A meta-analysis of the last decade STEM implementation: what to learn and where to go

M Tamur1,2, S Fedi1, E Sennen1, Marzuki1,2, A Nurjaman2,4 and S Ndiung1
1 Universitas Katolik Indonesia Santu Paulus Ruteng, Ruteng, Indonesia
2 Universitas Pendidikan Indonesia, Bandung, Indonesia
3 Institut Agama Islam Negeri Langsa, Aceh, Indonesia
4 Institut Keguruan dan Ilmu Pendidikan Siliwangi, Cimahi, Indonesia

E-mail: maximustamur@unikastpaulus.ac.id

Abstract. Comprehensive reviews of STEM (Science, Technology, Engineering, Mathematics) have not been explored much, causing many teachers not to realise the benefits of this approach for students. For that purpose, this meta-analysis study was conducted to understand the overall effect of STEM better. Empirical data were obtained from the ERIC, SAGE, SpringerLink, and Google Scholar databases. The search results found 89 articles published between 2010 and 2020. According to the inclusion criteria, 13 items were eligible for analysis. The analysis tool uses Comprehensive Meta-Analysis (CMA) software, and random effects models are estimated. The results revealed that the overall effect size of the study was 0.73 (standard error 0.14). These results indicate that the application of STEM has a moderate effect on students' academic abilities. The analysis of the mediator variables reveals that STEM implementation is effective by considering the level of education and sample size. The study's limitations and implications are discussed to provide vital information for performance and future STEM research.

1. Introduction

In recent years, STEM (Science, Technology, Engineering, Mathematics) education has highlighted the academic world. This substantial concern engages educators, stakeholders, and researchers who call for increasing student achievement and engagement [1]. Correspondingly, STEM education has faced new challenges to investigate empirical evidence supporting its effective implementation [2], [3]. STEM education is developed to educate a generation who questions, researches, produces and makes discoveries throughout the entire educational process [4]. However, due to the lack of comprehensive reviews, many teachers are not aware of the benefits of this integrative approach for students [2].

Examining the overall effect of studies among STEM domains is a possible research topic to guide and resolve such challenges. However, individual studies cannot address this issue. This gap can be closed by integrating the findings of individual studies to provide accurate and in-depth information [5]-[10]. For this reason, a meta-analysis study is needed to facilitate a better understanding of the overall effects of STEM education.
Different researchers from different countries have carried out previous meta-analyses on the overall effect of STEM (e.g., [2], [4], [11], [12]). These studies have not included differences in sample size and STEM domains as mediators that may change the study effect size. The level of study variability on learning supported by STEM education resulted in a study moderator analysis. This procedure will provide educators with accurate information about the specific conditions that need to be considered when applying STEM.

This study complements previous research by analysing the overall effect of STEM education on student academic achievement. Investigations into individual study variability were carried out by analysing the relationship between identified moderators, namely, "differences in study years," "differences in education levels," "treatment duration," and "STEM domains." The findings of this study provide comprehensive and in-depth information about STEM management and specific conditions to consider.

2. Method

2.1. Research design

This study is a meta-analysis design in which the statistical findings of individual studies on the effect of STEM education on student achievement are combined to integrate conclusions [13], [14]. Like most meta-analysis studies [7], [9], [15], this study has gone through three stages which include determining inclusion criteria, data collection and coding, and statistical analysis.

2.2. Inclusion Criteria

Based on the problem and research focus, this study has a standard of feasibility for analysis, namely; (a) study samples in the form of national and international journals; (b) whether the research on the effect of STEM on student academic achievement; (c) publication in at least the last decade, namely between January 2010 and September 2020; and (d) contains statistical data for the effect size transformation (ES).

2.3. Data Collection Process

Individual study samples on STEM education were collected from an electronic database that includes ERIC (https://eric.ed.gov/), SAGE Publishing (https://journals.sagepub.com/), and Google Scholar (HTTP://scholar.google.com/). The data obtained were filtered using the PRISMA protocol (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). This procedure starts by identifying 89 data. Then the screening stage is carried out by issuing double data. The eligibility stage was carried out by adjusting the 89 data with the inclusion criteria. This procedure resulted in 13 individual studies included in the analysis. However, there was a study involving more than one comparison group, so 19 ES was examined.

Data extraction from eligible studies was analysed using coding forms. The coding form is a research instrument developed to translate the sample data into numerical data covering the study year, education level, sample size, and STEM domain. In this study, we involved two independent researchers as coders. They have been provided with the best coding training beforehand. The level of agreement between coders used the Cappa Cohen coefficient ($\kappa$) [16]. Cohen's kappa formula is;

$$\kappa(7) = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$$

Where $Pr (a)$ represents an observable agreement, and $Pr (e)$ illustrates a coincidence agreement. A value of 0.85 or greater is predetermined to be considered high. This study's conformity level was 0.89, which reflects a substantial to almost perfect fit between coders. Thus the data in this meta-analysis are reliable.
2.4. Statistic analysis

This meta-analysis research uses ES as the unit of analysis. ES in this study reflects the magnitude of the influence of STEM application on student academic achievement. Hedges' g equation is used for ES transformation, while the interpretation uses classification [17]; that is, less than 0.2 (negligible), between 0.2 and 0.5 (small effect), between 0.5 and 0.8 (moderate effect), between 0.8, and 1.3 (large effect), and more than 1.3 (very large effect). The fact that individual studies show considerable variability, the estimation method uses a random-effect model. Calculation of $Q_b$ (Q between) and P values for heterogeneity test was performed using CMA. The null hypothesis ($h_0$), which states that all research results are the same (homogeneous), is rejected if the p-value is <0.05. Rejecting $h_0$ implies that the ES between studies differs or may not measure the same parameters [18].

3. Result and Discussion

3.1. STEM Overall Effect Analysis

The first objective of this study is to determine the magnitude of the learning effect supported by STEM. Figure 1 presents the distribution and interpretation of ES by classification [17].

![Number of Studies](image)

**Figure 1.** Interpretation of ES.

Based on figure 1, it reflects the influence of variability in STEM implementation on student academic achievement. The results of the overall analysis regarding the effects of STEM are illustrated in table 1.

| Model           | N  | Hedges’s g | Standard error | 95% Confidence Interval | Q   | P      | Decision       |
|-----------------|----|------------|----------------|-------------------------|-----|--------|----------------|
|                 |    |            |                | Lower                  |     | Upper  |                |
| Fixed-effects   | 19 | 0.48       | 0.03           | 0.39                   | 0.55| 187.83 | Reject H0     |
| Random-effects  | 19 | 0.73       | 0.14           | 0.46                   | 1.01|        |                |
Based on Table 1, the p-value was <0.05, which means that the ES of each study is different. This means that the study fits into a random-effect model, and the degree of study variability needs to be investigated. This procedure is carried out by analysing the identified mediator variables. However, publication bias should be investigated beforehand. This is done so that these meta-analysis results are not over-interpreted because of publication bias factors [18], [19]. Investigations for publication bias can be carried out by observing the study funnel plot. This study is resistant to publication bias if the ES studies are spread symmetrically [18]. ES studies are not entirely symmetrical, then the Rosental fail-safe N (FSN) statistic and the formula N / (5k + 10) [20] are used (where N is the FSN value, and k is the number of studies). This study is resistant to publication bias if the calculation result is greater than 1. Figure 2 presents a research funnel plot.

3.2. Results of Analysis of Mediator Variables
The difference in ES between individual studies reflects the high study variability. Therefore the mediator variables which are considered to alter the ES studies should be investigated [21]. The results of the analysis are illustrated in Table 2.

The analysis results revealed that the overall ES of this study was estimated at 0.73, which means that STEM application had a moderate impact on student academic achievement. These results are consistent with previous research on the effect of STEM on student achievement. Batdi et al., for example, found that the application of STEM had a moderate impact on students' academic abilities (ES = 0.65) [4]. However, the results of this study also differ greatly from several previous studies. For example, Siregar et al. reported ES = 0.24 (small effect) when they analysed the effect of STEM on student ability [12]. Meanwhile, Khoiri et al. reported ES = 0.93 (large effect) when they analysed STEM on student achievement. These findings appear to be inconsistent. However, this study has
implemented the PRISMA protocol in a high-quality meta-analysis study. This difference in results may be due to previous investigators not paying attention to the threat of publication bias and not rigorous study selection. As a result, the resulting ES can be interpreted excessively. However, further research is needed to verify this conjecture.

### Table 2. Results of Mediator Variable Analysis.

| Mediator Variable | Group                  | N  | Hedge's g | Heterogeneity | Decision |
|-------------------|------------------------|----|-----------|---------------|----------|
|                   |                        |    |           | Q            | df(Q)    | P        |
| Publication year  | 2010-2014              | 2  | 1.02      | 108.26       | 2        | 0.00     | Reject H₀ |
|                   | 2015-2017              | 7  | 0.47      |              |          |          |          |
|                   | 2018-2020              | 10 | 0.89      |              |          |          |          |
| Educational stage | Primary school        | 2  | 1.27      | 76.66        | 3        | 0.00     | Reject H₀ |
|                   | Junior high school    | 2  | 1.40      |              |          |          |          |
|                   | Senior high school    | 7  | 0.79      |              |          |          |          |
|                   | College               | 8  | 0.32      |              |          |          |          |
| Sample size       | 30 or less            | 6  | 0.83      | 12.92        | 1        | 0.00     | Reject H₀ |
|                   | 31 or more            | 13 | 0.56      |              |          |          |          |
| STEM domain       | Mathematics           | 7  | 0.41      | 9.27         | 2        | 0.01     | Reject H₀ |
|                   | science               | 9  | 0.65      |              |          |          |          |
|                   | Engineering Technology| 3  | 0.18      |              |          |          |          |

The analysis of the moderator variables in Table 2 shows that STEM is related to the four mediators. The analysis results revealed that different years of study changed the ES of the study (Q = 108.26; P <0.05). It appears that the oldest year study group had higher ES than the newest age group. This result is surprising because it differs from our previous assumption that STEM implementation year by year will have an increasingly positive effect. The novelty effect explains why students are motivated to make greater efforts simply because of the treatment [22], [23]. For more details, further studies are needed involving more samples.

The analysis results also revealed that differences in education levels could change the effect size of the studies (Q = 76.66; P <0.05). It appears that STEM is more effective when applied in primary and junior secondary schools. However, these findings are tentative in that the number of studies analysed in that group was two. However, these results are somewhat different from previous studies that analysed the same mediators [4], [12]. This may be due to differences in the number of studies analysed. To achieve a more comprehensive result, further studies are needed involving more data related to this mediator.

The sample size difference contributed to changing the studies' effect size (Q = 12.92; P <0.05). The analysis results revealed that STEM implementation was more effective in conditions where students were not more than 30 people (considered a small sample). This may be related to student-centred learning design and student-oriented activities ranging from thinking, discovering, researching, and others. It seems that this kind of learning process is not effective enough if it involves more students (for example, over 30 people). No previous meta-analyses have yet analysed this mediator. The analysis also revealed that the STEM domain also changed the studies' effect size (Q = 9.27; P <0.05). Learning supported by STEM seems to be more effective in the science domain. These results need to be verified further to include more individual studies.

### 4. Conclusion

Learning that is supported by the application of STEM has a moderate positive impact on student academic achievement. Effective application of STEM takes into account education level and sample size. The effect of novelty affects the application of STEM so that teachers must be more creative by using interactive and interesting learning. Thus the findings of this study contribute to teachers and related parties as a basis for implementing STEM in the future. However, the findings of this study are only supported by specific individual studies deemed worthy of analysis. Although there are still many
studies in the same domain, they do not contain sufficient information on ES transformation. Therefore, further research is needed to verify inconsistent findings involving more individual studies.

References

[1] MacDonald A, Danaia L and Murphy S 2020 STEM Education Across the Learning Continuum (Springer)

[2] Becker K and Park K 2011 Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students’ learning: A preliminary meta-analysis J. STEM Educ. 12 23–37

[3] Sirakaya M, Sirakaya D A and Korkamaz O 2020 The Impact of STEM Attitude and Thinking Style on Computational Thinking Determined via Structural Equation Modeling J. Sci. Educ. Technol. 29 61–572

[4] Batdi V Talan T and Semerci Ç 2019 Meta-Analytic and Analysis of STEM Education To cite this article: Meta-Analytic and Meta-Thematic Analysis of STEM Education Int. J. Educ. Math. Sci. Technol. 7 382–99

[5] Tamur M and Juandi D 2020 Effectiveness of Constructivism Based Learning Models Against Students Mathematical Creative Thinking Abilities in Indonesia: A Meta-Analysis Study Pervasive Heal. Pervasive Comput. Technol. Health. 1 pp 107–114

[6] Siddaway A P, Wood A M and Hedges L V 2019 How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses Ann. Rev. Psychol. 70 747–70

[7] Susanti N, Juandi D and Tamur M 2020 The Effect of Problem-Based Learning (PBL) Model On Mathematical Communication Skills of Junior High School Students – A Meta-Analysis Study JTAM (Jurnal Teor. dan Apl. Mat. 4 145–54

[8] Pigott T D and Polanin J R 2020 Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review Rev. Educ. Res. 90 24–46

[9] Tamur M, Juandi D and Adem A M G 2020 Realistic Mathematics Education in Indonesia and Recommendations for Future Implementation : A Meta-Analysis Study J. Teor. dan Apl. Mat. 4 17–27

[10] Tamur M, Juandi D and Kusumah YS 2020 The Effectiveness of the Application of Mathematical Software in Indonesia; A Meta-Analysis Study Int. J. Instr. 13 1–18

[11] Khouri A 2019 Meta-Analysis Study: Effect of STEM (Science Technology Engineering and Mathematics) towards Achievement Form. J. Iılm. Pendidik. MIPA 9 71–82

[12] Siragar N C, Rosli R, Maat S M and Capraro M M 2020 The Effect of Science, Technology, Engineering and Mathematics (STEM) Program on Students’ Achievement in Mathematics : A Meta-Analysis Int. Electron. J. Math. Educ. 15 1–12

[13] Schmidt F L and Hunter J E 2015 Methods of Meta-Analysis: Correcting Error and Bias in Research Findings Third 55 City Road, London: 55 City Road (London: SAGE Publications, Ltd.)

[14] Tamur M, Jehadus E, Nandi F, Mandur K and Murni V 2020 Assessing the effectiveness of the contextual teaching and learning model on students’ mathematical understanding ability: a meta-analysis study J. Phys.: Conf. Ser. 1657 012067

[15] Pigott T D 2012 Advances in Meta-Analysis (New York Dordrecht Heidelberg London: Springer)

[16] McHugh ML 2012 Lessons in biostatistics intrarater reliability : the kappa statistic Biochem. Medica 22 276–82

[17] Cohen J 1988 Statistical power analysis for the behavioural sciences 111 p 479

[18] Borenstein M, Hedges L V, Higgins J P T and Rothstein H R 2009 Introduction to Meta-Analysis A John Wiley and Sons, Ltd.)

[19] Park S and Hong S 2016 The empirical review of a meta-analysis published in Korea Asia Pacific Educ. Rev. 17 313–24
[20] Mullen B, Muellerleile P and Bryant B 2001 Cumulative meta-analysis: A consideration of indicators of sufficiency and stability Personal. Soc. Psychol. Bull. 27 1450–62
[21] Arik S and Yilmaz M 2020 The effect of constructivist learning approach and active learning on environmental education: A meta-analysis study Int. Electron. J. Environ. Educ. 10 44–84
[22] Bayraktar S 2001 A Meta-analysis of the Effectiveness of Computer-Assisted Instruction in Science Education J. Res. Technol. Educ. 34 173–188
[23] Nurjanah N, Latif B, Yuliardi R and Tamur M 2020 Computer-assisted learning using the Cabri 3D for improving spatial ability and self-regulated learning Heliyon 6 p e05536.