because it provides a less biased estimate of effect size in studies with small to moderate sample sizes, which are common in educational research [12]. Effect sizes were interpreted using Cohen's conventional benchmarks [6, 16]: small effect (g = 0.20 – 0.49), medium effect (g = 0.50 – 0.79), large effect ($g \geq 0.80$), and negligible effect ($g < 0.20$). A positive effect size indicates an increase in students' engagement in mathematics following the use of manipulative materials.

All meta-analytic procedures were conducted using Jamovi (version 2.6.44) with the MAJOR Meta-Analysis module [22]. A random-effects model was applied to estimate the pooled effect size, accounting for variability in study characteristics such as educational level, type of manipulative material, and engagement measures. Statistical heterogeneity was assessed using the I^2 statistic, and forest plots were generated to visually present individual and pooled effect sizes.

III. Results and Discussion

This section presents the results of the meta-analysis examining the effect of manipulative instructional materials on students' engagement in mathematics. The analysis synthesizes the effect sizes derived from the included studies to determine the overall magnitude of the intervention's impact. Although student engagement is widely conceptualized as a multidimensional construct encompassing behavioral, cognitive, and emotional dimensions, the present meta-analysis treated engagement outcomes as a single aggregated construct. The included studies measured engagement using various instruments that captured one or more of these dimensions; therefore, the outcomes were synthesized uniformly to estimate the overall effect of manipulative materials on students' engagement in mathematics. The findings are presented through pooled effect sizes, heterogeneity statistics, and graphical representations such as forest plots.

Table 1. Summary of studies included for meta-analysis

Authors	Specific Mathematics Lesson or Concept	Type of Manipulative Materials	Specific Materials
Roberts *et al.* (2020) School A [23]	Trigonometry (Pythagoras, sine, cosine, tangent)	Physical	Geostrips
Roberts *et al.* (2020) School B [23]	Trigonometry (Pythagoras, sine, cosine, tangent)	Physical	Action Figures, dowel rods, styrofoam blocks, angLegs
Roberts *et al.* (2020) School C [23]	Set Theory (sets, subsets, union, intersection)	Physical	String, plastic bottle caps, styrofoam cups, balloons, etc.
Irvine (2020) [24]	Quadratic Relations	Physical & Technology-Based virtual	Algebra tiles, individual whiteboards, hands-on activity materials
Llaneza & Magulod (2023) [25]	Addition	Physical	Tactual Learning Kit (TLK)
Alias *et al.* (2024) Class 8X [26]	Solving Algebraic Linear Equations	Technology-based Virtual	Didax Virtual Algebra Tiles
Alias *et al.* (2024) Class 8Y [26]	Solving Algebraic Linear Equations	Technology-based Virtual	Didax Virtual Algebra Tiles
Obut *et al.* (2024) [27]	Concepts of Angles	Physical	Real Object
Susaie & Tan (2024) High Achievers [28]	Logic Reasoning and Problem Solving	Technology-based Virtual	Pedagogical Agent
Susaie & Tan (2024) Medium Achievers [28]	Logic Reasoning and Problem Solving	Technology-based Virtual	Pedagogical Agent
Susaie & Tan (2024) Low Achievers [28]	Logic Reasoning and Problem Solving	Technology-based Virtual	Pedagogical Agent
Agricula *et al.* (2024) [29]	Calculus	Technology-based Virtual	Technology Integrated Tools
Da *et al.* (2024) [30]	Trigonometry	Technology-based Virtual	Geogebra-Game-RME
Ahmad (2024) [31]	Geometry	Physical and Technology-based Virtual	Base-Ten Blocks, Cuisenaire rods, fraction circlesz fraction squaresz etc.
Holland (2022) [32]	Multiplication Fluency	Physical	Counters, arrays, equal-group models, number lines, drawings, visual models

Table 1 summarizes the 15 effect sizes derived from 10 studies included in the meta-analysis. These studies investigated the effectiveness of manipulative instructional materials across a variety of mathematics topics, including trigonometry, set theory, quadratic relations, algebra, geometry, calculus, logic reasoning, and multiplication fluency. Several studies reported multiple independent outcomes across different schools, classes, or ability groups, contributing a total of 15 effect sizes to the analysis.

Manipulative materials used across studies varied in type and format. Physical manipulatives included GeoStrips, counters, base-ten blocks, Cuisenaire rods, fraction circles, action figures, and real objects. In contrast, technology-based virtual manipulatives included Didax Virtual Algebra Tiles, GeoGebra-based tools, and pedagogical agent software. Some studies combined both physical and technology-based virtual materials, such as algebra tiles paired with individual whiteboards, or base-ten blocks alongside fraction tools.

These materials were used in diverse mathematics lessons, ranging from basic arithmetic operations to complex concepts in trigonometry and calculus, as well as problem-solving activities. The studies encompassed different educational levels, with participants drawn from elementary, secondary, and tertiary contexts. The variety of manipulative types, mathematics topics, and participant groups demonstrates the versatility and broad applicability of manipulative instructional materials in enhancing student engagement in mathematics across multiple settings.

Table 2 presents the results of the random-effects meta-analysis of 15 effect sizes from 10 studies examining the impact of manipulative instructional materials on students' engagement in mathematics. Using a Restricted Maximum-Likelihood (REML) estimator for the between-study variance (τ^2), the overall effect size (intercept) was estimated at Hedges' $g = 1.680$ with a standard error of 0.437.

Table 2. Random-effects model ($k = 15$)

Model Term	Estimate	SE	Z	p	CI Lower Bound	CI Upper Bound
Intercept	1.680	0.437	3.830	<0.001	0.819	2.531

Note. Tau² Estimator: Restricted Maximum-Likelihood

This effect size was statistically significant ($Z = 3.830$, $p < 0.001$), indicating a large and positive effect of manipulative materials on student engagement. The 95% confidence interval ranged from 0.819 to 2.531, further supporting the reliability of this finding. According to conventional benchmarks for Hedges' g [20], an effect size above 0.80 is considered large [33], suggesting that manipulative materials substantially enhance student engagement in mathematics learning across the included studies.

It is important to note that the included studies primarily examined short-term interventions, with outcomes measured immediately after the implementation of manipulative instructional materials. As such, the findings of this meta- analysis reflect the immediate effects on students' engagement but do not provide sufficient evidence regarding the sustainability of these effects over time.

The use of a random-effects model accounts for potential variability among studies due to differences in participants, educational levels, mathematics topics, and types of manipulative materials. This approach ensures that the overall effect size represents the average effect across diverse study contexts, rather than assuming a single common effect.

The results provide robust evidence that manipulative instructional materials, whether physical, technology-based virtual, or a combination of both, are highly effective in promoting students' engagement in mathematics.

Table 3. Heterogeneity statistics

Tau	Tau²	I^2	H^2	*df*	*Q*	*p*
1.652	2.730 (SE = 1.082)	97.210%	35.798	14.000	193.625	<0.001

Table 3 presents the heterogeneity statistics for the 15 effect sizes included in the meta-analysis. The between-study variance was substantial, with $\tau^2 = 2.730$ (SE = 1.082) and a corresponding $\tau = 1.65$, indicating notable variability in effect sizes across studies. The I^2 statistic was 97.21%, suggesting that nearly all of the observed variability among effect sizes is due to true differences between studies rather than sampling error. Similarly, $H^2 = 35.80$ further confirms considerable heterogeneity among the included studies.

The Cochran's Q test for heterogeneity was statistically significant ($Q = 193.63$, $df = 14$, $p < 0.001$), providing additional evidence that the effect sizes were not homogeneous. This high level of heterogeneity is likely attributable to differences in participant characteristics, educational levels, mathematics topics, study designs, and types of manipulative materials employed across studies.

Although subgroup analyses, meta-regression, and sensitivity analyses are commonly used to explore sources of heterogeneity, these were not conducted due to the limited number of studies within comparable categories. Specifically, the included studies varied widely in the types of manipulative materials and outcome measures, resulting in insufficient numbers of studies per subgroup to support meaningful or statistically reliable comparisons. Conducting subgroup analyses under such conditions may produce unstable estimates and misleading interpretations. Similarly, meta-regression was not feasible given the small number of effect sizes ($k = 15$), which is below the recommended threshold for reliable model estimation.

As a result, the pooled effect size should be interpreted with caution, as the substantial heterogeneity indicates variability in the magnitude of the intervention's effects across studies.

Given the substantial heterogeneity, a random-effects model was appropriate, as it assumes that the true effect size varies across studies and yields a more general estimate of the intervention's effectiveness. These findings highlight the importance of considering study-level differences when interpreting the overall results.

The forest plot displays the individual effect sizes and corresponding 95% confidence intervals for the 15 effect sizes derived from the 10 studies included in the meta-analysis. All studies demonstrated positive effect sizes, indicating that the use of manipulative instructional materials consistently favored increased student engagement in mathematics, although the magnitude of the effects varied across studies.

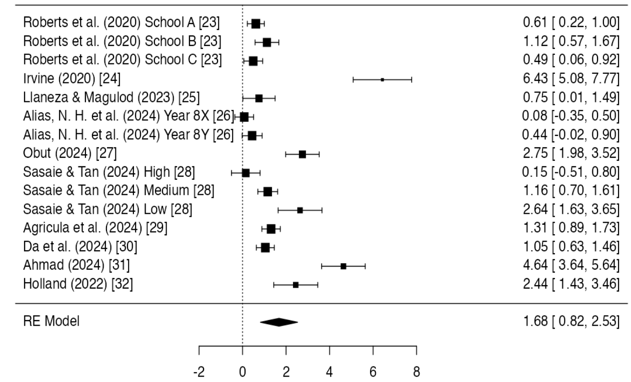

Fig. 2. Individual effect sizes and forest plot.

As shown in Fig. 2, the overall pooled effect size, represented by the diamond at the bottom of the plot, was large and statistically significant ($g = 1.68$, 95% CI [0.82, 2.53]). The width of the diamond reflects the precision of the pooled estimate under the random-effects model, accounting for variability among studies. The consistent direction of positive effects across all individual studies visually reinforces the conclusion that manipulative materials, whether physical, technology-based virtual, or combined, have a strong and favorable impact on students' engagement in mathematics. Despite differences in effect size magnitude, the forest plot provides clear visual evidence supporting the effectiveness of manipulative instructional materials across diverse educational contexts.

Table 4. Publication bias assessment

Test Name	value	*p*
Fail-Safe N	1564.000	<0.001
Kendalls Tau	0.429	0.027
Egger's Regression	6.107	<0.001

Note: Fail-safe N calculation using the Rosenthal approach.

To examine the potential presence of publication bias, several statistical procedures were conducted. As shown in Table 4, the Rosenthal Fail-safe N was 1564 ($p < 0.001$), indicating that 1564 unpublished or null-effect studies would be required to reduce the overall meta-analytic effect to statistical non-significance. This large Fail-safe N suggests that the overall effect size is robust and unlikely to be solely attributable to publication bias.

However, additional tests provided evidence of possible asymmetry. Kendall's [34] tau rank correlation test yielded a significant coefficient ($\tau = 0.429$, $p = 0.027$), indicating a significant association between effect sizes and their standard errors. This result suggests potential funnel plot asymmetry. Consistent with this finding, Egger's regression test was also statistically significant ($z = 6.107$, $p < 0.001$), further indicating asymmetry in the distribution of effect sizes.

Taken together, while the large Fail-safe N supports the stability and strength of the overall effect, the significant results from Kendall's tau [34] and Egger's [35] regression suggest that publication bias may be present. This potential bias may lead to an overestimation of the pooled effect size, as studies reporting larger or statistically significant effects are more likely to be published and included in the analysis. At the same time, smaller or non-significant findings may be underrepresented.

Although a trim-and-fill analysis could be conducted to estimate an adjusted effect size that accounts for missing studies, this procedure was not performed in the present study due to the limited number of included effect sizes, which may compromise the stability of the imputation. Nevertheless, the possibility of publication bias should be considered when interpreting the findings.

To minimize publication bias in future meta-analyses, it is recommended to include grey literature (e.g., theses, dissertations, and conference proceedings), conduct more comprehensive database searches, and apply broader inclusion criteria where appropriate.

Visual inspection of the funnel plot shown in Fig. 3 indicates noticeable asymmetry around the pooled effect size. The distribution of effect sizes shows a clustering of studies with larger effects and a relative scarcity of smaller or less precise studies in regions where they would typically be expected under conditions of minimal publication bias. This pattern suggests that studies reporting weaker or non-significant effects may be underrepresented.

The results of the statistical tests for publication bias support the observed asymmetry. Kendall's tau revealed a significant association between effect sizes and their standard errors ($\tau = 0.429$, $p = 0.027$), indicating potential small-study effects. In addition, Egger's regression test showed significant funnel plot asymmetry (intercept = 6.107, $p < 0.001$), further suggesting the presence of publication bias.

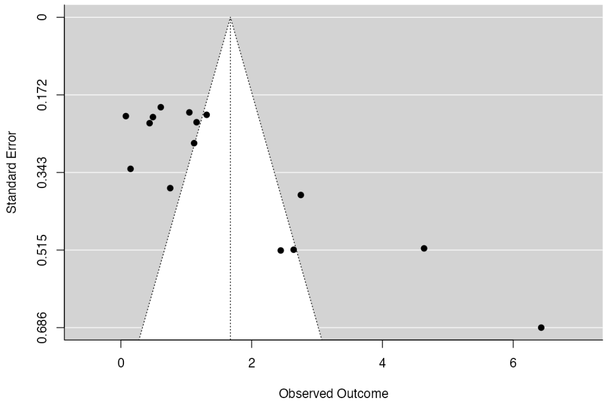

Fig. 3. Funnel plot.

Table 5. Two One-Sided Tests: Equivalence testing

Z-Value Lower Bound	*p*-Value Lower Bound	*Z*-Value Upper Bound	*p*-Value Upper Bound	LL_CI_TOST	UL_CI_TOST	LL_CI_ZTEST	UL_CI_ZTEST
4.978	<0.001	2.689	0.996	0.956	2.394	0.819	2.531

Despite these indications, the Rosenthal Fail-safe N presented in Table 5 is substantial ($N = 1564$, $p < 0.001$), implying that a large number of unpublished studies with null effects would be required to reduce the overall effect size to non-significance. Taken together, these findings suggest that while the overall effect of manipulative materials on students' engagement in mathematics is robust, the results should be interpreted with caution due to potential publication bias.

A Two One-Sided Tests (TOST) procedure was conducted to assess whether the pooled effect size was statistically equivalent to zero within a predefined equivalence interval [22]. Results indicated that the lower bound test was statistically significant ($Z = 4.978$, $p < 0.001$), while the upper bound test was not significant ($Z = 2.689$, $p = 0.996$). The 90% confidence interval for the equivalence test ranged from 0.956 to 2.394, whereas the 95% confidence interval from the standard z-test ranged from 0.819 to 2.531.

These findings demonstrate that the pooled effect size is significantly different from zero and exceeds the lower equivalence bound, confirming a meaningful positive effect of manipulative materials on student engagement in mathematics. However, because the upper-bound equivalence test was not supported, full statistical equivalence within the specified bounds was not established. This suggests that while the effect is robust and practically meaningful, it may extend beyond the upper limit of the predefined equivalence range.

The TOST results provide strong evidence for the effectiveness of manipulative materials in enhancing students' engagement in mathematics, and indicate that the effect size is substantial rather than merely at or below a minimal practical threshold.

Equivalence testing was conducted using the Two One-Sided Tests (TOST) procedure as shown in Fig. 4, to assess whether the pooled effect size was statistically equivalent to 0 within the predefined equivalence bounds of −0.50 to 0.50, corresponding to trivial or practically negligible effects [23]. The observed pooled effect size was $g = 1.675$, which lies far outside the equivalence region.

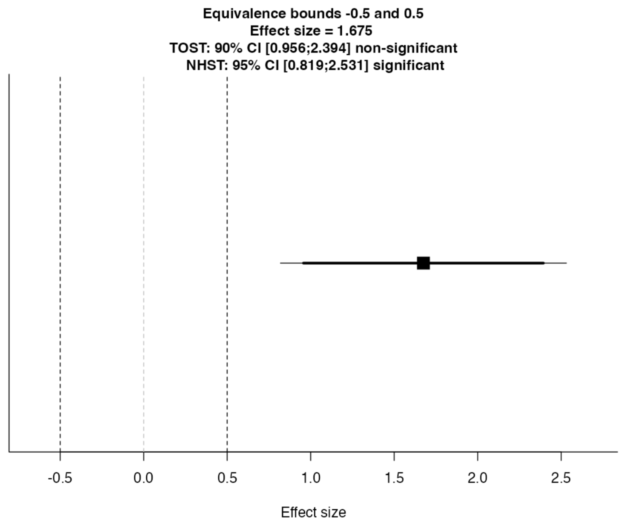

Fig. 4. Two One-Sided Tests: Equivalence testing.

The 90% confidence interval associated with the TOST ranged from 0.956 to 2.394, and both one-sided tests were statistically significant (lower-bound $z = 4.978$, $p < 0.001$; upper-bound $z = 2.689$, $p = 0.996$), indicating that the effect size is not statistically equivalent to a trivial effect. In contrast, the traditional Null Hypothesis Significance Test (NHST) yielded a statistically significant result, with a 95% confidence interval of 0.819 to 2.531 that does not include zero.

The equivalence test plot visually reinforces these findings, showing that while the pooled effect size is statistically significant under NHST, it lies well outside the equivalence bounds. This indicates that the observed effect is meaningfully large rather than practically negligible. Collectively, these results demonstrate that manipulative materials have a statistically significant and practically important impact on students' engagement in mathematics, and that the observed effect cannot be attributed to trivial or inconsequential differences.

This meta-analysis examined the effectiveness of manipulative instructional materials on students' engagement in mathematics and explored the types of manipulatives commonly used in prior research. The aggregated result (Hedges' g = 1.68, 95% CI [0.82, 2.53], $p < 0.001$) indicates a large and statistically significant effect, demonstrating that manipulative materials substantially enhance student engagement across various mathematics lessons and educational levels [22, 26, 28]. These findings provide strong empirical support for integrating hands-on and technology-based virtual manipulatives as effective instructional strategies, consistent with recent findings that blended manipulatives improve student engagement and achievement [10].

The included studies covered a wide range of mathematics topics, including trigonometry, set theory, quadratic relations, algebra, geometry, calculus, logic reasoning, and multiplication fluency. Physical manipulatives included GeoStrips, base-ten blocks, fraction circles, counters, and real objects, while technology-based virtual manipulatives included Didax Virtual Algebra Tiles, Geogebra-based tools, and pedagogical agents [23, 25, 27, 30, 31]. Some studies utilized both physical and technology-based virtual tools [24, 31]. This diversity illustrates the versatility of manipulatives in engaging students across different mathematical domains and classroom contexts, from elementary to tertiary levels.

The observed effectiveness of manipulatives aligns with constructivist learning theories emphasizing active knowledge construction and experiential learning [7, 8]. Manipulatives allow learners to visualize abstract concepts, explore relationships, and engage in problem-solving, fostering student engagement in behavioral, cognitive, and emotional domains [5, 6]. The large effect size suggests that active, hands-on learning strategies can meaningfully increase student motivation, focus, and participation during mathematics instruction [23, 26, 29].

Substantial heterogeneity was observed (I^2 = 97.21%), indicating considerable variability in effect sizes across studies. Differences in manipulative types (physical, technology-based virtual, or mixed), mathematics topics, educational levels, instructional design, intervention duration, and student characteristics likely contributed to this variation. These findings highlight the importance of examining potential moderating variables, such as manipulative type and educational level, to identify conditions under which manipulatives are most effective [28, 32].

The publication bias assessment suggested some asymmetry, with significant results from Kendall's tau and Egger's regression, indicating possible underrepresentation of smaller or null-effect studies. Nevertheless, the Fail-safe N of 1564 [36] demonstrates the robustness of the overall effect, showing that a large number of null studies would be required to negate the observed effect. Equivalence testing further confirmed that the pooled effect size is both statistically significant and practically meaningful, supporting the reliability and relevance of the findings.

This meta-analysis provides strong evidence that manipulative instructional materials are highly effective in enhancing student engagement in mathematics across various topics and educational levels. Educators and curriculum developers are encouraged to incorporate both physical and technology-based virtual manipulatives strategically, considering the mathematics content, students' abilities, and classroom context. Future research should explore long-term effects, additional moderators, and interactions with specific mathematics topics to strengthen the evidence base further. Overall, manipulatives represent a versatile, evidence-based approach to promoting active, meaningful, and sustained student engagement in mathematics learning.

IV. Conclusion

This meta-analysis examined the effectiveness of manipulative instructional materials in enhancing students' engagement in mathematics. Drawing from 10 empirical studies, which yielded 15 individual effect sizes, the findings revealed a large and statistically significant effect size (Hedges' g = 1.68, 95% CI [0.82, 2.53], $p < 0.001$), highlighting manipulative materials as a highly effective instructional strategy for promoting student engagement across different mathematical topics and educational levels. The included studies spanned a variety of mathematics lessons, manipulative types (physical, technology-based virtual, or combined), and classroom contexts, demonstrating that manipulatives are not only effective but also versatile and adaptable in diverse educational settings.

Despite the overall positive effect, substantial heterogeneity (I^2 = 97.21%) suggests that the effectiveness of manipulatives may be influenced by factors such as the type of manipulative, student ability, instructional design, intervention duration, and classroom context. While publication bias assessments indicated some asymmetry, the Fail-safe N of 1564 confirms the robustness of the observed effect, indicating that the positive outcomes are unlikely to be undermined by unpublished null results.

This meta-analysis provides compelling evidence that manipulatives, both physical and technology-based virtual, enhance students' engagement in mathematics. The findings underscore the importance of incorporating manipulatives thoughtfully into mathematics instruction to foster active learning and meaningful participation. Educators and curriculum designers are encouraged to consider the type of manipulatives, lesson content, and student characteristics when planning interventions.

Future research should explore long-term impacts, additional moderating factors, and the interactions between manipulatives and specific mathematical topics to further strengthen evidence-based instructional practices. Considering the findings of this meta-analysis, several recommendations are proposed. Teachers are encouraged to integrate manipulative materials, both physical and technology-based virtual, into mathematics instruction to enhance student engagement. Effective implementation depends on the careful selection of manipulatives appropriate to the lesson content, students' ages, and ability levels, ensuring that learning experiences are both meaningful and engaging.

Professional development opportunities should focus on strategies for using manipulatives effectively, including guidance on combining physical and technology-based virtual tools, designing interactive lessons, and fostering student-centered learning. Curriculum developers and school leaders are advised to embed manipulatives into instructional design, providing access to both physical and digital resources. Support for infrastructure, instructional materials, and training is particularly important in schools with limited resources to ensure equitable and effective implementation of manipulative-based instruction.

For researchers, further studies are recommended to examine potential moderating variables such as educational level, type of manipulatives, mathematics content, and intervention duration. Longitudinal and mixed-methods research designs could provide deeper insights into how manipulatives influence student engagement over time and identify factors that enhance or constrain their effectiveness in real-world classrooms. Collectively, these recommendations aim to strengthen the evidence-based application of manipulatives in mathematics education.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

J.E.A. led the conceptualization and design of the study, performed the formal data analysis, and oversaw the writing, reviewing, and editing of the final manuscript. Authors J.E.N., J.A.M.L., V.R.T.M., and V.B.S.B. contributed to the study's conceptualization, conducted the literature screening, and participated in the writing of the original draft. Authors J.V.V. and J.M.A. assisted in the final selection of articles, contributed to the writing of the results section, and helped review and edit the manuscript. All authors have read and approved the final version of the manuscript for submission.

Acknowledgment

The researchers sincerely thank God for the guidance and blessings throughout this study. They also extend their appreciation to their affiliated organization for the support and resources provided. The researchers likewise acknowledge the authors of the studies included in this meta-analysis for providing the basis of this research. Finally, they thank their families and peers for their continued support and encouragement.

References

[1] J. A. Fredricks, M. Filsecker, and M. A. Lawson, "Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues," *Learn. Instr.*, vol. 43, pp. 1–7, 2016. doi: 10.1016/j.learninstruc.2015.12.002

[2] J. Reeve, "How students create motivationally supportive learning environments for themselves: The concept of agentic engagement," *J. Educ. Psychol.*, vol. 105, no. 3, pp. 579–595, 2013. doi: 10.1037/a0032690

[3] E. A. Skinner and J. R. Pitzer, "Developmental dynamics of student engagement, coping, and everyday resilience," in *Handbook of Research on Student Engagement*, S. L. Christenson *et al.*, Eds. Boston, MA, USA: Springer, 2012, pp. 21–44. doi: 10.1007/978-1-4614-2018-7_2

[4] M. Bond, S. Bedenlier, K. Buntins, and M. Kerres, "Facilitating student engagement through educational technology in higher education: A systematic review in the field of arts and humanities," *Australasian Journal of Educational Technology*, vol. 36, no. 2, pp. 126–150, 2020. doi: 10.14742/ajet. 5477

[5] K. J. Carbonneau, S. C. Marley, and J. P. Selig, "A meta-analysis of the effectiveness of teaching mathematics with concrete manipulatives," *J. Educ. Psychol.*, vol. 105, no. 2, pp. 380–400, 2013. doi: 10.1037/a0031084

[6] P. S. Moyer-Packenham and A. Westenskow, "Effects of virtual manipulatives on student achievement and mathematics learning," *Int. J. Virtual Pers. Learn. Environ.*, vol. 4, no. 3, pp. 35–50, 2013. doi: 10.4018/jvple.2013070103

[7] J. Piaget, *Science of Education and the Psychology of the Child*, New York, NY, USA: Orion Press, 1970.

[8] L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*, Cambridge, MA, USA: Harvard University Press, 1978.

[9] J. Sarama and D. H. Clements, *Early Childhood Mathematics Education Research: Learning Trajectories for Young Children*, New York, NY, USA: Routledge, 2009. doi: 10.4324/9780203883785

[10] H.-S. Siller, "The effect of concrete and virtual manipulative blended instruction on mathematical achievement for elementary school students," *Canadian Journal of Science, Mathematics and Technology Education*, vol. 24, pp. 229–266, 2024. doi: 10.1007/s42330-024-00336-y

[11] J. Chen, Z. Fu, H. Liu, and J. Wang, "Effectiveness of virtual reality on learning engagement," *International Journal of Web-Based Learning and Teaching Technologies*, vol. 19, no. 1, pp. 1–14, 2024. doi: 10.4018/IJWLTT.334849

[12] M. Borenstein, L. V. Hedges, J. P. T. Higgins, and H. R. Rothstein, *Introduction to Meta-Analysis*, Chichester, U.K.: Wiley, 2011. doi: 10.1002/9780470743386

[13] M. W. Lipsey and D. B. Wilson, *Practical Meta-Analysis*, Thousand Oaks, CA, USA: SAGE Publications, 2001.

[14] J. P. T. Higgins, J. Thomas, J. Chandler *et al.*, *Cochrane Handbook for Systematic Reviews of Interventions*, 2nd ed. Chichester, U.K.: Wiley, 2019. doi: 10.1002/9781119536604

[15] F. L. Schmidt and J. E. Hunter, *Methods of Meta-analysis: Correcting Error and Bias in Research Findings*, 3rd ed. Thousand Oaks, CA, USA: SAGE Publications, 2015.

[16] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1988.

[17] D. Gough, S. Oliver, and J. Thomas, *An Introduction to Systematic Reviews*, 2nd ed. London, U.K.: SAGE Publications, 2017.

[18] L. Cohen, L. Manion, and K. Morrison, *Research Methods in Education*, 6th ed. London, U.K.: Routledge, 2007. doi: 10.4324/9780203029053

[19] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed. Thousand Oaks, CA, USA: SAGE Publications, 2013.

[20] L. V. Hedges and I. Olkin, *Statistical Methods for Meta-Analysis*, Orlando, FL, USA: Academic Press, 1985. doi: 10.1016/C2009-0-03396-0

[21] S. B. Morris and R. P. DeShon, "Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs," *Psychol. Methods*, vol. 7, no. 1, pp. 105–125, 2002. doi: 10.1037/1082-989X.7.1.105

[22] The jamovi project. (2024). *jamovi (Version 2.6)*. [Online]. Available: https://www.jamovi.org

[23] J. Roberts, S. Phipps, D. Subeeksingh, S. J. Jaggernauth, N. R. Jodha, and Z. Dedovets, "Reflection on the effects of concrete mathematics manipulatives on student engagement and problem solving in three secondary schools in Trinidad and Tobago," *Caribbean Curriculum*, 2020.

[24] J. Irvine, "Positively influencing student engagement and attitude in mathematics through an instructional intervention using reform mathematics principles," *Educ. Res. Int.*, pp. 48–75, 2020. doi: 10.5539/jel.v9n2p48

[25] P. D. R. Llaneza and G. C. Magulod Jr., "Abot kamay and pagsasanay: Effectiveness of tactual learning kit as a localized manipulative instructional material in Grade 1 mathematics," *J. Soc. Sci. Humanit.*, pp. 1–13, 2023.

[26] N. H. Alias, M. Shahrill, F. Nurhasanah, and M. Adnan, "The effects of virtual algebra tiles on the performance of Year 8 students in solving algebraic linear equations," *Int. J. Math. Math. Educ.*, vol. 2, no. 3, pp. 175–192, 2024. doi: 10.56855/ijmme.v2i3.1067

[27] L. B. Obut, L. B. Denopol, and G. R. Baluyos, "The effectiveness of manipulating real objects in enhancing retention among Grade 7

students struggling with concepts of angles," *Int. J. Innov. Sci. Res. Technol.*, 2024. doi: 10.38124/ijisrt/ijisrt24sep528
[28] C. Susaie and C. K. Tan, "Examining the impact of pedagogical agents on student performance in mathematics," *Pak. J. Learn. Soc. Sci.*, 2024. doi: 10.57239/PJLSS-2024-22.2.00569
[29] J. Agricula, C. Maderazo, R. Abejuela *et al.*, "Effects of technology-assisted instruction on the level of performance and engagement of BSEd mathematics students," in *SHS Web of Conferences*, 2024. doi: 10.1051/shsconf/202418202002
[30] N. T. Da, D. D. Anh, P. T. C. Giang, P. P. Dung, and D. B. K. Linh, "Teaching trigonometry based on realistic mathematics education and game-based learning approach," *J. Instr. Math.*, vol. 5, no. 2, pp. 168–179, 2024. doi: 10.37640/jim.v5i2.2173
[31] S. Ahmad, "An instructional model utilizing hands-on learning, concrete and virtual manipulatives on fifth-grade students' academic achievement in mathematics," M.S. thesis, University of Würzburg, Würzburg, Germany, 2024. https://opus.bibliothek.uni-wuerzburg.de/opus4-wuerzburg/frontdoor/deliver/index/docId/37524/file/Ahmad_Sagheer_Thesis.pdf
[32] W. R. Holland, "Using manipulatives and visual models to support students' multiplication fluency, attitudes, and retention," Ph.D. dissertation, Trevecca Nazarene University, 2022.
[33] D. Lakens, "Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses," *Soc. Psychol. Personal. Sci.*, vol. 8, no. 4, pp. 355–362, 2017. doi: 10.1177/1948550617697177
[34] M. Egger, G. Davey Smith, M. Schneider, and C. Minder, "Bias in meta-analysis detected by a simple, graphical test," *BMJ*, vol. 315, no. 7109, pp. 629–634, 1997. doi: 10.1136/bmj.315.7109.629
[35] C. B. Begg and M. Mazumdar, "Operating characteristics of a rank correlation test for publication bias," *Biometrics*, vol. 50, no. 4, pp. 1088–1101, 1994. doi: 10.2307/2533446
[36] R. Rosenthal, "The 'file drawer problem' and tolerance for null results," *Psychol. Bull.*, vol. 86, no. 3, pp. 638–641, 1979. doi: 10.1037/0033-2909.86.3.638