

The Effect of Science, Technology, Engineering and Mathematics (STEM) Program on Students' Achievement in Mathematics: A Meta-Analysis

Nur Choירו Siregar¹, Roslinda Rosli^{1*}, Siti Mistima Maat¹, Mary Margaret Capraro²

¹ Center of Teaching and Learning Innovations, Faculty of Education, The National University of Malaysia, MALAYSIA

² Department of Teaching and Learning, Texas A&M University, USA

* CORRESPONDENCE: ✉ roslinda@ukm.edu.my

ABSTRACT

The positive impact of Science, Technology, Engineering, and Mathematics (STEM) programs on student achievement, attitude, interest, communication skills and problem-solving has alerted the education community to reform instructional approaches in STEM subjects. This meta-analysis study aims to analyze previous studies' results of STEM program impacting students' mathematics achievement. The criteria for inclusion of literature in the meta-analysis were: published between 1998 and 2017, employed experimental research design, and reported data necessary for computing effect sizes. Based on the Hedges-g effect size values, three of the studies were categorized into the large effect category (n=3), two studies were classified under the medium effect group (n=2), and twelve studies were considered to be small effect sizes (n=12). The overall weighted average effect size was 0.242 with a corresponding p-value of 0.023 demonstrating that STEM had an impact on student mathematics achievement. However, no evidence was found concerning the impact of STEM program on students' mathematics achievement based on the three moderator variables: education level, publication source, and length of intervention. Limitations and implications of the study were discussed.

Keywords: STEM education, mathematics achievement, meta-analysis, synthesis, intervention program

INTRODUCTION

During the 21st century, workforce related to science, technology, engineering, and mathematics (STEM) fields has become increasingly important (Ashford, 2016; Khalil & Osman, 2017; Wilhelm, 2014). Many countries have integrated STEM education into the school curricula providing a meaningful learning environment. According to McCaslin (2015), STEM education is a critical tool for improving students' knowledge and understanding in related fields. Also, the integration of STEM project-based learning embraces the constructivism and cognitive principles in the learning process (McCaslin, 2015). This principle is believed to provide benefits for students because they learn more by actively engaging rather than merely listening and focusing on the critical thinking and understanding problems conceptually. In STEM PBL classrooms, students are accustomed to cooperative learning and discussion, learning by questioning and exploring, investigating into various tasks, and applying the knowledge, they possess (Olivarez, 2012; Olusegun, 2015; Shahali et al., 2017). Classroom environments should be focused on collaboration and exchange of ideas

Article History: Received 28 May 2019 ♦ Revised 29 June 2019 ♦ Accepted 29 June 2019

© 2020 by the authors; licensee Modestum Ltd., UK. Open Access terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>) apply. The license permits unrestricted use, distribution, and reproduction in any medium, on the condition that users give exact credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if they made any changes.

(Olusegun, 2015). These PBL activities can increase students' interest in the STEM fields that eventually produce STEM-literate citizenry (Becker & Park, 2011).

The integration of STEM activities can cultivate student thinking skills which can help students form the ability to analyze, evaluate, make conclusions and arguments correctly and logically about problems to be solved (Chia & Maat, 2018; Dwyer, Hogan, & Stewart, 2014). Tolliver (2016) stated that students need to possess useful innovation and creativity skills in finding solutions to any related problems. STEM integration can create active, creative, critical, and communicative human beings (Bahri, Suryawati, & Osman, 2014; Tolliver, 2016). During STEM activities, students are learning contextually and focusing on the applied knowledge of STEM to solve real-world problems (Berland, Steingut, & Ko, 2014). Hence, meaningful STEM activities are challenging for educators to develop and integrate for raising students' interest and eventually to boost their academic potential in STEM subjects (Shahali et al., 2015).

Improving students' academic achievement is one of the long-term goals of any educational institution (Brown, 2012). Prior research has shown the integration of STEM has a positive impact on elementary, middle, and high school student achievement (Han et al., 2016; Hansen & Gonzalez, 2014; Ing, 2013, 2014; James, 2014; Judson, 2014; McCaslin, 2015; Tolliver, 2016). For example, McCaslin's (2015) experimental studies with fourth-grade schools in Georgia has shown the effects of STEM education on student achievement on number and operation, data measurement, and analysis, geometry and algebra. The assessment showed that there were improved results in the achievement of students following STEM-based learning. Ashford (2016) discovered that the STEM after-school program increased the academic achievement of 75 third, fourth, and fifth-grade students. Similarly, Olivarez (2012) found that a STEM program had a positive impact on 176 eighth grade students' mathematics, science and reading achievement.

Several studies have reported STEM integration can improve student involvement during classroom instruction. Educators who integrate STEM into the learning process encourage students to be active learners. Meaningful activities should include all the STEM disciplines and the real world application so that students can see the connection between the content they are learning with their daily life context (Kuenzi, 2008; Osman & Saat, 2014). In fact, Thomas (2013) stated that more coordinated activities in the STEM could cultivate students' positive attitude in encouraging them to pursue further mathematics. Also, Wong and Wong (2010) claimed meaningful STEM activities do not solely improve the understanding of concepts but increase student interest in these subjects (Kutch, 2011; Lee, 2013).

The integration of STEM in school curricula aims to strengthen the ability of students to be critical thinkers and analytical problem-solvers (Nasarudin, Halim, & Zakaria, 2014) through interactive learning experiences (Fortus et al., 2005). Although STEM positively improved student achievement outcomes, there are still some negative results when integrating STEM activities that are evidenced by James' (2014) study which showed that there were no effects using the STEM approach on the mathematics achievement of seventh graders in central Tennessee. This failure might have occurred due to the teacher being less trained in the STEM approach or the lack of school support. Additionally, the study participants were not on the same mathematics level as indicated by the findings. Along the same lines, when applying a specific STEM approach in the classroom, the role of the teacher is essential (McCaslin, 2015). As teachers have a significant role in student learning, they must understand the content, prior knowledge, challenge and support for students to learn actively in building new understanding and being able to solve various mathematical problems (Walker, 2008). A well-designed STEM approach is needed to modernize classroom instruction to help teachers and students solve problems relevant to the 21st century (Wilhelm, 2014).

Previous studies utilized empirical data that focused on the STEM in schools, starting from primary to high schools (Han et al., 2016; Hansen & Gonzalez, 2014; Ing, 2013, 2014; James, 2014; Judson, 2014; McCaslin, 2015; Tolliver, 2016). All these studies integrated STEM activities in the teaching and learning process to improve student achievement in STEM subjects. One review conducted by Kulturel-Konak, D'Allegro, and Dickinson (2011) focused on gender differences based on STEM programs, but no prior research has examined the effects of the STEM approach on students' mathematics achievement statistically. Therefore, there is a need for a meta-analytic study that discusses the effect of STEM activities on students' mathematics achievement.

METHODS

The Purpose of the Study

Many scholars believe STEM programs are an appropriate curriculum at primary school to university level because of its interdisciplinary component (Han et al., 2016). Thus, it is essential for us to understand the effectiveness of STEM programs (independent variable) in improving student mathematics achievement (dependent variable). We also investigated the factors (moderator variables) that could affect the incorporation of STEM activities during the learning process of mathematics. This meta-analysis was conducted to answer the following questions:

- (a) To what extent do STEM programs affect students' achievement in mathematics?
- (b) What factors have affected the effectiveness of STEM program in learning mathematics?

The Journal Selection Criteria

One of the challenging issues in mathematics education is the integration of STEM for increasing student achievement and to cultivate a positive attitude and interest toward the STEM field (Berlin & Lee, 2005; Brown, 2012; Dejarnette, 2012; Kuenzi, 2008). Many researchers have conducted studies related to the integration of STEM in the classroom during the past 20 years. The findings showed that STEM education is an appropriate curriculum to be integrated into the 21st-century classroom instruction. Hence, this meta-analysis included studies starting from 1998 to 2017 that focused on the impact of STEM on students' mathematics achievement.

The identification of the keywords used in the search process was the first step in conducting a meta-synthesis. Keywords used were "Science, Technology, Engineering, and Mathematics," "STEM," "STEM and Achievement." The following databases were included in this meta-analysis study: a) PsycINFO (ProQuest), b) ProQuest Digital Dissertations and Theses Full text, and c) Educational Resources Information Center (ERIC, through EBSCOhost). Also, the search for studies in mathematics education was done manually through the leading journals: Teaching Children Mathematics, Educational Studies in Mathematics, Mathematical Thinking and Learning, American Educational Research Journal, School Science, and Mathematics, For Learning of Mathematics, and Journal for Research in Mathematics Education, Google Scholar. Finally, search for journals was done to locate articles which were included in the university's institutional repository.

The first stage of the search process yielded a total of 5064 articles, but many studies were not suitable for the topic. We next located 134 research sources from dissertations, journals, and proceedings related to STEM education. The last step was to examine in depth every research article thoroughly based on the inclusion and exclusion criteria. The eligibility criteria for the inclusion were (a) the research design (i.e., quasi-experiment with treatment and control groups or pretest-posttest group), b) mathematics achievement as the outcome; c) reporting the integration issues of STEM learning, and (d) statistical data for calculating effect size (e.g., frequency, average, proportion, t-test, standard deviation, coefficient).

Many articles were rejected because a) the outcome measure was not a mathematical topic (n=45; e.g., Richardson, 2016; Robinson et al., 2014), b) the studies were based on pedagogical strategies for teachers (n=38; e.g., DeBiase, 2016; Glavich, 2016), c) they used a non-experimental designs (n=21; e.g., Alumbaugh, 2015; Petersen, 2014), d) they were focused on literature reviews and framework (n=9 e.g., Becker & Park, 2011; Harper, 2010), and e) the articles did not report statistical data (n=4; e.g., Uttal, Miller, & Newcombe, 2013; Walker & Sherman, 2017). After the complete examination of articles, only 17 articles were selected for further analysis as shown in **Table 1**. Many of the studies were from the United States because STEM education has become a regular part of the U.S. school curriculum in comparison to other countries.

Table 1. Studies included in the Meta-Analysis

Study (Years)	Source	N	Participants (Location)	Instrument	STEM Program
Shepherd (2016)	Doctoral Dissertation	2,071	M. Gr 7, 8 (West Tennessee, U.S.)	TCAP Achievement Test Data	STEM approach
Tolliver (2016)	Doctoral Dissertation	117	E. Gr 5 (Maryland, U.S.)	MSA MSA Reading and Math Assessment Scores	STEM education
Ashford (2016)	Doctoral Dissertation	550	E. Gr 3, 4, 5 (Southeastern U.S.)	North Carolina math and science end-of-Grade test scores and The Science Process Skills Inventory (SPSI)	Discovery Place After-School STEM Program
McCasin (2015)	Doctoral Dissertation	900	E. Gr 3, 4 (Georgia, U.S.)	The Criterion-Referenced Competency Test (CRCT)-summative assessment).	STEM education
James (2014)	Doctoral Dissertation	2,216	M. Gr 7 (Middle Tennessee, U.S.)	TCAP Achievement Test-NCE Data	Integration of STEM
Thomas (2013)	Doctoral Dissertation	1,824	E. Gr 4 (Southeastern U.S.)	Fennema-Sherman Mathematics Attitude Scale (FS MAS) and Tennessee Comprehensive Assessment Program (TCAP)-NCE	Integrated Science, Technology, Engineering, and Mathematics (STEM) program
Olivarez (2012)	Doctoral Dissertation	176	M. Gr 8 (South Texas, U.S.)	Est) test scores	STEM-Project-Based Learning (PBL), collaborative learning, and hands-on strategies
Kutch (2011)	Doctoral Dissertation	565	M. Gr 7, 8 (Wilmington, U.S.)	The Delaware State Testing Program (DSTP) results for mathematics grades 7 and 8 and NWEA measures of academic progress (MAP) in mathematics test the data	Integrated mathematics and science instructional approach
Guzey et al. (2016)	Journal of Science Education and Technology	4,450	E. Gr 4-8 (Midwest U.S.)	TCAP Achievement Test Data and engineering, mathematics assessments	Engineering integration approach /integrated science and engineering experiences
Han et al. (2016)	Journal of Turkish Science Education	1,187	High Schools (Texas, U.S.)	Est in areas: algebra, geometry, probability, and problem-solving)	(STEM) Project Based Learning (PBL) instructional approach
Ing (2014)	Journal of Career Development	3,116	M. Gr 7, 8, 9, 10, 12 (U.S.)	Basic skills, algebra, geometry, and quantitative literacy from NAEP	The actual shape of the trajectory
Judson (2014)	The Journal of Educational Research	11,800	E, Gr 2, 3, 4, 5 (Arizona, U.S.)	NCE-TerraNova assessment of data mathematics, reading, and language arts)	STEM-Focused Charter and Magnet Schools
Ing (2013)	International Journal of Science and Mathematics Education	3,116	M. Gr 7, 8, 9, 10, 11, 12 (U.S.)	Basic skills, algebra, geometry, and quantitative NAEP literacy from	the actual shape of the trajectory
Hansen & Gonzalez (2014)	American Journal of Education	100,000	M, 6, 7, 8 (North Carolina, U.S.)	The End-of-Grade (EOG) tests)	integration of technology with instruction
Melguizo & Wolniak (2011)	Research in Higher Education	4,000	High School (Chicago, U.S)	SAT score	STEM pipeline
Leuwerke et al. (2004)	Journal of Career Assessment	1,108	First Semester of College (U.S.)	Mathematics test of the ACT Assessment	Mathematics Achievement x congruence model
Nasarudin et al. (2014)	Eurasia Journal of Mathematics, Science & Technology Education	193	Primary School (Malaysia)	State metacognitive Inventory and student achievement in mathematical word problem solving- UPSR	VStops approach through MVStops modules

Note. Gr. = Grade, E = Elementary, M = Middle, U.S. = United States of America

Coding Process

The information derived from each of the studies was coded using a coding sheet developed iteratively based on the information provided. The first researcher coded the articles according to the predetermined information such as author name, publication type, publication year, the study participant, sample size, the instrument used, statistical analysis performed, and treatment type. The coding sheet was the essential instrument of the meta-analysis study for gathering information to compute the effect size and compare the results among studies. Effect size (I would put a citation here) is a calculated statistic based on quantitative information from each study and is essential for the moderator analysis in exploring the variability of the studies. The first researcher frequently discussed with the second researcher throughout the coding process and effect sizes computation. Many articles did not provide complete information; hence we contacted the authors to obtain relevant data for calculating the effect sizes.

Statistical Analysis

Results were obtained from the effect sizes derived from nine published journal articles and eight doctoral dissertations. All studies utilized the experimental design with posttest means for the treatment and control groups. Standardized mean difference effect size, Cohen's d is required to synthesize and compare the mean value of research results between these treatment and control groups (Lipsey & Wilson, 2001). The effect size is the value that can be used to compare studies that used different measurement procedures. The calculation of effect size is based on the mean and standard deviation between two contrast groups. To correct for small sample bias, d was converted to the unbiased estimator g (Hedges & Olkin 1985). The Hedges- g effect size is interpretable and a meaningful value that helps researchers in making assessments across the variety of studies.

The Comprehensive Meta-Analysis 3.0 software (CMA) was utilized to compute the Hedges- g effect sizes, standard error, variance, lower limit, upper limit, Z-value, and a p-value of each study. Besides, the CMA was used to calculate the distribution of homogeneity (Q statistics), I-squared, Tau-squared and moderator analysis that have been identified based on coding sheets (Borenstein et al., 2005). Also, in processing the data analysis, Microsoft Office Excel (Ms-Excel) version 2010 was applied to ease of computation of the prediction intervals.

In the study, we hypothesized three potential variables that might influence students' achievement such as a) education level (elementary, secondary, and university); b) publication source (dissertation and journal); and, c) length of intervention (long term, for program >1 year and short-term, for program < 1 year). The moderator analysis was performed in CMA by comparing the average effect sizes of mathematics achievement in different categories that formed the stages of the moderator. The moderator analysis utilized a mixed effects analysis using a Q test based on the analysis of Variance (ANOVA).

RESULTS

Description of Selected Studies

The present meta-analytic study reviewed 17 articles that were published between 2004-2016 (see **Table 1**). The studies were published in two countries, the U.S. ($n=16$) and Malaysia ($n=1$). The total number of participants in all studies was approximately 137,389 students. On average, studies utilized sample sizes of more than 100 students at elementary (grade 3-8), secondary (grade 6-12), and university levels. The studies covered the mathematics content areas of algebra, geometry, probability, and problem-solving. The STEM program approaches were: Discovery-STEM, STEM-PBL, and VSTops-STEM Module. All studies used the quasi-experiment research design.

In the next section, we discuss results from the first research question: What effect does STEM integration have on improving student achievement in mathematics?. In general, the effect size of the STEM programs on students' mathematics achievement was categorized into three groups of Becker (2000) criteria: 0.2 (small effect), 0.5 (medium effect) and 0.8 (large effect). Results from the analysis are displayed in **Table 2** based on the Hedges- g effect size values. The CMA results showed that two studies fall into the large effect category ($n=2$), two studies fall under the medium effect group ($n=2$), and 13 studies are considered to have small effects ($n=13$).

Table 2. The Effect Sizes of STEM Program

Study Name	Hedges-g	Standard error	Variance	Lower limit	Upper limit	Z-value	p-value
Nasarudin et al. (2014)	1.103	0.155	0.024	0.799	1.407	7.120	0.000
Ashford (2016)	0.180	0.071	0.005	0.041	0.319	2.546	0.011
Han et al. (2016)	0.214	0.089	0.008	0.039	0.389	2.393	0.017
James (2014)	0.055	0.110	0.012	-0.160	0.270	0.502	0.616
Olivarez (2012)	0.568	0.167	0.028	0.240	0.896	3.394	0.001
Thomas (2013)	-0.473	0.055	0.003	-0.580	-0.366	-8.636	0.000
Tolliver (2016)	0.374	0.187	0.035	0.007	0.741	1.999	0.046
Melguizo & Wolniak (2011)	0.004	0.063	0.004	-0.120	0.128	0.063	0.950
Ing (2014)	1.571	0.045	0.002	1.483	1.659	35.129	0.000
Leuwerke et al. (2004)	0.521	0.100	0.010	0.325	0.717	5.210	0.000
Hansen & Gonzalez (2014)	0.118	0.003	0.000	0.112	0.124	37.315	0.000
Ing (2013)	0.028	0.032	0.001	-0.034	0.090	0.885	0.376
Guzey et al. (2016)	-0.179	0.065	0.004	-0.306	-0.052	-2770	0.006
McCaslin (2015)	0.127	0.122	0.015	-0.113	0.366	1.037	0.300
Kutch (2011)	0.292	0.102	0.010	0.092	0.492	2.864	0.004
Ashford (2016)	0.153	0.049	0.002	0.056	0.250	3.094	0.002
Judson (2014)	-0.429	0.089	0.008	-0.604	-0.254	-4.796	0.000
Random	0.242	0.106	0.011	0.034	0.450	2.277	0.023

Table 2 displays the overall weighted average effect size as 0.242, which shows the mean difference between STEM and non-STEM programs. The 17 studies were sampled based on specific inclusion/exclusion criteria as outlined. The confidence intervals for the difference in means ranged from 0.034 to 0.450 indicating that the mean difference could fall anywhere in this range. This range does not include a difference of zero, which tells us that the true mean difference is probably not zero. Similarly, the Z-value for testing the null hypothesis (that the mean difference is 0) is 2.277, with a corresponding p-value of 0.023 demonstrating that the STEM program has an impact on students' mathematics achievement.

Heterogeneity of Effect Sizes

The observed effect size might vary from study to study, but a certain amount of variation is expected due to sampling error. The Q statistic is a measure of weighted squared deviation providing a test of the null hypothesis that all studies in the analysis share a common effect size wherein the expected value of Q would be equal to the degree of freedom (the number of studies minus 1). For this meta-analysis, the value of Q is 1311.583, (df=16, $p < 0.05$) where the statistical test is statistically significant indicating that the studies are heterogeneous (the true effect size varies from study to study). The value of I-squared is 98.780 demonstrating that 98% of the variance in observed effects reflects the percentage of variability attributable to true heterogeneity (i.e., over and above sampling error) with an estimate of variance of true effect sizes (in log units), Tau-squared = 0.182 and the estimate of standard deviation of 0.426 in true effect size. The prediction interval of 0.08 to 0.75 is based on the calculation from Microsoft Excel showing the STEM group experienced the effect of the program on mathematics achievement, with an impact ranging from 0.08 unit in some studies to 0.75 unit in others.

Moderator Analyses

The result of significant heterogeneity provides that other variables related to the characteristics of studies might differ either systematically or otherwise within the set of effect sizes. For the second research question: What factors are affecting the effectiveness on STEM integration in learning mathematics?, we explored further the result of the heterogeneity analysis and identified several factors such as the education level, publication source, and the length of the intervention on influencing the STEM program on mathematics achievement.

Level of education

The moderator analysis was run based on seven studies focusing on the STEM program at the elementary level, eight studies at the secondary level, and two studies at university levels. The values of Hedges-g, 95% confidence interval, Z, and p were computed through the CMA (see **Table 3**) and were described as:

Table 3. Moderator Analysis Based on Level of Education

Groups	#	Effect size and 95% confidence interval			Test of null [2-Tail]		Heterogeneity			Tau-squared						
		Point estimate	Standard error	Variance	Lower Limit	Upper limit	Z-value	p-value	Q-value	df (Q)	p-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis																
Elementary	7	-0.095	0.032	0.001	-0.158	-0.033	-3.008	0.003	135.697	6	0.000	95.578	0.169	0.125	0.016	0.411
Secondary	8	0.124	0.003	0.000	0.118	0.130	39.701	0.000	1109.058	7	0.000	99.369	0.244	0.233	0.054	0.494
University	2	0.152	0.053	0.003	0.047	0.256	2.838	0.005	19.092	1	0.000	94.762	0.127	0.189	0.036	0.356
Total within									1263.847	14	0.000					
Total between									47.736	2	0.000					
Overall	17	0.122	0.003	0.000	0.116	0.128	39.313	0.000	1311.583	16	0.000	98.780	0.182	0.150	0.023	0.426
Mixed effects analysis																
Elementary	7	0.156	0.187	0.035	-0.210	0.521	0.833	0.405								
Secondary	8	0.314	0.172	0.030	-0.024	0.651	1.820	0.069								
University	2	0.259	0.344	0.119	-0.416	0.934	0.753	0.452								
Total between									0.390	2	0.823					
Overall	17	0.243	0.119	0.014	0.010	0.476	2.046	0.041								

Note. # = Number Studies

Table 4. Moderator Analysis Based on Publication Source

Groups	#	Effect size and 95% confidence interval			Test of null [2-Tail]		Heterogeneity			Tau-squared						
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	p-value	Q-value	df (Q)	p-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis																
Dissertation	8	-0.035	0.028	0.001	-0.089	0.020	-1.256	0.209	136.855	7	0.000	94.885	0.127	0.089	0.008	0.356
Journal	9	0.124	0.003	0.000	0.118	0.130	39.702	0.000	1142.405	8	0.000	99.300	0.246	0.230	0.053	0.496
Total within									1279.260	15	0.000					
Total between									32.323	1	0.000					
Overall	17	0.122	0.003	0.000	0.116	0.128	39.313	0.000	1311.583	16	0.000	98.780	0.182	0.150	0.023	0.426
Mixed effects analysis																
Dissertation	8	0.089	0.172	0.030	-0.249	0.426	0.516	0.606								
Journal	9	0.376	0.160	0.026	0.062	0.691	2.348	0.019								
Total between									1.494	1	0.222					
Overall	17	0.243	0.117	0.014	0.013	0.473	2.070	0.038								

Note. # = Number Studies

a) Is the STEM program at the elementary level more effective?

The Hedges-g value is 0.156 with a 95% confidence interval of -0.210 to 0.521. The Z-value for a test of the null is 0.833 with $p = 0.405$.

b) Is the STEM program at the secondary level more effective?

The Hedges-g value is 0.314 with a 95% confidence interval of -0.024 to 0.651. The Z-value for a test of the null is 1.820 with $p = 0.069$.

c) Is STEM program at the university level more effective?

The Hedges-g value is 0.259 with a 95% confidence interval of -0.416 to 0.93. The Z-value for a test of the null is 0.753 with $p = 0.452$.

The subgroup analysis was run for the studies and was grouped based on the level of education. The mean effect was approximately the same for elementary, secondary and university (Hedges-g of 0.156, 0.314, and 0.259 respectively). The test to compare the three levels yielded a Q-value of 0.390 with $df = 2$ and $p = 0.823$. The analysis shows that there is no evidence that the impact of STEM programs vary by educational levels.

Publication source

Table 4 displays the results of the moderator analysis based on the publication source (eight dissertation studies and nine journal article studies). The values of Hedges-g, 95% confidence interval, Z, and p were computed through the CMA and are described below:

a) Is STEM program research published in a dissertation format more effective?

For dissertations, the mean effect size is a mean difference of 0.089 with a 95% confidence interval of -0.249 to 0.426, a Z-value of 0.516 and a corresponding p-value of 0.606.

b) Is STEM program research published in journals more effective?

Table 5. Moderator Analysis Based on Length of Intervention

Groups	#	Effect size and 95% confidence interval				Test of null [2-Tail]			Heterogeneity			Tau-squared				
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	p-value	Q-value	df (Q)	p-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed effect analysis																
Long term	12	0.122	0.003	0.000	0.116	0.128	39.112	0.000	1263.405	11	0.000	99.129	0.202	0.174	0.030	0.450
Short term	5	0.195	0.045	0.002	0.106	0.284	4.286	0.000	45.606	4	0.000	91.229	0.130	0.115	0.013	0.360
Total within												1309.011	15	0.000		
Total between												2.572	1	0.109		
Overall	17	0.122	0.003	0.000	0.116	0.128	39.313	0.000	1311.583	16	0.000	98.780	0.182	0.150	0.023	0.426
Mixed effects analysis																
Long-term	12	0.192	0.131	0.017	-0.065	0.448	1.465	0.143								
Short term	5	0.370	0.208	0.043	-0.038	0.777	1.778	0.075								
Total between												0.524	1	0.469		
Overall	17	0.242	0.111	0.012	0.025	0.459	2.188	0.029								

Note. # = Number Studies

For journals, the mean effect size is a mean difference of 0.376 with a 95% confidence interval of 0.062 to 0.691, a Z-value of 2.348 a corresponding p-value of 0.019. The p-value (<0.05) shows the STEM program published in the journal were more effective than the non-STEM.

The test to compare the two effect sizes (0.089 and 0.376) yielded a Q-value of 1.494 with one df and a corresponding p-value of 0.222. The statistical values indicated that there was no effect for STEM programs on students' achievement in mathematics based on publication source (dissertation versus journal).

Length of intervention

The analysis moderator (see **Table 5**) was based on the effect sizes that were computed on 12 studies categorized as a long-term program and five studies as a short-term program. The values of Hedges-g, 95% confidence interval, Z, and p were computed through the CMA and was described below:

a) Are long-term STEM programs more effective than long-term non-STEM programs?

The Hedges-g value was 0.192 with a 95% confidence interval of -0.065 to 0.448. The Z-value for H0 was 1.465 with a p-value of 0.143.

b) Are short-term STEM programs more effective than short-term non-STEM programs?

The Hedges g value was 0.370 with a 95% confidence interval of -0.038 to 0.777. The Z-value for H0 was 1.778 with a p-value of 0.075.

The subgroup analysis was run for studies based on the length of intervention. The mean effect was approximately the same for the long term and short term programs (Hedges-g of 0.192, and 0.370). The test to compare the two effect sizes yielded a Q-value of 0.524 with df = 1 and p = 0.469. The analysis shows that there was no evidence that the impact of STEM programs vary by the length of intervention.

DISCUSSION

This meta-analysis provides a standard measure for the outcome of each study and produces an overall weighted outcome for the studies. Studies were selected from electronic databases and a review of relevant references from studies and articles. In this present study, the results demonstrate significant outcomes regarding the effects of the independent variable - STEM program on the dependent variable - students' mathematics achievement. The overall weighted average effect size of 0.242 indicated that STEM programs are educationally important for student achievement in mathematics. When examining individual studies, the analysis shows the majority, 10 of 17 studies yielded statistically significant positive effect sizes between 0.118 and 1.571. These findings illustrate that the STEM program approach utilized in these ten studies might have improved students' achievement in mathematics in some way. Besides, another four studies showed a positive effect but did not show statistically significant results with small effect size values between 0.004 and 0.127, and the remaining three studies produced negative effect size values. Interestingly, the three moderator

variables (i.e., education level, publication source, and length of intervention) did not statistically affect STEM and non-STEM programs on students' achievement in mathematics learning.

These meta-analytic results are supportive of previous individual studies that show the effectiveness of STEM programs in improving student mathematics achievement (Han et al., 2016; Hansen & Gonzalez, 2014; Ing, 2013, 2014; James, 2014; Judson, 2014; McCaslin, 2015; Tolliver, 2016). The overall results indicate that STEM programs have a positive impact on students' mathematics achievement. The indicators might be used by teachers to make decisions about whether to integrate a STEM program and to help teachers in developing materials (McCaslin, 2015; Wilhelm, 2014). It is a critical issue for schools and teachers, particularly when considering their budgets for implementing a STEM program. The large positive effect sizes of 1.103 and 1.571 indicate reasonable standards in conducting future research, especially when deciding the effective programs, sample sizes, and research designs.

Several limitations are noted in this meta-analysis study. First, some studies did not provide relevant data for effect size computation and failed to provide detailed information about the STEM program they had conducted in their studies. Also, the results presented in the selected studies were not explicit and not systematic. Thus, data analysis was performed based on available data even after contacting the author. We have excluded studies which the extraction of the information was not possible.

Also, the findings might encourage future researchers to plan comprehensive studies in examining the effect of STEM programs on other learning outcomes for example attitude, interest, parental support, and motivation as well as different moderator variables such as time period and research design.

CONCLUSION

In summary, the results of this meta-analytic study are promising and provide an overall effect for STEM programs on students' mathematics achievement as positive and statistically significant. Policy makers and teachers should utilize this evidence in reforming instructional approaches in a classroom for improving student achievement at all levels. It is hoped that more research will be conducted to answer many unsolved questions and to enhance our understanding of the complex nature of teaching and learning of STEM in providing students with 21st-century skills.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Nur Choירו Siregar – Center of Teaching and Learning Innovations, Faculty of Education, The National University of Malaysia, Malaysia.

Roslinda Rosli – Center of Teaching and Learning Innovations, Faculty of Education, The National University of Malaysia, Malaysia.

Siti Mistima Maat – Center of Teaching and Learning Innovations, Faculty of Education, The National University of Malaysia, Malaysia.

Mary Margaret Capraro – Texas A&M University, USA.

REFERENCES

- Alumbaugh, K. M. (2015). *The perceptions of elementary STEM schools in Missouri* (Unpublished doctoral dissertation). Lindenwood University, Saint Charles, Missouri.
- Ashford, V. D. (2016). *STEM after school programming: The effect on student achievement and attitude* (Unpublished doctoral dissertation). Wingate University, Charlotte, NC.
- Bahri, N. M., Suryawati E., & Osman, K. (2014). Students' biotechnology literacy: The pillars of STEM education in Malaysia. *Eurasia Journal of Mathematics, Science and Technology Education*, 10(3), 195-207. <https://doi.org/10.12973/eurasia.2014.1074a>

- Becker, K., & Park, K. (2011). Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. *Journal of STEM Education*, 12(5), 23-38. <https://doi.org/10.1037/a0019454>
- Becker, L. A. (2000). *Effect size (ES)*. Retrieved from <https://www.uccs.edu/lbecker/effect-size>
- Berland, L., Steingut, R., & Ko, P. (2014). High school student perceptions of the utility of the engineering design process: Creating opportunities to engage in engineering practices and apply math and science content. *Journal of Science Education and Technology*, 23(6), 705-720. <https://doi.org/10.1007/s10956-014-9498-4>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2005). *Comprehensive meta-analysis version 3*. Englewood, NJ: Biostat.
- Brown, J. (2012). The current status of STEM education research. *Journal of STEM Education: Innovations & Research*, 13(5), 7-11. Retrieved from <http://search.proquest.com/docview/1266493670?accountid=16369>
- Chia, P. L., & Maat, S. M. (2018). An exploratory study of teachers' attitudes towards integration of STEM in Malaysia. *International Journal of Electrical Engineering and Applied Sciences*, 1(1), 45-50. Retrieved from <http://smk.ukm.my/erep/fail3.cfm?komp=TGG2018484>
- DeBiase, K. (2016). *Teacher preparation in science, technology, engineering, and mathematics instruction* (Unpublished doctoral dissertation). California State University, Los Angeles.
- Dejarnette, N. K. (2012). America's children: Providing early exposure to STEM (science, technology, engineering and math) initiatives. *Education*, 133(1), 77-84. Retrieved from <https://www.researchgate.net/publication/281065932>
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, 12, 43-52. <https://doi.org/10.1016/j.tsc.2013.12.004>
- Fortus, D., Krajcik, J., Dersheimer, R. C., Marx, R. W., & Mamlok-Naaman, R. (2005). Designbased science and real world problem-solving. *International Journal of Science Education*, 27(7), 855-879. <https://doi.org/10.1080/09500690500038165>
- Glavich, C. (2016). Growing strong STEMs reflections of a beginning teacher's preservice program. *Issues in Teacher Education*, 25(2), 89-102. Retrieved from <https://search.proquest.com/openview/6e23b55cb5b18d24dd6a22573003d3ef/1?pq-origsite=gscholar&cbl=28752>
- Guzey, S. S., Harwell, M., Moreno, M., Peralta, Y., & Moore, T. J. (2016). The impact of design-based STEM integration curricula on student achievement in engineering, science, and mathematics. *Journal of Science Education and Technology*, 26(2), 207-222. <https://doi.org/10.1007/s10956-016-9673-x>
- Han, S., Rosli, R., Capraro, M. M., & Capraro, R. M. (2016). The effect of science, technology, engineering and mathematics (STEM) project based learning (PBL) on students' achievement in four mathematics topics. *Journal of Turkish Science Education*, 13, 3-29. <https://doi.org/10.12973/tused.10168a>
- Hansen, M., & Gonzalez, T. (2014). Investigating the relationship between STEM learning principles and student achievement in math and science. *American Journal of Education*, 120(2), 139-171. <https://doi.org/10.1086/674376>
- Harper, S. R. (2010). An anti-deficit achievement framework for research on students of color in STEM. *New Directions for Institutional Research*, 2010(148), 63-74. <https://doi.org/10.1002/ir.362>
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Ing, M. (2013). Gender differences in the influence of early perceived parental support on student mathematics and science achievement and STEM career attainment. *International Journal of Science & Mathematics Education*, 12(5), 1221-1239. <https://doi.org/10.1007/s10763-013-9447-3>
- Ing, M. (2014). Can parents influence children's mathematics achievement and persistence in STEM careers. *Journal of Career Development*, 41(2), 87-103. <https://doi.org/10.1177/0894845313481672>
- James, J. S. (2014). *Science, technology, engineering, and mathematics (STEM) curriculum and seventh grade mathematics and science achievement* (Unpublished doctoral dissertation). Grand Canyon University, Phoenix, Arizona.
- Judson, E. (2014). Effects of transferring to STEM-focused charter and magnet schools on student achievement. *The Journal of Educational Research*, 107(4), 255-266. <https://doi.org/10.1080/00220671.2013.823367>

- Khalil, N. M., & Osman, K. (2017). STEM 21CS Module: Fostering 21st century skills through integrated STEM. *K-12 STEM Education*, 3(3), 225-233. Retrieved from <http://smk.ukm.my/erep/fail3.cfm?komp=TGG2017804>
- Kuenzi, J. J. (2008). Science, technology, engineering, and mathematics (STEM) education: Background, federal policy, and legislative action. *Congressional Research Service Reports, Paper 35*, 1-35. Retrieved from <https://digitalcommons.unl.edu/crsdocs/35/>
- Kulturel-Konak, S., D'Allegro, M. L., Dickinson, S. (2011). Riview of gender differences in learning styles: Suggetion for STEM education. *Contemporary Issues in Education Research*, 4(3), 9-18. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1072833.pdf>
- Kutch, M. (2011). *Integrating science and mathematics instruction in a middle school STEM course: The impact on attitudes, career aspirations and academic achievement in science and mathematics* (Unpublished doctoral dissertation). Wilmington University, DE.
- Lee, A. (2013). Determining the effects of pre-college STEM contexts on STEM major choices in 4-year postsecondary institutions using multilevel structural equation modeling. *Journal of Pre-College Engineering Education Research*, 3(2), 13-30. <https://doi.org/10.7771/2157-9288.1059>
- Leuwerke, W. C., Robbins, S., Sawyer, R., & Hovland, M. (2004). Predicting engineering major status from mathematics achievement and interest congruence. *Journal of Career Assessment*, 12(2), 135-149. <https://doi.org/10.1177/1069072703257756>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE.
- McCaslin, S. D. (2015). *The influence of stem initiative programs for middle and high school students on female STEM college majors* (Unpublished doctoral dissertation). Capella University, Minnesota, US.
- Melguizo, T., & Wolniak, G. C. (2011). The earnings benefits of majoring in STEM fields among high achieving minority students. *Research in Higher Education*, 53(4), 383-405. <https://doi.org/10.1007/s11162-011-9238-z>
- Nasarudin, A, Halim, L., & Zakaria, E. (2014). VStops: A thinking strategy and visual representation approach in mathematical word problem solving toward enhancing STEM literacy. *Eurasia Journal of Mathematics, Science & Technology Education*, 10(3), 165-174. <https://doi.org/10.12973/eurasia.2014.1073a>
- Olivarez, N. (2012). *The impact of a STEM program on academic achievement of eighth grade students in a south Texas middle school* (Unpublished doctoral dissertation). Texas A&M University, Corpus Christi, Texas.
- Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *IOSR Journal of Research & Method in Education Ver. I*, 5(6), 2320-7388. <https://doi.org/10.9790/7388-05616670>
- Osman, K., & Saat, R. M. (2014). Editorial. Science technology, engineering and mathematics (STEM) education in Malaysia. *Eurasia Journal of Mathematics, Science & Technology Education*, 10(3), 153-154. Retrieved from <http://smk.ukm.my/erep/fail3.cfm?komp=TGG2014397>
- Petersen, A. M. (2014). *Females and STEM: Determining the K-12 experiences that influenced women to pursue STEM fields* (Unpublished doctoral dissertation). The College of William and Mary, Williamsburg, Virginia.
- Richardson, S. S. (2016). *The effect of an integrated STEM course on middle school students' interest and career aspirations in STEM Fields* (Unpublished doctoral dissertation). The University of Kansas, Lawrence, Kansas.
- Robinson, A., Dailey, D., Hughes, G., & Cotabish, A. (2014). The effects of a science-focused STEM intervention on gifted elementary students' science knowledge and skills. *Journal of Advanced Academics*, 25(3), 189-213. <https://doi.org/10.1177/1932202X14533799>
- Shahali, E. H. M. S., Halim, L., Rasul, M. S., Osman, K., & Zulkifeli, M. A. (2017). STEM learning through engineering design, impact on middle secondary students interest towards STEM. *EURASIA Journal of Mathematics Science and Technology Education*, 13(5), 1189-1211. <https://doi.org/10.12973/eurasia.2017.00667a>
- Shahali, E. H. M., Halim, L., Rasul, S., Osman, K., Ikhsan, Z., & Rahim, F. (2015). Bitara-STEM™ training of trainers' programme: Impact on trainers' knowledge, beliefs, attitudes and efficacy towards integrated STEM teaching. *Journal of Baltic Science Education*, 14(1), 85-95. Retrieved from http://www.scientiasocialis.lt/jbse/files/pdf/vol14/85-95.Shahali_JBSE_Vol.14_No.1.pdf

- Shepherd, A. A. W. (2016). *The effect of middle school STEM curriculum on science and math achievement* (Unpublished doctoral dissertation). Union University, Tennessee, US.
- Thomas, M. E. (2013). *The effects of an integrated S.T.E.M. curriculum in fourth grade students' mathematics achievement and attitudes* (Unpublished doctoral dissertation). Trevecca Nazarene University, Tennessee, US.
- Tolliver, E. R. (2016). *The effects of science, technology, engineering and mathematics (STEM) education on elementary student achievement in urban schools* (Unpublished doctoral dissertation). Grand Canyon University, Arizona, US.
- Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013). Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics?. *Current Directions in Psychological Science*, 22(5), 367-373. <https://doi.org/10.1177/0963721413484756>
- Walker, L. H., & Sherman, H. J. (2017). Common core and STEM opportunities. *The Mathematics Enthusiast*, 14(1-3), 413-434. Retrieved from <http://schoarworks.umt.edu/tme/vol14/iss1/23>
- Walker, R. J. (2008). Twelve characteristics of an effective teacher: A longitudinal, qualitative, quasi-research study of in-service and pre-service teachers' opinions. *Educational Horizons*, 61-68. Retrieved from <https://files.eric.ed.gov/fulltext/EJ815372.pdf>
- Wilhelm, J. (2014). Project-based instruction with future STEM educators: An interdisciplinary approach. *Journal of College Science Teaching*, 43(4), 80-90. Retrieved from <http://www.jstor.org/stable/43632018>
- Wong, H. K., & Wong, R. T. (2010). *Developing and retaining effective teachers and principals*. Mountain View, CA: Harry K. Wong.

<http://www.iejme.com>