

REVIEW

Open Access



A meta-analytic investigation of the impact of middle school STEM education: where are all the students of color?

Danielle R. Thomas^{1*†}  and Karen H. Larwin^{2†}

Abstract

Background Integrated science, technology, engineering, and mathematics (STEM) education initiatives are becoming an increasingly popular approach to narrow the opportunity gap among underrepresented minority (i.e., Black, Hispanic, and first-generation) students. However, there are limited studies on the impact of exposure to integrated STEM education on academic achievement and an even lesser amount on performance among underrepresented minority (URM) groups. Students exposed to STEM programming in middle school are more likely to pursue a STEM field in college or a STEM-related career. However, despite increases in middle school STEM programming initiatives, STEM college graduation rates have declined, particularly among URM populations. This meta-analysis aims to determine the effectiveness of STEM education in middle school, focusing on URM students.

Results A total of 20 studies containing 45 independent samples met the study criteria. The studies included were published from January 1, 2011 to May 1, 2022, and identified from the following academic databases: ERIC, Google Scholar, ProQuest Dissertations and Theses, and SCOPUS. Integrated STEM programming was most impactful when: engineering was incorporated into science courses and at full STEM integration, occurring over one academic year ($d=0.89$) and occurring in 8th grade ($d=1.55$). Overall, the effect size estimate demonstrated heterogeneity, with a large positive significant effect across the studies ($d=0.558$, 95% CI [0.514–0.603], $p < 0.001$), indicating a significant impact on student achievement. The most notable finding was the lack of empirical studies involving URM groups, with only one effect size estimate reported for Black students and other minority groups and 40 effect size estimates for non-minority groups revealing a non-significant difference in effect size estimates.

Conclusions Students benefit from STEM program participation, with the average STEM student outperforming approximately 70% of their same-age, same-grade peers not participating in STEM programming. In particular, URM students benefit even more from quality integrated STEM education initiatives, given one caveat—students must be given the opportunity. We conclude that the issue is not that URM students are not academically benefiting from middle school integrated STEM education programs, based on the available research—they are merely not participating. We highlight the need and suggest interventions for providing collaborative and focused attention on the societal and cultural factors impacting URM student participation and retention in integrated STEM education programs.

Keywords STEM education, Integrated STEM, Career awareness, Middle school, Student achievement, Educational equity, Underrepresented minority students

[†]Danielle R. Thomas and Karen H. Larwin contributed equally to this work.

*Correspondence:

Danielle R. Thomas
drthomas@cmu.edu

Full list of author information is available at the end of the article

Introduction

STEM education is becoming increasingly popular in public education as a way of gaining student interest in STEM subjects, improving technological skills, and preparing students for future careers (Bryan & Guzey, 2020; Moore et al., 2020). According to the National Middle School Association, STEM education effectively engages teachers and students in active, purposeful learning—a crucial component of educating young minds (Lounsbury, 2010). The argument among researchers is not whether integrated STEM education is effective at the elementary, middle, or high school level but rather at which grade level is the introduction of STEM practices the most impactful on student achievement and, subsequently, future success (Bybee, 2010). The elementary and high school years have positive effects on shaping students' perceptions of their learning and future career choices (Bryan & Guzey, 2020). However, the middle school years demonstrate an optimal time to implement STEM education initiatives and programs (Christensen et al., 2015; Lesseig et al., 2017). Students interested in STEM in middle school are more likely to pursue a STEM field in college (Bryan & Guzey, 2020; Maltese et al., 2014). However, despite increases in middle school STEM programming, diversity among STEM college graduates appears stagnant (Premraj et al., 2021). There has been dramatic growth in the number of STEM graduates from U.S. colleges in the past decade. However, only 7 percent of STEM bachelor's degrees are earned from Black students and 12 percent earned from Hispanic students, given their share of all bachelor's degrees at 10 and 15 percent, respectively (Pew Research Center, 2021). Underrepresented minority (URM) groups (i.e., Black, Hispanic, and first-generation students) exhibit lower enrollment and graduation rates in STEM fields (Pew Research Center, 2021; Premraj et al., 2021). Interestingly, women make up exactly half of those employed in STEM jobs. However, they are overrepresented among health-related jobs (i.e., nursing) and vastly underrepresented in technical careers making up nearly 15 percent of engineers and architects and 25 percent of jobs involving computer science (Pew Research Center, 2021). Over 19 million U.S. workers are employed in STEM occupations, with two-thirds made up of White workers and the remaining third mostly composed of Asian (13%), Black (9%), and Hispanic (8%), respectively. In response to the lower proportion of students of color representing STEM careers, schools are attempting to expose students to STEM initiatives and programs to spark career interest and retain students in STEM fields, maintaining students in the pipeline. STEM careers and jobs are defined broadly as involving science, technology, engineering, and math. STEM careers are identified solely of 74

defined occupations in the sciences (i.e., life, Earth, physical), engineering and architecture, computers and math, and health and healthcare-related occupations (Pew Research Center, 2021).

The National Middle School Association (Lounsbury, 2010) recommends integrated STEM curriculum and instruction at the middle school level. At the middle school level, it offers engaging and holistic instruction for all learners with studies finding the integration of mathematics and science having a positive influence on student's attitudes toward school, their motivation to learn, and academic performance. Middle school is a pivotal time for cultivating student's interest and preparedness for STEM careers (Moreno et al., 2016) and an impressionable time for students as their viewpoints on STEM education and self-efficacy with respect to math and science are greatly impacted by their environment and access to learning opportunities (Blotnick et al., 2018).

Recently, Le Thi Thu et al. (2021) published a two-decade review monitoring the development of middle school STEM education from 2000 to 2020. The review of 272 academic journal articles determined a boom in quantitative analyses on the effectiveness of STEM education in the last five years. Although interest in providing and assessing high-quality STEM education in the middle years has become increasingly popular, research interests are remarkably diverse, focusing on many issues from gender studies, technology and engineering education, and curriculum. The proposed meta-analysis focuses on the impact of STEM programming on middle school academic achievement. In addition to determining the consistency of treatment effects of STEM education, this work unveils the lack of research on the participation of underrepresented minority (URM) students in STEM education programs. The lack of middle school students engaging in STEM education programs contributes to the deficiency of URM students enrolling in STEM majors in college—further depleting not only the number of students but the diversity of students to fuel the STEM pipeline.

Defining integrated STEM education

Currently, there is no single definition encompassing integrated STEM education outside of the interdisciplinary instruction of science, technology, engineering, and math (Moore et al., 2020). In response to the lack of cohesive understanding of integrated STEM education among educators and policy-making stakeholders, several researchers have operationalized a conceptual framework defining key concepts and learning theories surrounding integrated STEM education (Bryan & Guzey, 2020; Kelley & Knowles, 2016; Moore et al., 2020; Roehrig et al., 2021). Due to the ambiguity in defining STEM education,

the term *integrated STEM education* was created to incorporate all disciplines as a whole (Giasi, 2018) with researchers lacking a concise definition (Bryan & Guzey, 2020). Other terms used to describe STEM integration include: *interdisciplinary*, *cross-disciplinary*, *connected*, *fused*, or *transdisciplinary* with no definitive boundaries separating each discipline (Honey et al., 2014; Morrison, 2006). Recently, the *Handbook of STEM Education* was published, reviewing 109 sources providing definitions and conceptual frameworks of integrated STEM education (Moore et al., 2020). Researchers narrowed down six common themes pertinent to the description of integrated STEM education, listed below:

STEM integration:

- should be centered around real-world problems,
- applies concepts, principles, and ideas across disciplines,
- frequently uses student-centered learning approaches and peer collaboration,
- requires at least two disciplines,
- can exist on a wide continuum from little (or no) to full integration,
- often contains active learning, student-centered, problem- and project-based teaching pedagogies (Moore et al., 2020).

Studies used within this meta-analysis contain intentional subject integration applying many, if not all, of the common themes listed above. For example, some studies satisfied the themes defining integrated STEM education but focused primarily on engineering and technology integration, with other disciplines playing varying minor roles (i.e., *StEM* and *StEM*). It is important to note that studies analyzing the incorporation of *arts* into STEM integration, such as *STEAM*, were not included in this meta-analysis due to a lack of studies meeting the study criteria and containing a quantitative measure of student achievement.

Levels of STEM integration

There are several integrative approaches to incorporating STEM subjects reported and explained in the literature (see Becker & Park, 2011). Full integration, represented by the notation S-T-E-M, involves incorporating aspects of science, technology, engineering, and mathematics into all coursework using the previously defined components of STEM integration. Other integrative approaches representative of research articles in this meta-analysis include: S-E defined as an engineering-based STEM curriculum taught in science courses (i.e., Moreno et al., 2016), S-M defined as science and math concepts in an integrated STEM course or STEM concepts embedded

in math and science courses (i.e., Kutch, 2011), S-E-M described as an engineering design-based curriculum integrated into science and math courses (i.e., Harlan et al., 2014), S described as STEM concepts embedded in science courses. The latter has varied descriptions with researchers defining S integration as: STEM-related lessons in science courses (Gazibeyoglu & Aydin, 2019), engineering-based science curriculum with technology, engineering, and math imbedded in science courses (Selcen Guzey et al., 2017), and a STEM approach based on a 5E model in science courses (i.e., Izgi & Kalayci, 2020). The 5E learning model, based on constructivist theory, and created by Bybee, applies stages of experimentation to students' learning (i.e., engagement, exploration, explanation, elaboration, and evaluation) (Bybee et al., 2006) and was used in several international studies in this meta-analysis.

Measures of academic achievement

State standardized tests are the gold standard for measuring academic achievement as they provide a universal standard assessing the same constructs across the state. In addition, regarding college acceptance, state standardized tests provide universities with a common standard to evaluate individual student achievement. Math and science achievement assessments were used in the meta-analyses as both subjects are predictors of student academic achievement and interest in STEM fields (Blotnick et al., 2018). Standardized assessments are not without criticism. Because they are normed based on most students, they can create cultural bias among URM groups (Kim & Zabelina, 2015). Although not a perfect measure for quantifying student knowledge base, they provide the most readily available and consistent measure for determining student academic achievement across schools and states (William, 2010). Researchers and critics alike are working on methods of bias reduction and test optimization (Kim & Zabelina, 2015). Adherence to the study selection criteria (i.e., containing both independent and control comparison groups, pretest–posttest design) lessens testing bias and ensures testing outcomes are reflective of the learning process.

The majority of studies used in this meta-analysis measured academic achievement using at least one learning outcome. State standardized test scores were preferred, but common assessment measures were allowed when having both an independent comparison group and pretest–posttest design, given they met all other selection criteria.

Conceptual framework

We expand on the conceptual framework developed by Kelley and Knowles (2016), which provides much-needed

clarification into the operationalization and blending learning theories comprising integrated STEM education. Kelley and Knowles (2016) conceptualize the cognitive “load” of “situated STEM learning” by illustrating a pulley system, connecting the four common practices of scientific inquiry, technological literacy, mathematical thinking, and engineering design (p. 4). This conceptual framework provides insight to the relationship between all four STEM domains and the importance of the community of practice, which acts as a rope carrying the load of providing integrated STEM education. The community of practice, or group of practitioners, educators, students, and members of the community, fuel the mechanical pulley by roping social discourse and shared practices into the providing of STEM education (Kelley & Knowles, 2016). Designed for secondary education, and particularly for high school, this conceptual framework translates appropriately to middle school and encapsulates the complexity and integrity of the coordinated relationship between all four STEM domains. Kelley and Knowles (2016) emphasize that although the mental model illustrates a pulley system consisting of all four STEM subjects, this does not mean all domains must occur within every STEM learning experience.

We combine the conceptual framework of Kelley and Knowles (2016) and the individual efforts of describing different variations of STEM integration (i.e., S-E-M, S-M, S-T-E-M) from the following: Becker & Park, 2011; Gazibeyoglu & Aydin, 2019; Harlan et al., 2014; Izgi & Kalayci, 2020; Kutch, 2011; Moreno et al., 2016; Selcen Guzey et al., 2017. Figure 1 illustrates the conceptual framework of this meta-analytical study. *Integrated STEM education* is the independent variable and *student achievement* is the dependent variable, with moderators illustrated. Integrated STEM education is described aligning with the description of Kelley and Knowles’ (2016) awareness and understanding of the relationship across domains connected by the community of practice. The moderator variable, *level of STEM integration*

is added to provide more granular information into the description of STEM implementation across studies included in this meta-analysis and their particular student outcomes. Similar to Kelley and Knowles (2016), we affirm that not all STEM learning experiences that comprise integrated STEM education programs incorporate the four domains of STEM. *Level of STEM integration* as a moderator variable is intended to provide insight into the particular impact of the different levels of integration on student achievement. The other moderator variables of *dosage*, *grade level*, and *student demographics* are commonly provided in studies reporting student outcome measures as a function of participation in integrated STEM education (e.g., Becker & Park, 2011; Kazu & Yalcin, 2021).

Peripheral moderators, such as the year of study publication, assessment type, assessment subject, and other information used to measure student outcomes are not direct moderators.

Related work

Past research has determined the impact of STEM education policies and initiatives on student achievement having varying degrees of success (Dugger, 2010; Gonzalez & Kuenzi, 2012; Snyder, 2018; White, 2014). Although Gonzalez and Kuenzi (2012) report there is no single statistic that can fully quantify the success of STEM education on a national, state, or local level this study attempted to gain insight into the impact of STEM education on middle school student achievement. In a 2021 meta-analysis comprising 56 quantitative studies on the effect of STEM education on academic performance by education level (i.e., primary, secondary, high school, university), researchers determined large effect sizes across grade bands (i.e., primary level; $g=1.055$). However, no statistical significance was found across education levels (Kazu & Kurtoglu Yalcin, 2021). In addition, short STEM program intervention (2–5 weeks) produced the largest

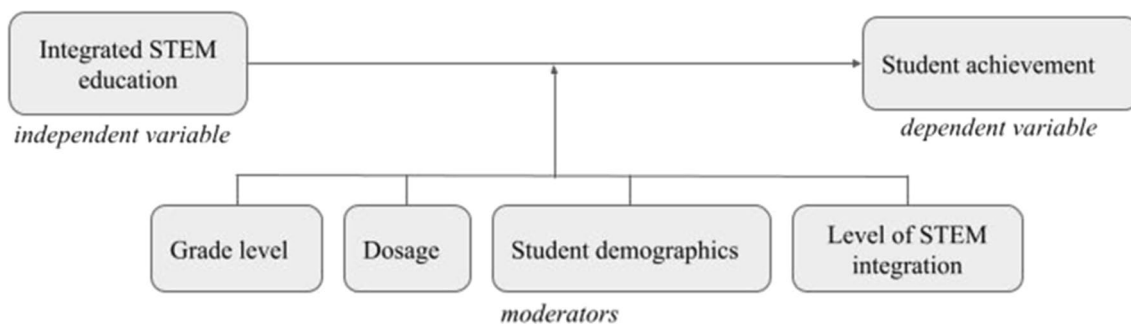


Fig. 1 Conceptual framework

effect size on student achievements, further supporting the importance of short-term or extracurricular STEM initiatives (Kazu & Kurtoglu Yalcin, 2021).

Several research studies have sought to determine the impact of STEM integration and programming on student achievement, sparking the interest in creating a meta-analysis of recent research particularly on math and science achievement. First, Wade-Shepard (2016) investigated the effect of the middle school STEM curriculum on both science and math achievement scores. The research was conducted among four schools of seventh and eighth grade students in Tennessee using the Tennessee Comprehensive Assessment Program (TCAP). The study found a significant, strong, and positive correlation between math and science test scores of students participating in STEM classes compared to those that were not taking STEM classes (Wade-Shepard, 2016). Hansen and Gonzalez (2014) investigated the relationships between STEM learning principles, such as project-based learning (PBL) and student achievement in math and science and found specific STEM practices were associated with performance gains in those subjects. For example, projects and science experiments were associated with higher scores in science, and using calculators, computers, and listening and taking notes were associated with higher scores in math.

In addition, these significant and positive correlations were also found among racial minorities (Hansen & Gonzalez, 2014). Last, Han et al. (2015) analyzed both STEM curriculum and project-based learning (PBL) strategies on student mathematics performance disaggregated by low, middle, and high achieving students to determine the degree of effect as a function of student achievement level (Han et al., 2015). Students in three Texas high schools participated in STEM project-based learning activities every 6 weeks over the course of 3 years. Han et al. (2015) concluded lower-achieving students showed a statistically significant higher rate of growth on math scores compared to middle and high performing students over three years. They also found student race and socioeconomic status were strong predictors of student academic achievement with low-income students exhibiting negative impacts due to participation in STEM-related PBL programs after the first year of implementation. Han et al. (2015) hypothesize the lack of learning gain among low-income students was due to unequal access to PBL materials. Analysis of student ethnicity found mixed results with Hispanic students benefiting more than Black students, hypothesized to be because of additional mathematical terminology exposure and opportunities for peer-to-peer and student-teacher relationship building. Both URM populations (i.e., low-income students and Black students) displayed a lack of learning gain in a

STEM PBL environment having had unequal access and opportunity (Han et al., 2015), hence lack of academic achievement gains. Bracey (2013) states that this is primarily due to the fact that many URM students attend subpar schools that focus on prescriptive academic remediation, which prevents access to the type of creative, inquiry-based learning that is foundational to STEM achievement. However, when conditions are optimal as suggested by Han et al. (2015), URM students can succeed in STEM education programs.

This meta-analysis is intended to combine the results of many studies to provide a more consistent estimate of the impact of integrated STEM programming on middle school student achievement. In addition, the review and exhaustion of literature will also gain perspective on the relative amount of research being conducted on URM groups participating in STEM education programs.

Underrepresented minority (URM) groups

The underrepresentation of minorities in the STEM workforce is a direct byproduct of the ever-present achievement gaps evident among minority youth in kindergarten through high school graduation (Gonzalez & Kuenzi, 2012). Gonzalez and Kuenzi (2012) comment that researchers have identified dozens of variables responsible for the achievement gap in STEM-related fields among minority populations, such as socioeconomic status (SES), “a lack of resources (underfunding), less qualified teachers at schools that serve minority students, teachers’ low expectations, stereotype threat, and racial oppression” (p. 24). The many factors responsible for the achievement gap in STEM education among minorities is a complex, multi-faceted issue fueling much of the research within this study. Milner (2020) refers to the achievement gap as an opportunity gap. In essence, many students are not succeeding due to not being given the opportunity to have a robust learning experience. In the development of the Opportunity Gap Framework, he references that in most cases, URM youth have not been provided the opportunity to learn and have positive experiences with school. By addressing the opportunity gaps that often exist for educators and students, teachers become reflective on the practices, policies, and experiences that directly impact student learning outcomes.

There are many proposed reasons for the lack of students from URM backgrounds enrolling in STEM college majors. In addition, among a small number of URM students enrolled in STEM-fields there is a lower rate of college completion among Black, LatinX, and Native American students (Chen, 2013; Williams et al., 2019). Williams et al. (2019) emphasizes the phenomenon is not due to a lack of STEM career interest among URM students. On the contrary, URM students

are the same or, perhaps, more likely than White students to choose a STEM major upon entering college (Gelbgiser & Alon, 2016; O'Brien et al., 2015). There are several hypothesized reasons for low retention among URM students in STEM-related majors (i.e., racial phenotypic stereotyping). Historically, common academic interventions, such as student support, mentoring, and tutoring sessions were prescribed for URM students to combat student attrition. However, the foundation of the problem is not academic in nature but social (Williams et al., 2019; Van Sickle et al., 2020). Researchers recommend a social context approach through addressing stereotypic bias among peers, academic advisors, and professors (Williams et al., 2019). This is supported by McGee (2021), who describes the STEM college experience for Black, Indigenous, and Latinx students as “chilly waters”. She posits that URM students often experience isolation, racial stereotyping, and impostor syndrome, all hindering success and decreasing retention in STEM fields. Further, URM STEM students at the university level often experience psychological and physical stress from harsh conditions that manifest themselves in higher drop-out rates.

Historically marginalized students, such as students in URM groups, often lack access to quality instructional materials and services and experience lesser opportunities for learning—thus, creating an opportunity gap (Chine et al., 2022; Schaldenbrand, 2021). The discrepancy in the number of learning opportunities between marginalized and non-marginalized students in education subsequently lowers student achievement. Researchers propose that to combat the opportunity gap in education, society must attend to other gaps, such as the teacher quality gap, school funding gap, digital divide gap, health care gap, quality child-care gap—and the list goes on (Irvine, 2010; Milner, 2020). Shirley Malcom, director of STEM Equity Achievement Change, an initiative of the American Advancement of Science, states “If you’re Black, you may have the drive, you may have the passion, but you also have deficiencies that were born of differential opportunities”, oftentimes, the focus is on “fixing the student rather than fixing the system” (Suran, 2021, p. 2). This is further supported by Bracey (2013), who contends that the behaviorist-reductionist teaching and learning model also contributes to the lack of interest and achievement of students in STEM areas. Many societal and cultural factors contribute to the opportunity gap evident among URM students. The introduction of STEM education and components of STEM integration in the early years and into middle school is associated with increased student interest in STEM fields into college and beyond. However, little research exists attesting

to the impact of such programming on URM students, in particular during their formative education years.

Impact of STEM education by grade level

There are many misconceptions surrounding the best time to introduce STEM education principles with a common fallacy being, “The belief that ‘real’ science, technology, engineering, and mathematics learning doesn’t occur until children are older, and that exposure to STEM concepts in early childhood is only about laying a foundation for the serious learning that takes place later” (McClure, 2017, p. 84). The belief that authentic STEM education does not occur until later years is a hot topic among researchers. What is the best time to introduce integrated STEM education models into students’ formal education?

Many researchers argue that early STEM programming, particularly from birth to 8 years old, is just as important as early literacy in the practice of critical thinking, persistence, and systematic experimentation. Being naturally born scientists, students are “never too young for STEM” with even the youngest learners able to think critically, conduct experiments and investigations, and make sense of the world around them (McClure, 2017, p. 84). A 2-year research analysis among preschoolers determined these young learners can carry out scientific practices using the scientific method matching that of high schools strengthening the belief that early STEM foundations are just as important as early literacy, with both emerging skills predicting future academic achievement (McClure, 2017). Early STEM literacy assists young students with developing attitudes toward STEM education and the exploration of future STEM-based careers (McClure et al., 2017; Christensen et al., 2015).

Adversely, some researchers and experts providing integrated STEM education feel high school is a particularly impactful time to introduce STEM learning models, with many current STEM initiatives beginning at the high school level (Barakos et al., 2012). The North Carolina New Schools Project has redesigned over one hundred high schools with the goal of every student graduating “ready for college, a career, and life” (Barakos et al., 2012, p. 5). Most of these schools specifically aim to teach high school students using integrated STEM instruction, project-based learning, real-life issues, and collaboration. These researchers view STEM-focused high schools as the most effective route to generating students’ interests in STEM fields and preparing them for STEM-related careers (Barakos et al., 2012). Although students make a career- and higher education-based decisions in high school, perhaps schools are missing an important opportunity to spark students’ interest in

STEM fields with lesser STEM programming options during middle school?

There are many reasons to introduce and support STEM programming and initiatives at the middle school level (approximately grades 5–8). With two-thirds of U.S. students failing to achieve proficiency in both math and science by the 8th grade (NAEP, 2019), the lack of knowledge hinders students and prevents future interest in STEM and technical careers. Cohen (2020) states students' academic interest tends to wane in the middle school years with many students who enjoy school losing interest in traditional schooling. Integrated STEM education revives many students' interest in school subjects. Second, many students begin to form career aspirations in their middle school years. The project-based learning methods and real-life applications involved in integrated STEM education assist students with future career exploration. Cohen (2020) states, "Exposure to STEM careers during this time triggers students to seriously consider jobs in engineering, technology, manufacturing, biology, etc." (website). Third, integrated STEM education often facilitates hands-on learning, which wanes in the middle school years with an increase in long lectures and many subjects taught in isolation. Fourth, STEM training teaches problem-solving principles which is particularly important in middle school as subjects begin to be taught in isolation. Lastly, integrated STEM education assists with closing the gender gap by exposing STEM principles to girls and boys before making definitive decisions regarding future careers (Cohen, 2020).

Inclusive STEM high schools (ISHSs) have recently popped up in states across the country, such as California, Massachusetts, Texas, and Ohio. These exclusive STEM-focused secondary schools accept students based on interest and not on achievement or aptitude (Spillane et al., 2016). Although the practice of choosing students based on STEM interests is a powerful method of recruiting invested students it can further decrease the number of girls and underrepresented populations. For this reason, many ISHSs intentionally recruit a larger proportion of minority groups often underrepresented in other STEM-related high schools. ISHSs are promising to enrich student STEM understanding, boost self-confidence in STEM subjects, and increase awareness in STEM college majors and careers (Lynch et al., 2017; Spillane et al., 2016).

Purpose and research questions

The aim of the meta-analysis is to determine the impact of STEM education programs and initiatives on academic achievement compared to students in a traditional setting not exposed to STEM interventions. Integrated STEM education is becoming an increasingly popular option to

improve student learning with limited studies on the benefits of STEM education concerning academic achievement and a lack of underrepresented minority (URM) groups participating in middle school STEM education and going into college STEM fields. This meta-analysis aims to determine the effectiveness of STEM education in middle school and highlights recent research:

1. What moderators (i.e., demographics, level of STEM integration, grade levels, etc.) are included in the research investigating the effect of STEM education programs on student's achievement?
2. What moderators (i.e., demographics, level of STEM integration, grade levels, etc.) or assessment types (i.e., math or science) demonstrate a larger effect of STEM programming on a student's academic achievement?
3. What differences exist in academic achievement between students participating in STEM education programs compared to students participating in a traditional setting?
4. What differences exist in academic achievement between underrepresented minority (URM) students or marginalized students participating in STEM programming compared to similar students in a traditional setting?

Methods

Meta-analysis was first invented by Glass (1976) and is a secondary analysis method used to answer research questions with improved statistical analysis. By integrating the quantitative results and effect sizes of past empirical studies, researchers can get a clearer picture of the research being studied (Glass, 1976). For purposes of this meta-analysis a systematic review of the literature was conducted using both primary and secondary databases to identify studies that met all study selection criteria.

Study selection criteria

A meta-analysis was conducted with the following criteria for studies to be included:

1. Studies had to use a randomized, true experimental design or quasi-experimental design.
2. Studies had to be empirical investigations of the effects of STEM programming and curriculum on student learning. Secondary data analyses, other meta-analyses, and literature reviews were excluded.
3. Studies had to be published within the reporting window from January 1, 2011 to May 1, 2022, and they had to be published in English.

4. Studies had to concentrate on students in any or all grades 5–8 and include students of all performance levels (e.g., high, middle, and low achieving students). Studies focusing on only a specific subgroup (e.g., students with disabilities) were excluded.
5. Studies had to contain an independent control or comparison group. Studies without a comparison group or containing one treatment group pretest–posttest design were excluded.
6. Studies had to quantify or measure academic achievement using at least one learning outcome. State standardized test scores were preferred but common assessment measures were allowed when having both an independent comparison group and pretest–posttest design.
7. Studies had to have at least 17 students in both the treatment and control group. Studies with sample sizes smaller than 17 students were excluded.
8. Studies had to include at least the minimum information and data necessary to estimate or calculate effect sizes.

Study search

The studies included in the meta-analysis were published from January 1, 2011 to May 1, 2022. Researchers searched the following academic databases: SCOPUS, ERIC, Google Scholar, and ProQuest Dissertations and Theses. One of the main research databases, SCOPUS, was used to search the terms “student achievement” and “STEM education” within the abstract, title, or author-specified keywords. Of the 49 articles found, researchers manually determined if each article met the study criteria. Upon manual selection, only two articles met the study criteria. The second main database, ERIC, was used with the same search criteria and displayed 290 articles with 5 of those articles meeting study search criteria upon hand-matching. The secondary scholarly database, Google Scholar, yielded 150 results meeting the search criteria upon searching “STEM education”, “student achievement”, “quantitative”, and “middle school” within the title using the “OR” Boolean operator and manually selecting 11 articles meeting the search criteria. Lastly, ProQuest Dissertations and Theses search found article results when filtering using the same terms as SCOPUS and ERIC, isolating 20 articles with four that met all criteria. All keywords were searched using a combination of Boolean operators (AND, OR). A total of 22 studies were found with 20 meeting all study criteria and containing 45 independent samples. Figure 2 displays the data collection process highlighting the search, screening, and selection of qualified articles meeting the eight criteria requirements.

Effect size calculation

Meta-analysis is used to synthesize effect size estimates across several studies, with primary and secondary data. Effect size estimates measure the impact of treatment, such as integrated STEM education programming. Effect size estimates are standardized values, making it possible to compare the direction and magnitude of the variables of interest. There are several methods for calculating effect sizes (Glass et al., 1981; Hedges & Olkin, 1985; Hunter & Schmidt, 2004; Rosenthal, 1991; Wolf, 1986). For this meta-analytic study, all statistics from each study will be converted to Hedges’ *d*. Hedges’ *d* statistic is defined as the difference between the means of the experimental and control groups divided by the inter-group standard deviation. Means and standard deviations were available in order to calculate effect size measures for several of the studies included in the current investigation. Effect size measures were calculated with means and standard deviations using the formula from Johnson (1989):

$$d = (M_E - M_C) / S_{\text{pooled}}$$

considering:

$$S_{\text{pooled}} = \left[(n_E - 1) (s_E)^2 + (n_C - 1) (s_C)^2 \right] / (n_E + n_C - 2),$$

where M_E is the mean for the experimental group, M_C is the mean for the control group, n_E is the number of participants in the experimental group, n_C is the number of participants in the control group, s_E is the standard deviation of the experimental group, and s_C is the standard deviation of the control group. When means and standard deviations were not available, the effects sizes were calculated using other formulas for calculating effect size. In this meta-analysis, the effect size estimate for studies providing an *F* statistic was calculated using the following formula:

$$d = F[(n_E + n_C) / (n_E - n_C)]^{1/2}.$$

The effect size estimate for studies providing the Chi-square statistic was calculated using this formula to acquire the *r* value:

$$r = \left(\chi^2 / n \right)^{1/2},$$

which can then be used to calculate the *d* value using the following formula:

$$d = 2r / \left(1 - r^2 \right)^{1/2}.$$

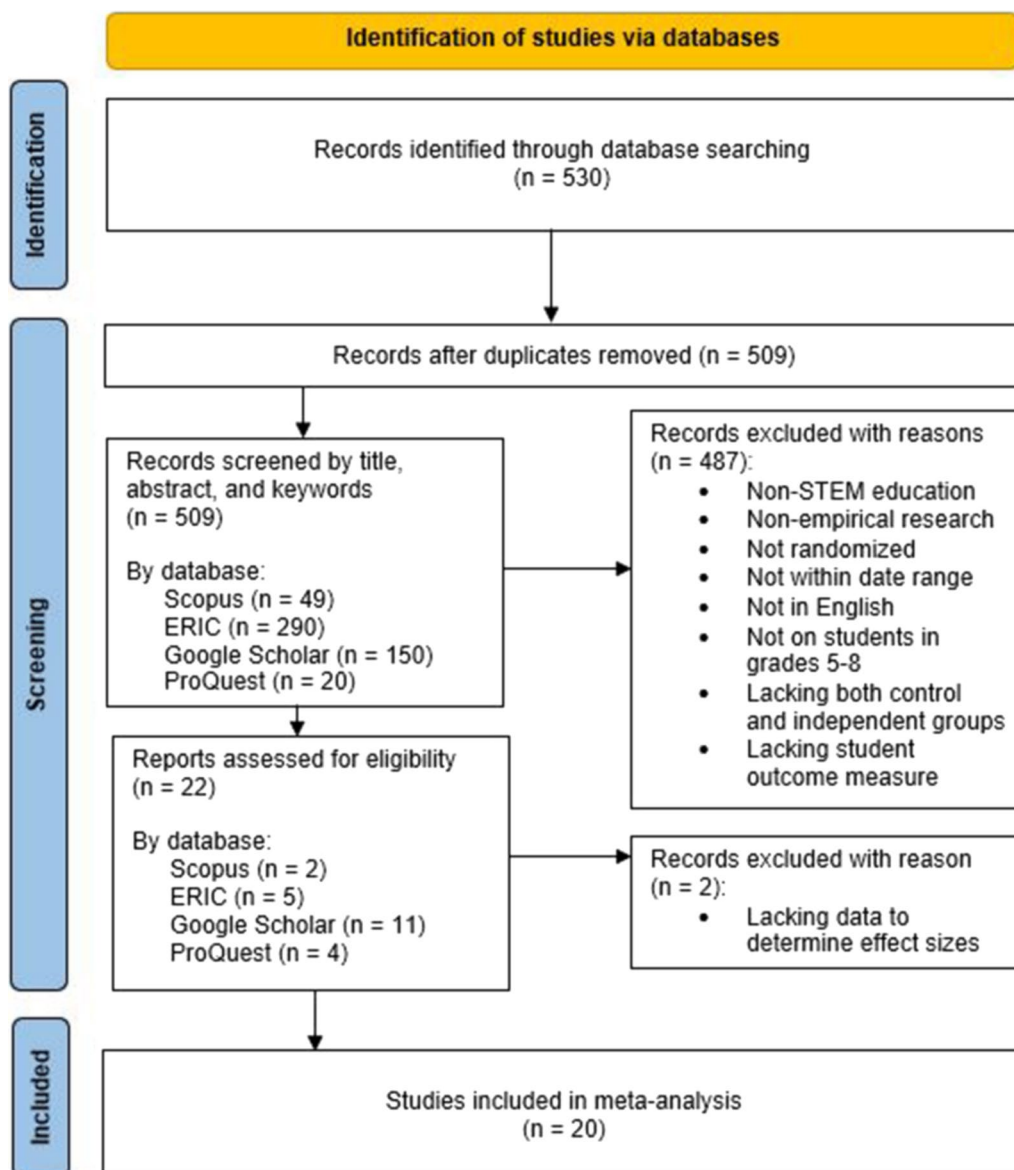


Fig. 2 PRISMA flow diagram of data collection

The effect size estimate for studies providing *t* statistics were calculated using this formula:

$$d = t[N/(n_E * n_C)]^{1/2}$$

Johnson (1989).

Lastly, some studies provided effect size estimates, and no additional computation was necessary. Once the effect sizes are calculated for the individual studies, the overall

effect size measure for all the studies combined can be calculated. This can be done, according to Glass et al. (1981), by simply calculating the mean of the individual effect size measures. However, this approach does not take into consideration the fact that the studies vary in sample size. Hedges and Olkin (1985) provide a formula for calculating the overall mean effect size as an unbiased weighted estimate (weighted by sample size) of the population effect size:

$$d_+ = \frac{\sum_{i=1}^k \frac{d_i}{\hat{\sigma}^2(d_i)}}{\sum_{i=1}^k \frac{1}{\hat{\sigma}^2(d_i)}} = \frac{\sum_{i=1}^k w_i d_i}{\sum_{i=1}^k w_i}$$

where the variance of d is calculated using the following formula:

$$\sigma^2(d_+) = \frac{\sum_{i=1}^k w_i^2 \hat{\sigma}^2(d_i)}{\left(\sum_{i=1}^k w_i\right)^2} = \frac{\sum_{i=1}^k \left(\frac{1}{\hat{\sigma}^4(d_i)} \cdot \hat{\sigma}^2(d_i)\right)}{\left(\sum_{i=1}^k \frac{1}{\hat{\sigma}^2(d_i)}\right)^2} = \frac{1}{\sum_{i=1}^k \frac{1}{\hat{\sigma}^2(d_i)}}$$

as well as the corresponding confidence intervals using this formula:

$$d_+ - Z_{\alpha/2}\sigma(d_+) \leq \delta \leq d_+ + Z_{\alpha/2}\sigma(d_+)$$

in order to calculate a 100(1- α) confidence interval (p.111). The overall mean effect sizes for this meta-analysis were calculated according to the procedures recommended by Hedges and Olkin (1985) within the Comprehensive Meta-Analysis, a dedicated meta-analytic software.

Results

Description of studies

Data were extracted from the studies that met the inclusion criteria. Cohen’s d was computed from the studies that provided means, standard deviations, and sample sizes for the control and treatment groups. For studies reporting t-test outcomes, F-test results, Chi-square data, p values, r , and R^2 values, and sample sizes were used to compute Cohen’s d . This data is recomputed as a Cohen’s d for each study. If the study did not provide enough information, it was not included. Comprehensive Meta-Analysis® was used for this analysis. Effect size estimates were based on random effects as opposed to fixed effects. Random effects were used since the student gains measures were inconsistent across the different studies. The random-effects estimate does result in a more conservative estimate than the fixed effects estimate.

The results from the 20 studies were extracted and resulted in 45 effect size estimates. Overall, the effect size estimate demonstrated heterogeneity, with a large positive significant effect across the studies, Cohen’s $d=0.558$. $p<0.001$, $CI_{95}[0.514: 0.603]$. The Z-value is $Z=5.601$, $p<0.001$, suggesting that the mean effect size differs from zero (Borenstein et al., 2021; Hedges & Olkin, 1985). Additionally, the Q-statistic, which provides a test of the null hypothesis that all studies in the analysis share a common effect size, was computed. The results for Q indicate that $Q(44)=3080.71$, $p<0.001$, indicating that the true effect size is statistically different across the analyzed studies. Likewise, the I-squared

statistic is 99%, which indicates that 99% of the variance in observed effects reflects variance in true effects rather than sampling error. The variance of true effect sizes, $\tau^2=0.414$. Finally, the resulting prediction interval is -0.729 to 1.904 , indicating that the true effect size in 95% of all comparable studies will fall within this interval

(Borenstein et al., 2021).

Specifically, these results indicate that students benefit from their participation in STEM, and the average STEM student will outperform approximately 70% of their same-age, same-grade peers who are not in STEM programming. Since the results indicate significant heterogeneity (variation in study outcomes between studies) is indicated for the full model, $R=0.55$, $SE=0.20$, $CI_{95} [0.15, 0.96]$, $p=0.007$, additional analyses will identify the study differences moderating this large effect estimate outcome. Table 1 provides a breakdown of each study included in the analyses.

As indicated in Table 1, the number of effect size estimates extracted from the studies ranges from one to seven. The effect size estimates range from a large $d=-0.19$ to a large $d=6.41$. Figure 3 illustrates the forest plot of all included effect sizes in the random-effects model allowing for heterogeneity to yield an average treatment effect across studies.

The studies were each examined for potential moderators to the reported outcomes. The identified moderators include grade level of the student group, reported race, type of STEM integration (see Becker & Park, 2011), dosage or time in STEM programming, assessment type (English Language Arts-ELA, Math, or Science), state or local assessment data, year of publication, data source, and location of research (domestic or international). The effect size estimates by reported race are provided in Table 2.

As indicated above, there is a significant difference across the studies with designated Black student, Hispanic student, and minority student data, relative to the effect size data without, $d=0.745$, $p=0.002$. Further examination of the data by the minority (Black, Hispanic, and Minority) relative to non-minority reveals a non-significant difference in the effect size estimates, $d=0.611$, $p=0.301$.

Thirty potential effect size estimates provided specific grade level data with its reported data. The results by grade level are in Table 3.

Table 1 Overall effect size estimate by study

Study	# Of effect estimates	Overall effect	SE	Lower	Upper	<i>p</i>
Adams	7	- 0.30	0.17	- 0.63	0.04	0.085
Bentley	2	- 0.50	0.10	- 0.69	- 0.31	0.000
Chine	2	0.70	0.10	0.51	0.89	0.000
Gazibeyoglu & Aydin	1	0.68	0.29	0.12	1.24	0.017
Gerlach	4	0.79	0.36	0.08	1.50	0.028
Gülen & Yaman	1	0.72	0.33	0.08	1.36	0.028
Harlan et al.	3	0.31	1.21	- 2.06	2.68	0.798
Hidge et al.	1	- 0.81	0.31	- 1.43	- 0.20	0.010
Izgi & Kalaysci	1	0.99	0.30	0.40	1.58	0.001
James	4	0.32	0.27	- 0.22	0.85	0.247
Kurt et al.	1	1.04	0.42	0.22	1.86	0.013
Kutch	3	0.06	0.09	- 0.11	0.23	0.486
Moreno et al.	1	1.17	0.13	0.91	1.43	0.000
Olivarez	3	0.59	0.10	0.40	0.79	0.000
Özcan & Koca	1	0.72	0.37	0.00	1.44	0.049
Price et al.	2	0.87	0.20	0.48	1.25	0.000
Sarican & Akgunduz	1	0.22	0.30	- 0.37	0.82	0.457
Selcen Guzey et al.	1	0.23	0.05	0.14	0.32	0.000
Tolliver	2	0.19	0.56	- 0.90	1.28	0.733
Wade-Shepard	4	2.98	1.33	0.38	5.59	0.025

Robust variance estimates (RVE) were not computed because of the limited number of studies in the current investigation that are used for multiple effect size estimates. Hedges et al. (2010) simulations suggest that RVE are valid when there are a larger number of studies with multiple estimators, suggesting twenty or more. The current investigation includes nine studies that produce a single effect size estimation, leaving eleven students with potentially non-independent estimates. Data are broken down by study and by the various moderators, in order to provide the estimates for all possible levels of effect

As indicated in Table 3, significant differences are present based on grade level, $p < 0.001$. The most significant effect size estimate is reported for eighth grade students ($d = 1.55$) followed by multiple grades studies ($d = 1.17$). STEM integration of each study was established based on the guidelines used in Becker and Park (2011). The results for STEM integration are provided in Table 4.

As indicated above, “S-E” and “S-T-E-M” produce the largest positive effect size estimates for STEM integration. The effect size differences across STEM integration are statistically significant, $p < 0.001$. Dosage of STEM experiences are provided in Table 5.

Effect size estimates for dosage indicate that the largest effect is seen with a one-year program ($d = 0.89$), followed by a 4-year program ($d = 0.87$). Results indicate that the differences in dosage are statistically significant, $p < 0.001$. Table 6 provides the effect size estimates based on the type of outcome measure used.

As indicated above, the greatest impact is found in the ELA assessments ($d = 2.02$) with a large significant positive effect estimate, followed by a large positive effect on science ($d = 0.50$). These effect size estimates by Assessment Type were statistically significant, $p < 0.001$. Additionally, the outcome data were

examined by whether the outcome was a state or local assessment. Results are in Table 7.

The state and local assessment results indicate large effect size estimates; however, the state assessments produce a significantly larger estimate, $d = 0.60$, $p < 0.001$. Year of publication is a peripheral moderator that is examined to understand if there is a trend that presents, across time, regarding the effectiveness of STEM in middle school. The results are presented in Table 8.

While there is variability in the data, the decline in 2022 results in a non-significant relationship between year and effect size estimates ($p > 0.05$). However, if 2022, which is based on one study that was likely impacted by the COVID-19 pandemic, is removed, the association between year and effect size estimate increases to $r = 0.345$, $R^2 = 0.119$, $p < 0.001$ (Additional file 1).

Data were extracted from a program evaluation report, peer-reviewed publications, and dissertations/theses. The results for the data sources moderator are provided in Table 9.

Significant differences were found across publication type, $p < 0.001$, with the largest effect reported in a program evaluation report data, followed by Dissertations/Thesis data. Finally, data were examined by location of

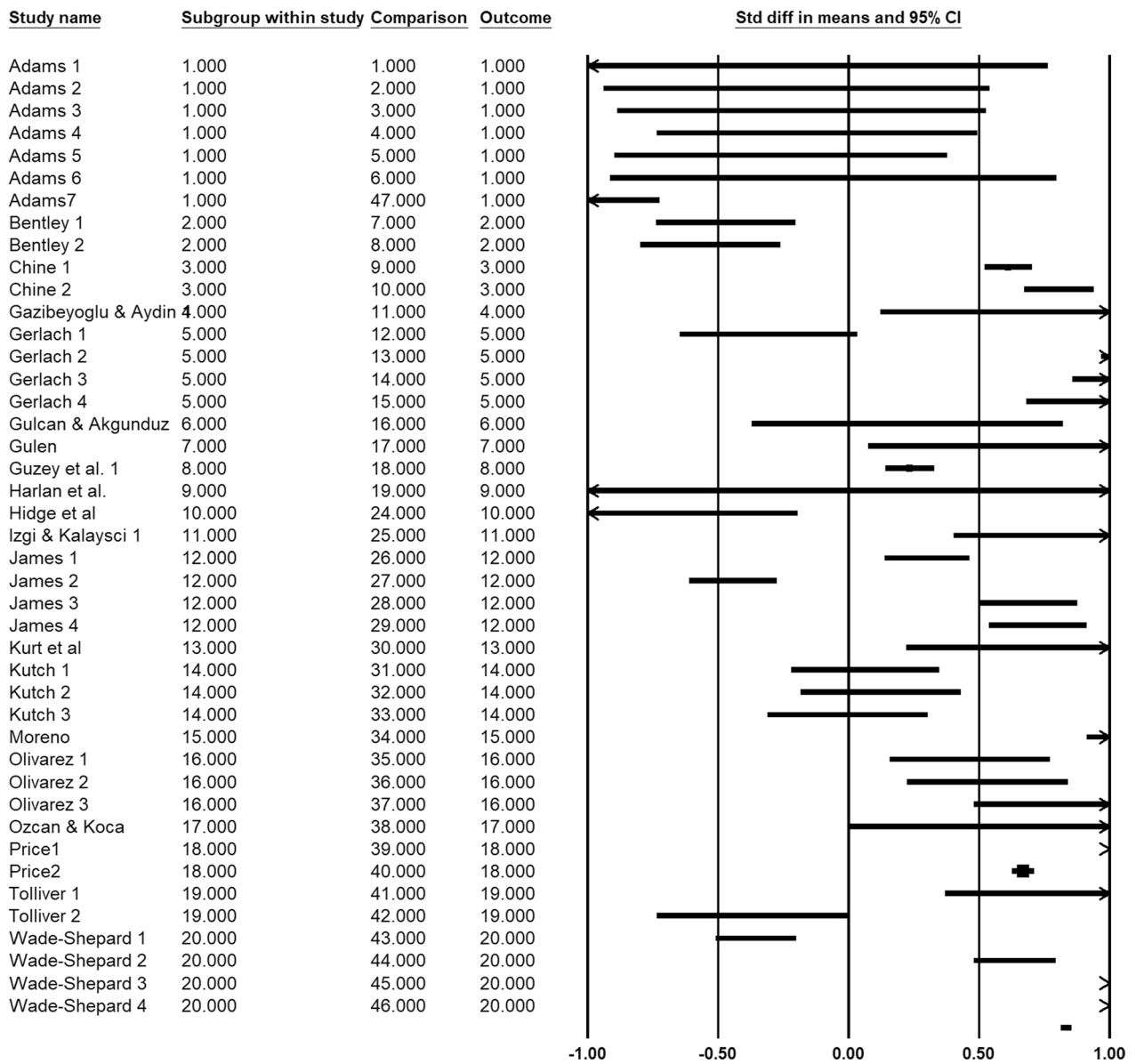


Fig. 3 Forest plot of all included effect sizes in the random-effects model

Table 2 Effect size estimate by reported race

Race	# Of effect estimates	Overall effect	SE	Lower	Upper	p
Race: Black	1	- 1.90	0.60	- 3.07	- 0.73	0.001
Race: Hispanic	1	- 0.18	0.36	- 0.88	0.52	0.616
Minority	1	1.17	0.13	0.91	1.43	0.000
Non-Minority	40	0.63	0.11	0.41	0.84	0.000

Table 3 Effect size estimates by reported grade level

Grade level	# Of effect estimates	Overall effect	SE	Lower	Upper	<i>p</i>
5th Graders	5	0.09	0.45	− 0.78	0.97	0.838
6th Graders	4	0.27	0.40	− 0.51	1.05	0.499
7th Graders	10	0.34	0.18	− 0.01	0.68	0.054
8th Graders	10	1.55	0.32	0.92	2.19	0.000
Multiple grades	1	1.17	0.13	0.91	1.43	0.000

Table 4 STEM integration

	# Of effect estimates	Overall effect	SE	Lower	Upper	<i>p</i>
S	7	0.58	0.15	0.28	0.89	0.000
S-E	1	1.17	0.13	0.91	1.43	0.000
S-E-M	5	0.32	0.02	0.28	0.37	0.000
S-M	9	0.05	0.17	−0.28	0.39	0.752
S-T-E-M	24	1.09	0.15	0.79	1.39	0.000

Table 5 Dosage (based on year)

	# Effect size estimates	Overall effect	SE	Lower	Upper	<i>p</i>
0.23	7.00	0.80	0.15	0.51	1.10	0.000
0.38	1.00	− 0.81	0.31	− 1.43	− 0.20	0.010
0.5	3.00	0.06	0.09	− 0.11	0.23	0.486
1	21.00	0.89	0.26	0.39	1.39	0.001
2	2.00	0.70	0.10	0.51	0.89	0.000
3	7.00	− 0.30	0.17	− 0.63	0.04	0.085
4	2.00	0.87	0.20	0.48	1.25	0.000

Table 6 Effect size estimates by subject area of assessment

Assessment type	# Of effect estimates	Overall effect	SE	Lower	Upper	<i>p</i>
ELA	4	2.02	0.85	0.36	3.68	0.02
MATH	16	0.34	0.13	0.09	0.59	0.01
SCIENCE	9	0.50	0.13	0.24	0.76	0.00
COMPREHENSIVE	13	0.23	0.04	0.14	0.32	0.00

Table 7 Effect size estimates by state or local assessment

Assessment level	# Of effect estimates	Overall effect	SE	Lower	Upper	<i>p</i>
State	28	0.81	0.129	0.017	0.562	0.001
Local	14	0.096	0.210	− 0.315	0.508	0.001

Table 8 Year of publication

Year	# Effect size estimates	Overall effect	SE	Lower	Upper	<i>p</i>
2011	3	0.06	0.09	- 0.11	0.23	0.49
2012	3	0.59	0.10	0.40	0.79	0.00
2014	5	0.32	0.27	- 0.21	0.82	0.24
2015	2	- 0.50	0.10	- 0.69	- 0.31	0.00
2016	8	1.71	0.59	0.55	2.87	0.00
2018	1	0.22	0.30	- 0.37	0.82	0.46
2019	5	0.80	0.15	0.50	1.10	0.00
2020	6	0.86	0.28	0.31	1.41	0.00
2021	8	0.67	0.04	0.60	0.74	0.00
2022	1	- 0.81	0.31	- 1.43	- 0.20	0.01

Table 9 Data source

Data source	# Of effect estimates	Overall effect	SE	Lower	Upper	<i>p</i>
Evaluation report	2	0.87	0.20	0.48	1.25	0.000
Publication	17	0.21	0.15	- 0.09	0.51	0.179
Dissertation/Thesis	24	0.79	0.22	0.37	1.22	0.000

the study (international or domestic) and no significant differences were found, $p = 0.658$.

Limitations

While this study focused on the impact of integrated STEM education on middle school STEM academic achievement, a few limitations exist in this meta-analytic research. First, given the necessary criteria for articles to be included in this meta-analytic study, our analysis excludes several empirical studies that have substantial value in providing insight into our research questions, however, they did not meet article selection requirements. This study included 20 studies containing 45 independent effect sizes. Second, the small quantity of studies meeting the selection criteria affects study generalizability. The conclusion that integrated STEM education is beneficial for URM students should be approached with caution. This result was interpreted based on three effect size estimates and two of these outcomes are from the same study (Adams, 2021), demonstrating the lack of generalizability of findings related to the effect of integrated STEM education and URM student achievement.

An analysis of gender could not be determined due to a lack of studies meeting the selection criteria reporting gender as a moderator. Lastly, we could not break apart Minority groups into demographic subgroups due to individual studies not reporting subgroups (see Table 2). The overall lack of empirical studies in the literature

reporting on race used was small with few articles in this meta-analytic investigation reporting at least the minimum information and data necessary to estimate effect sizes. Although the lack of studies on race is a limitation of this work, it was also a substantial finding.

Publication bias

Publication bias is assessed to ensure that published studies do not dominate the effects found in a meta-analytic study. The Egger's test of the Intercept suggests that bias is assessed by using precision (the inverse of the standard error) to predict the standardized effect (effect size divided by the standard error). In this equation, the size of the treatment effect is captured by the slope of the regression line (B1) while bias is captured by the intercept (B0). This approach may offer a number of advantages over the rank correlation approach. Under some circumstances, this may be a more powerful test. Additionally, this approach can be extended to include more than one predictor variable, which means that we can simultaneously assess the impact of several factors, including sample size, on the treatment effect. In this, the results indicate $t(44) = 1.73$, $p = 0.091$, CI95 [- 0.54: 7.02], suggesting no significant publication bias exists. Figure 4 illustrates the funnel plot supporting the finding that there is a lack of complete asymmetry suggesting the absence of bias (Lin & Chu, 2018).

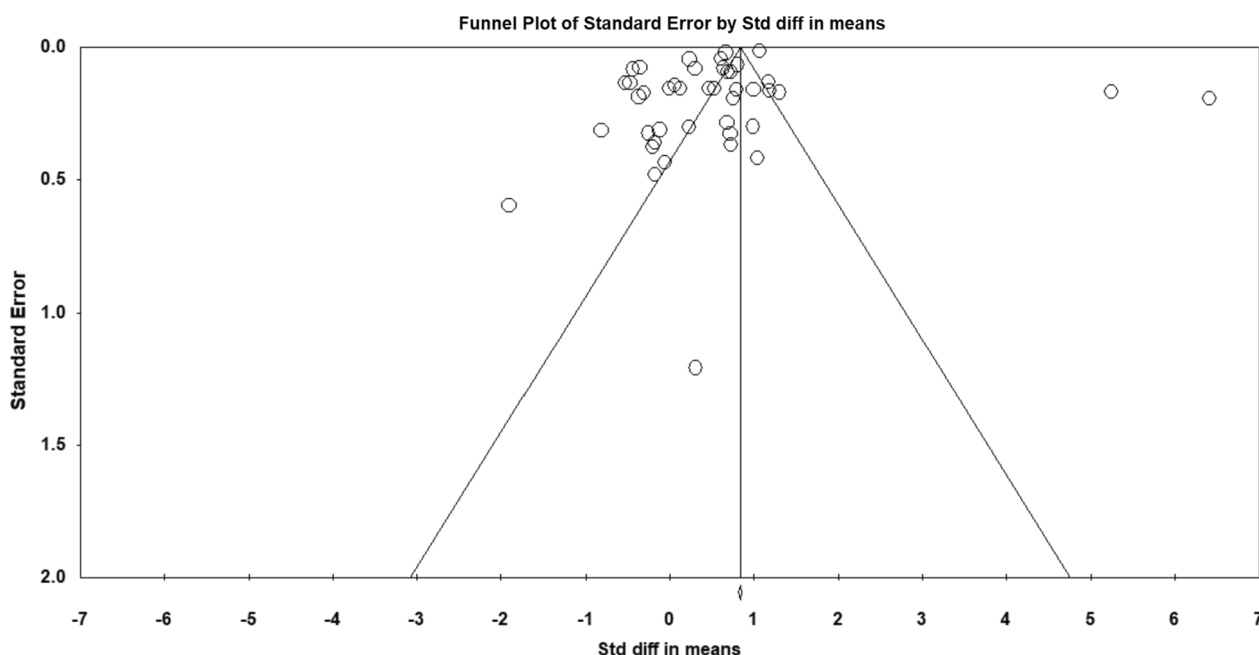


Fig. 4 Funnel plot of effect sizes with 95% confidence interval boundaries

Discussion

This meta-analysis determined the overall effect of twenty studies and resulted in 45 effect size estimates. Overall, the effect size estimate demonstrated heterogeneity, with a large positive significant effect across the studies, Cohen's $d=0.558$, $p<0.001$, $CI_{95}[0.514: 0.603]$. Specifically, this indicates that students benefit from their participation in STEM, and the average STEM student will outperform approximately 70% of their same-age, same-grade peers who are not participating in STEM programming. Since the results indicate heterogeneity exists across the studies, additional analyses identified the study differences moderating this large effect estimate outcome. We further expand on the difference in findings among moderators and attend to each research question below.

Moderators of student achievement

The first research question sought to identify moderators (i.e., demographics, level of STEM integration, grade levels, etc.) or assessment types (i.e., math or science) of student achievement and, in particular, the most impactful influencers of student achievement. We isolated the following moderators of achievement: grade level, student race, level of integration, dosage, data source (i.e., dissertation, publication), and publication year.

The majority of studies occurred in 7th and 8th grade with 12 and 10 studies, respectively (see Table 2). This is not surprising given many integrated STEM initiatives

begin during peak middle school years (i.e., 7th and 8th grade). The 8th grade was the most impactful and statistically significant ($d=1.55$) with 7th grade considerably less impactful showing comparatively weak impact ($d=0.31$). As previously discussed, Cohen (2020) supports the introduction of STEM initiatives during the middle school years as it can provide a resurgence of student investment in school when interest in traditional schooling begins to wane. In addition, STEM programming can spark career interests, facilitate hands-on learning, and encourage problem-solving across subjects during a time when subjects are often taught in isolation. Given all these explanations for the increased impact in middle school, why is there a substantially stronger impact of STEM programming in 8th grade as opposed to other middle school grades? We propose that the influences proposed by Cohen (2020) are stronger in the 8th grade, perhaps, the oldest among middle school students has the most vested career interests. Additionally, we infer that many middle school STEM programs span across middle school years concluding in 8th grade. A common implementation model is to begin with one grade only for the first year and each year add a subsequent grade ending with 8th grade. The larger effect size evidenced in the 8th grade may be due to overcoming an implementation dip in the lower grades. Fullan (2007, p. 40) describes an "implementation dip" as a drop in performance and, sometimes, confidence as a function of an innovation that requires new skills. Our findings indicative of a

possible implementation dip is not surprising, particularly during early adoption of any systemic program, policy, or initiative involving collective change (Fullan, 2007). Implementing STEM programs often requires a change in teaching strategies and techniques, which can initially cause confusion and difficulty for students. The acquisition of new teaching techniques and training on interdisciplinary instruction can be difficult for educators and comes with implementation challenges.

Despite these initial implementation challenges, Fullan (2007) emphasizes overcoming these obstacles is imperative to implementing positive change and continued academic growth. STEM educators and educational leaders should be aware of the two types of problems when experiencing an “implementation dip”: the social-psychological fear of change, common when facilitating new educational policies, programs, or practices that require a shift in collective thinking; and the lack of technical knowledge or skills required to ensure successful outcomes (Fullan, 2007). In relation to STEM educator training and knowledge, the Technology Pedagogical Content Knowledge (TPACK) framework is often used to support, facilitate learning, and assess STEM educators, claiming the interplay between technology, pedagogy, and content is necessary to ensure successfully integrated STEM education (Morales et al., 2022; Schmidt et al., 2009). It involves an understanding of the content knowledge (CK), pedagogical knowledge (PK), and technological knowledge (TK) necessary to design effective STEM learning experiences in a meaningful way (Schmidt et al., 2009). Educational leaders can support the development of TPACK in STEM educators several ways: being aware of valid instruments to assess TPACK skills and using them as a measure educator knowledge (Schmidt et al., 2009); providing professional development (PD) opportunities for educators to participate in ongoing PD, such as workshops, online courses, and peer-to-peer mentoring (Major & McDonald, 2021); supporting curriculum development by providing resources and funding for the creation of integrated STEM units that incorporate technology, and ensuring that the curriculum aligns with the latest standards and best practice STEM educational practices; encouraging collaboration and sharing among STEM educators; and, lastly, ensuring that school have the necessary technology infrastructures to support technology integration in STEM education, such as devices, software, and hardware, along with adequate educator training to effectively use these resources (Major & McDonald, 2021).

Nonetheless, teacher enthusiasm, confidence, and pedagogy development improve over time with increasing implementation year (Tytler et al., 2019). In addition, students may be adapting to new learning methods

and using critical thinking skills that may not have been used prior to exposure to STEM programming, particularly URM students or students lacking prior opportunities and exposure. URM students engaging in integrated STEM education may demonstrate lower shifts in achievement and performance due to implementation dips compared to their non-URM peers. We posit decreases in achievement among URM students participating in STEM education programs, evidenced by the implementation dip phenomenon, contribute to the “pipeline” leakage of more URM students withdrawing from STEM programs compared to white or Asian students (Estrada et al., 2016). We propose several methods STEM leaders can leverage to retain URM students and increase achievement and performance, particularly during the early stages of STEM initiatives. First, STEM educators and leaders should design and implement culturally responsive teaching using curricula that are culturally responsive and inclusive, which involves acknowledging and valuing the diverse experiences, perspectives, and backgrounds of students (Villegas & Lucas, 2002). STEM educators can embrace culturally responsive teaching practices by being socioculturally conscious, upholding the viewpoints of diverse students in the classroom, recognizing themselves as responsible parties to create change and promote equitable outcomes, and design instruction that builds on what their URM student already know (Villegas & Lucas, 2002) and not a construct of what curricular organizations and developers think URM students know. Second, Villegas and Lucas (2002) express the importance of educator PD that helps model responsive educator characteristics evident in progressive curriculum. In addition, they emphasize honoring multicultural perspectives and responsiveness in a way that is embedded in the vision of the school and collective teacher capacity, further providing an organizing framework to achieve this complex task (Villegas & Lucas, 2002).

There was only one effect size estimate for both the Black and Hispanic race categories and one effect size estimate defined as a minority group with 40 estimates for non-minority student groups. Effect size examination revealed a non-significant difference among minority groups (Black, Hispanic, and Minority) relative to non-minority effect size estimates, $d=0.611$, $p=0.301$. We can infer from this finding that it is not the case that Black, Hispanic, and perhaps students from URM groups are not academically benefiting from integrated STEM education programs—they are merely not participating! The limited number of research studies providing race or minority group data on student achievement is staggering. This finding further showcases the lack of URM student participation in STEM and STEM-related programs

previously reported by Moreno et al. (2016) and Estrada et al. (2016).

The most impactful level of integration occurred with the incorporation of science and engineering (S-E, $d=1.17$). However, the effect size estimate was determined from one study. The highest frequency of effect size estimates ($n=24$) occurred at full integration (S-T-E-M, $d=1.09$) with a significant effect size difference across integration types. A previous meta-analysis analyzing the impact of 28 studies across seven forms of integration (E-M, S-T-E-M, S-E, S-T-E, S-M, S-T-M, and S-T) determined the effects of integrative approaches among STEM subjects (Becker & Park, 2011). Although the meta-analysis is antiquated for purposes of analysis of findings for this work (published in 2011 using empirical studies spanning ten years prior), the method of coding studies by subject integration was used. Similar to Becker and Park (2011), we conclude it is difficult to analyze the results of the meta-analysis given the few empirical studies for particular integration types (i.e., S-E, S-E-M). Due to few empirical studies at certain levels of integration, further research needs to be conducted along with more diversified STEM integration methods. However, the large effect size of greater than one standard deviation ($d=1.09$) across 24 independent samples supports the positive impact of full STEM integration on student achievement.

Full STEM integration, indicated by the notation S-T-E-M, is described in detail with sample curricula explained in several meta-analysis articles included in this investigation. Adams (2021) emphasizes the weaving of interdisciplinary PBL to engage real-world problem-solving, providing the sample project: students building a wind turbine in science while concurrently writing a technical manual in ELA. This project could be further expanded to incorporate math standards (i.e., calculating the circumference and area of the circular rotation of the turbine blades or creating scaled drawings to include in the technical manual), with elements

of technology included by expanding on the use of next-generation wind energy and applications. Interdisciplinary curricula are strategically implemented to increase the academic achievement of students across all STEM subjects (Bybee, 2013). Chine (2021) provides another example of full STEM integration by referencing Bybee’s five aspects of the learning cycle theory: engagement, exploration, explanation, elaboration, and evaluation (1997) and Dewey’s constructivist learning-by-doing approach (1897). Chine (2021) describes two comprehensive and interdisciplinary annual projects completed by students participating in a fully integrated STEM education program: the building and racing of Soap Box Derby cars and the assembling, launching, and retrieving of a weather balloon. The former involves students engaging in math, science, and engineering curriculum grade-level standards while building the cars, with the topics in the following order: collection and analysis, ratio and proportion, geometry, simple machines, gravity, energy, friction, and speed. Taking over two months to complete with students participating approximately 45 min per school day, the “Masters of Gravity” curriculum includes optional competition projects involving a photography contest, infomercial creation, and press release design, which includes ELA standards allowing for an immersive, fully integrated STEM experience (Masters of Gravity, n.d.). A last example of fully integrated STEM programming involved a 10-week, activity-based education program with students participating in approximately one activity per week (Hiğde & Aktamiş, 2022). Table 10 displays a few of the STEM activities and describes the relationships to each of the STEM disciplines.

The majority of integrated STEM programming happened over one year (21 studies), with the largest effect size estimates occurring at that dosage ($d=0.89$). Additionally, a statistically significant impact occurred at the longest term of four years ($d=0.87$). However, short-cycle STEM programming, occurring over a few weeks,

Table 10 Selected STEM activities and relationships among STEM disciplines (adapted and modified from Hiğde & Aktamiş, 2022)

Activity	Science	Technology	Engineering	Mathematics
Recycled Car	Understanding and application of recycled materials and force of air pressure	Using recycled materials, designing wheels, vehicle body type, and other components	Designing of a recycled car	Measurement and calculations of force, air pressure, and motion
Solar Oven	Temperature and heat energy	Understanding of space technology	Designing a heat-conserving solar oven	Calculating the relationship between temperature and time under sunlight, understanding of greenhouse effect
Biosphere	Understanding and importance of biosphere	Designing a self-sufficient biosphere and preserving biodiversity	Designing a biosphere to preserve biodiversity	Measurement and calculations of temperature, water, and living species in biosphere

short-cycle STEM programming reported a large effect size estimate ($d=0.80$). This is similar to the reported finding of Kazu and Kurtoglu Yalcin (2021), which reported a significantly higher impact for short STEM program interventions (2–5 weeks), emphasizing the importance of short-cycle programs and extracurricular STEM initiatives. These findings support the potential short-term and long-term impact of integrated STEM programming on student achievement and support the need for further research.

Analyses of effect size estimates of the remaining moderators are described. The subject area of assessment, there was a large significant, positive effect estimate ($d=2.02$) using ELA assessments. This may be suggestive that students with the strongest reading abilities have access to more resources or simply put, better students are getting involved in STEM programs. On the other hand, math assessments showed a relatively small effect ($d=0.34$). ELA and math assessments are the preferred measures of assessing academic achievement, as most studies used standardized state tests in both subjects to determine impact on student achievement. This finding is surprising with deeper investigation and further research needed to better analyze these results. More research must also be conducted to explain the substantial differences in effect size estimates between state ($d=0.81$) and local ($d=14$) assessments. Given the strict, standards-driven approach evident in state assessments, the effect size estimates carry more weight than the local assessment estimates, which are locally regulated at the school level and often exhibit issues in validity and reliability. Furthermore, data source is an area of subsequent research with differences in effect size estimates ranging from $d=0.21$ (publications) and $d=0.87$ (evaluation reports). We hypothesize evaluation reports have potential motivation to report more promising or better results to due pressures involving funding. Lastly, there was variability in effect size estimates for publication year with a non-significant decline in 2022 results demonstrating a non-significant relationship between year and effect size estimates ($p>0.05$). However, if 2022, which is based on one study that was likely impacted by the COVID-19 pandemic, is removed, the association between year and effect size estimate increases to $r=0.345$, $R^2=0.119$, $p<0.001$. More research needs to be conducted to determine the effect of integrated STEM education on student achievement as a function of year of implementation and publication.

Student achievement and integrated STEM participation

The second research question was to determine what differences exist in academic achievement between students participating in STEM education programs compared to

students participating in a traditional setting. The overall effect size estimate demonstrated heterogeneity, with a large positive significant effect across the studies Cohen's $d=0.558$, $p<0.001$, $CI_{95}[0.514: 0.603]$. Specifically, this indicates that students benefit from their participation in STEM, and the average STEM student will outperform approximately 70% of their same-age, same-grade peers who are not in STEM programming. A meta-analysis of STEM education's impact on student achievement across all grade levels, not only isolating middle school, found statistically higher ($g=1.150$) using a random-effects model (Kazu & Kurtoglu Yalcin, 2021). Similar meta-analyses reported from Turkish studies have reported varying effect sizes ranging from small to large, indicating a weak to strong effect (Ayverdi & Öz-Aydın, 2020; Saraç, 2018; Yüceliyiğit & Toker, 2021). Researchers reporting on mostly U.S. empirical studies have found consistently moderate effect sizes: $d=0.63$ (Becker & Park, 2011); $d=0.62$ (D'Angelo et al., 2014); $d=0.46$ (Belland et al. 2017). Our findings indicate a moderately strong impact on middle school achievement aligning with similar U.S. studies analyzing all grade levels. However, prior research has not analyzed the impact of STEM integration on middle school achievement only, nor has it sought to determine the impact on URM students, particularly students of color.

Tending to opportunity gaps into college and beyond

Lastly, this meta-analysis sought to provide deeper insight into the differences in academic achievement between URM students or marginalized students participating in STEM programming to similar students in a traditional setting. Students from URM groups have a higher risk of dropping out of STEM education programs earlier when not exhibiting success, determined by good grades or positive feedback from peers and teachers. The use of "early warning systems" to catch struggling or "at-risk" students early, before they stop participating in STEM programs, is important to ensure URM group retention in STEM education programs in middle school and high school (Bernacki et al., 2020). In addition, addressing opportunity gaps related to race is a systemic problem in schools requiring educators, teachers, and staff members to gain knowledge on how to address this gap. Milner (2020) states a three-pronged approach to attending to the racial opportunity gap through building knowledge of: (1) their own racial identity and their students, (2) their own experiences with racism and their students, and (3) how experiences with racial discrimination create and contribute to trauma, which greatly influences student learning and, subsequently, achievement. Milner (2020) suggests an Opportunity Gap Framework that

focuses on how educators conceptualize and reflect on their teaching and learning by focusing on providing opportunities and experiences for students over valuing outcomes (i.e., achievement, test scores). The main takeaway from this principle attends to the interrelatedness of the achievement and opportunity gaps—student achievement improves when opportunity gaps are addressed (Milner, 2020).

Educational frameworks and programs similar to integrated STEM education initiatives need to attend to both students' social-emotional needs and peer relationship building, not merely supplemental content alone. Van Sickle et al. (2020) analyzed the impact of comprehensive (i.e., social networking and peer relationship-building opportunities) versus supplemental (i.e., math content support) STEM programming on the achievement of college students in STEM majors. They determined for URM students, comprehensive programming was associated with substantial learning gain with supplemental instruction alone having little effect on student achievement. For non-URM students, the opposite was found—student learning gains occurred mostly during supplemental instruction. The concept of marginalized and URM students needing social connection and the feeling of belonging in order to attain academic success is known among researchers (McGee, 2021; Milner, 2020; Williams et al., 2019), and we postulate it applies to students participating in STEM programs from kindergarten through high school and beyond!

The lack of representation of empirical studies on the impact of integrated STEM education on students of color sparks debate—is it students of color are not being included in studies, or is it students of color are not participating in STEM-related programming? More research needs to be conducted to answer this question.

Regardless of the root cause for the lack of research on the academic performance among students of color, we need to increase URM student participation and retention in integrated STEM programs, thus increasing opportunities for students. Several methods increase and retain students in STEM and STEM-related (e.g., computer science education, robotics, math extracurricular programs, etc.) programs and subsequently increase engagement. The intentional over-recruitment of students of color preparing for the anticipated high mortality of students participating in STEM programming is one solution. However, overcompensating does not solve the core problem: Why are students of color dropping out of STEM programs and, at the college level, leaving STEM and STEM-related majors? Further research needs to be conducted to determine the impetus behind high student STEM attrition rates among students of color. However, many researchers are focusing not on the

“why” but on “how” to include all people, regardless of background, in advancing technologies and driving new innovations. In *STEM Education for the Future: A Vision Report*, the National Science Foundation's progressive vision statement for STEM education experts cite a dire need for STEM role models for Black and Hispanic youth, a population projected to encompass half of all school children by 2060 (NSF, 2020). Researchers adamantly discuss three top priorities synthesized from the triangulation of individual experts and students, the NSF's 10 Big Ideas for Future Investment, and the National Science Board's Vision 2030 that ensure all learners are prepared and have the skills to succeed in STEM careers.

Priority One: All learners at all stages of their educational pathways must have access to and opportunities to choose STEM careers and contribute to the innovation economy.

Priority Two: We must build an ethical workforce with future-proof skills.

Priority Three: We must ensure that the appropriate technological innovations make it into learning spaces, whether face-to-face classrooms or not, guided by educators who understand how modern technology can affect learning and how to use technology to enhance context and enrich learning experiences for students.

(NSF, 2020, p. 12).

The first priority sparks the most difficult challenge through the lens of educational equity, with high-quality STEM education being inherently unequal, with a student's family income and zip code being the biggest predictor of STEM program quality in kindergarten through high school. Poor and under-resourced communities, both rural and urban, are left behind, struggling to make a positive impact on student outcomes. The *STEM Education for the Future: A Vision Report* proposes several actions to “create opportunities for all students to receive an accessible and high-quality STEM education and help them foster a love and curiosity for science and mathematics from an early age” (NSF, 2020, p. 13). We must challenge past beliefs by making equally accessible and sustainable program changes, train and incentivize STEM educators, and, in a particular effort to reach URM students and students of color, use culturally relevant teaching practices and context-appropriate learning experiences. The use of culturally relevant STEM learning practices (Thevenot, 2022) and cultural competence among teachers and educational staff members (Estrada et al., 2016) are evidence-based strategies shown to increase both student interest and commitment to STEM education. A challenge to ensuring equitable access to opportunities in STEM education is the presence of bias, which creates an uninviting environment and contributes to STEM attrition, particularly among college students

pursuing STEM majors. Gender bias, or implicit bias against women in STEM, has been long known; however, racial bias or bias toward first-generation students or students experiencing poverty is a more recent topic among researchers (Pusey, 2020). Further investigation is needed to determine the influence of staff, teachers, and institutional biases on URM students and, in particular, students of color and STEM program participation at the middle school level.

Experts have devised similar plans to increase URM participation and retention in STEM education at the undergraduate level. In *Improving Underrepresented Minority Student Persistence in STEM*, researchers propose why the STEM pipeline and other academic pathways leak more among URM than among Asian or White students (Estrada et al., 2016). Using Lewin's approach to change, Estrada et al. (2016) posits five recommendations to increase STEM persistence in college. They suggest creating strategic partnerships, attending to student resource disparities, using interventions that increase students' interests and commitment in STEM fields, and focusing on removing institutional barriers (Estrada et al., 2016). The positive correlation between a student's sense of belonging and academic success and motivation is long known (Freeman et al., 2007), with students who feel they belong being more apt to report a purpose and value in their work and higher self-efficacy (Verschelden, 2017).

URM students suffer from gender- and race-related stereotypes affecting participation in academic interventions and programs such as STEM education initiatives, with lack of participation and retention, not an academic problem but a social problem (Williams et al., 2019; Van Sickle et al., 2020). Gender-related stereotypes about one's intellectual ability emerge as early as the age of five, in the way that children tend to perceive males as brilliant and smart and females less so (Bian et al., 2017). Ever-present racial stereotyping is common, with researchers reporting that students of color experience racial microaggressions from instructors and peers, with Black students even more exposed to racial stereotypes (Lee et al., 2020). On the other hand, Asian and Asian American students, as another group of minority students, are facing so-called "positive stereotypes", namely they are often assumed to be good at math and science, extremely hardworking and competent, but disliked at the same time (Lee et al., 2017, p. 225). Though some investigations found that this kind of "positive stereotype" posits students' academic performance, it can also lead to pressure and anxiety due to high or even unrealistic expectations. In response, these students "miss out on important aspects of life" to meet such expectations (McGee et al., 2017, p. 226).

Stereotypes can be harmful to students' motivation and learning outcomes. According to Cheryan et al. (2011), when a stereotype mismatches with a student's self-concept, it will negatively affect an adolescent's interest and motivation in STEM, which becomes a barrier to their entry into STEM-related domains. The negative impact of a mismatch between the social stereotype and a student's self-perception of their self-efficacy in STEM fields is particularly harmful to URM students. After URM students are recruited in STEM domains, the existence of racial microaggressions negatively affects their emotions, confidence, and retention. Subsequently, racial stereotyping contributes to the increased number of STEM students of color leaving STEM college majors, and we propose that may be a contributive reason for the lack of students of color participating in middle school STEM education programs.

Many research works have confirmed that URM students' sense of belonging affects their decision to pursue a career in STEM, as well as their persistence in related areas (Rainey et al., 2018). Female high school students are less likely to report fitting in or feeling accepted in STEM courses than their male peers, but female students reporting a sense of belonging had increased intentions to major in STEM in college (Ito & McPherson, 2018). Researchers are attempting to identify the factors affecting the sense of belonging and, accordingly, seeking ways of intervention. McKoy (2019) reported that for African American engineering students, the lack of role models, namely faculty members of the same race, hinders their will to continue their study in this field. This indicates that instructor–student homophily, or the extent to which students consider their instructors to share similar attitudes (e.g., shared beliefs and values) and backgrounds (e.g., shared experiences) (McCroskey et al., 2006), can affect their engagement and persistence in STEM areas.

Much recent work discusses the impact of instructor–student (or mentor–mentee) homophily on students' participation and persistence in educational institutions. First, perceived instructor homophily is strongly related to students' willingness to participate in class (Myers et al., 2009)—the more the instructor shares similarity with the students, the more credible they will be perceived. The credibility and authenticity of teachers have a positive impact on student's motivation to learn (Wheelless et al., 2011) and have a strong influence on URM students. As Spears (2016) and Kricorian et al. (2020) stated, if URM students feel supported by the environment where they have meaningful, positive contact with faculty members and instructors, they will become more likely to persist in school, especially when STEM URM

students are mentored by those of their same gender and ethnicity.

URM faculty members play a crucial role in enhancing diversity and inclusivity in educational institutions at the college level with the increasing spotlight. Miriti (2020) states, “there has been much investment in diversifying the STEM workforce, but scholars of color continue to be strongly underrepresented” (p. 4). According to the National Center for Education Statistics (2022), the percentage of faculty of color in degree-granting post-secondary institutions hardly increased from the year of 2018 (24.4%) to 2020 (25.8%).

Similarly, the percentage of minority teachers has not increased significantly from 1999 to 2018, either. In recent years, the percentage of Black and Native American teachers has dropped, exemplifying an ever-present problem (National Center for Education Statistics, 2019; U.S. Dept. of Education, 2021). Though non-White, non-male, and first-generation faculty are more involved in diversity and inclusivity-related activities through, for instance, recruiting URM students and faculty, participating in diversity-focused academic works, serving on inclusivity committees, etc., their efforts are still restricted due to insufficient resources (Jimenez et al., 2019). In other words, instead of not having enough knowledge or training, the real barrier to their engagement in diversity and inclusion-related activities is not regarded as a part of their professional evaluation criteria.

Thus, how could educators reduce the impact of stereotypes and improve the sense of belonging and self-efficacy of URM students in STEM? The effect of role models should never be ignored. When assigning students to mentors, their preferences and matching background with faculty should be considered, and we propose ensuring teacher–student homophily within middle school STEM contexts. Highlighting the achievements of STEM professionals from diverse backgrounds and utilizing digital media to increase URM students’ exposure to those role models are also promising ways to reduce their concerns about being recruited in STEM domains (Master & Meltzoff, 2020) and solidify their sense of belonging and STEM self-efficacy (Kricorian et al., 2020). As an implication of this work, we propose collaborative and focused attention among all stakeholders (e.g., teachers, practitioners, educational administrators, policymakers, and families) on the societal and cultural factors impacting both URM students’ participation and retention in integrated STEM education programs.

In summary, many factors contribute to the lack of students of color (i.e., Black, Hispanic, Multi-racial) participating in STEM education programs in K-12, contributing to the smaller proportion of undergraduates

of color majoring in and graduating from STEM-related fields. Racial stereotyping, biases among educational faculty, the inadequacy of culturally responsive teaching practices, and students lacking a sense of belonging all contribute to the opportunity gap evidenced among URM students. To ensure all students are exposed to integrated STEM education and academic interventions in general, attention needs to focus on attending to social and cultural factors. This work highlights the lack of empirical studies on the impact of integrated STEM education among URM populations and, in particular, students of color. In addition, our investigation further exposes the opportunity gap evident among URM students calling attention to the need for interventions attending to cultural inclusivity, attention to social-emotional wellness, sense of belonging, and awareness of biases.

Conclusion

The findings indicate that integrated STEM programming in middle school has a positive, statistically significant effect across multiple grade levels, particularly 8th grade with integrative STEM programming interventions most impactful at dosages of one academic year or over the course of four years. Due to the lack of empirical studies at various levels of STEM integration, further research needs to be conducted. However, the large effect size of greater than one standard deviation ($d=1.09$) across 24 independent samples is supportive of the positive impact full STEM integration can have on student achievement over other integration levels (e.g., S-E, S-M, S-E-M). ELA assessments showed the greatest impact on student achievement suggesting that students with the strongest reading abilities have access to more resources or simply put, better students are getting involved in STEM programs. A deeper investigation into students’ performance by assessment subject area is warranted to provide insight into these findings. In addition, subsequent research on the following moderators is needed: publication year, assessment type (i.e., state or local), and data source (i.e., evaluation report, publication, dissertation/thesis).

Students in middle school overall benefit from STEM program participation, with the average STEM student outperforming approximately 70% of their same-age, same-grade peers who are not participating in STEM programming. In particular, URM students benefit even more from quality integrated STEM education initiatives, but there is one caveat—they must be given the opportunity. This work highlights the lack of empirical studies on URM performance suggesting insufficient student participation and exposure to middle school integrated STEM initiatives. We discuss the need for collaborative and focused attention on the societal (e.g., racial

stereotyping) and cultural (e.g., lack of cultural competence among educators and faculty) factors impacting both URM student participation and retention in integrated STEM education programs.

Abbreviations

CK	Content knowledge
ELA	English Language Arts
PBL	Project-based learning
PD	Professional development
PK	Pedagogical knowledge
RVE	Robust variance estimates
STEM	Science, technology, engineering, and mathematics
TK	Technological knowledge
URM	Underrepresented minority groups

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40594-023-00425-8>.

Additional file 1: Research articles in the meta-analysis.

Acknowledgements

The authors would like to express their gratitude to: Dr. Carmen Thomas-Browne for her thoughts and feedback; Xinyu Yang for her guidance within the initial draft of the Discussion attending to diversity, equity, and inclusion; Dr. Kenneth R. Koedinger and the Personalized Learning Squared (PLUS) team for their support; and the anonymous reviewers for providing their welcomed comments to refine this manuscript.

Author contributions

DT designed the research, investigated articles, and drafted the work as an extension of her prior doctoral research. KL researched analyzed articles, interpreted the data, and provided guidance in creating the work. Both authors read and approved the final manuscript.

Authors' information

Danielle R. Thomas is a Systems Scientist and Special Faculty member at Carnegie Mellon University in the Human-Computer Interaction Institute within the School of Computer Science. A former middle school teacher and school administrator, her research interests include STEM education, human-computer interaction, and learning engineering with an emphasis on ensuring educational equity.

Karen H. Larwin is a full Professor at Youngstown State University in the cross-disciplinary Education Leadership Program. Her research focuses on quantitative applications in Structural Equation Modeling and Hierarchical Linear Modeling, as well as in learning theories focused on educational equitable opportunities for underrepresented groups.

Funding

Submission and publication fees are supported by Carnegie Mellon University.

Availability of data and materials

Not applicable.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Carnegie Mellon University, Pittsburgh, PA, USA. ²Youngstown State University, Youngstown, OH, USA.

Received: 29 July 2022 Accepted: 2 May 2023

Published online: 15 June 2023

References

- Adams, E. L. (2021). The effect of a middle grades STEM initiative on students' cognitive and non-cognitive outcomes. *Studies in Educational Evaluation*, 68, 100983. <https://doi.org/10.1016/j.stueduc.2021.100983>
- Ayverdi, L., & Öz Aydin, S. (2020). Meta-analysis of studies examining the effect of STEM education on academic success. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education (EFMED)*, 14(2), 840–888. <https://doi.org/10.17522/balikesirnef.755111>
- Barakos, L., Lujan, V., & Strang, C. (2012). Science, technology, engineering, mathematics (STEM): Catalyzing change amid the confusion. *RMC Research Corporation, Center on Instruction*. <https://eric.ed.gov/?id=ed534119>
- Becker, K. H., & Park, K. (2011). Integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A meta-analysis. *Journal of STEM Education: Innovations and Research*, 12(5). <https://www.jstem.org/jstem/index.php/JSTEM/article/view/1509>
- Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, 87(2), 309–344. <https://doi.org/10.3102/0034654316670999>
- Bernacki, M. L., Chavez, M. M., & Uesbeck, P. M. (2020). Predicting achievement and providing support before STEM majors begin to fail. *Computers & Education*, 158, 103999. <https://doi.org/10.1016/j.compedu.2020.103999>
- Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, 355(6323), 389–391. <https://doi.org/10.1126/science.aah6524>
- Blotnicky, K. A., Franz-Odenaal, T., French, F., & Joy, P. (2018). A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *International Journal of STEM Education*, 5(1), 1–15. <https://doi.org/10.1186/s40594-018-0118-3>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to Meta-Analysis* (2nd ed.). Wiley.
- Bracey, J. M. (2013). The culture of learning environments: Black student engagement and cognition in math. In J. Leonard & D. B. Martin (Eds.), *The brilliance of black children in mathematics: Beyond the numbers and toward new discourse* (pp. 171–194). Information Age Publishing Inc.
- Bryan, L., & Guzey, S. S. (2020). K-12 STEM Education: An overview of perspectives and considerations. *Hellenic Journal of STEM Education*, 1(1), 5–15. <https://doi.org/10.51724/hjstemed.v1i1.5>
- Bybee, R. W. (1997). *Achieving scientific literacy*. Heinemann.
- Bybee, R. W. (2010). Advancing STEM education: A 2020 vision. *Technology and Engineering Teacher*, 70(1), 30. <https://doi.org/10.1007/s11422-015-9713-5>
- Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. NSTA press.
- Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness. *Colorado Springs, Co: BSCS*, 5, 88–98. <https://fremonths.org/ourpages/auto/2008/5/11/1210522036057/bscs5efullrepor t2006.pdf>
- Chen, X. (2013). *STEM Attrition: College Students' Paths Into and Out of STEM Fields* (NCES 2014–001). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. <https://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=2014001rev>
- Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science*, 2(6), 656–664. <https://doi.org/10.1177/1948550611405218>
- Chine, D. R. (2021). *A Pathway to Success? A Longitudinal Study Using Hierarchical Linear Modeling of Student and School Effects on Academic Achievement in a Middle School STEM Program* [Doctoral dissertation, Youngstown State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=yysu1619525650839685

- Chine, D. R., Brentley, C., Thomas-Browne, C., Richey, J. E., Gul, A., Carvalho, P. F., Branstetter, L., & Koedinger, K. R. (2022). Educational equity through combined human-AI personalization: A propensity matching evaluation. In *International Conference on Artificial Intelligence in Education: 23rd International Conference, AIED 2022, Durham, UK, July 27-31, 2022, Proceedings, Part 1* (pp. 366-377). Cham: Springer International Publishing.
- Christensen, R., Knezek, G., and Tyler-Wood, T. (2015). Alignment of hands-on STEM engagement activities with positive STEM dispositions in secondary school students. *Journal of Science Education and Technology*, 24(6), 898–909. <https://www.jstor.org/stable/43867753>
- Cohen, A. (2020). "5 Reasons Why STEM Career Training Should Start in Middle School". Retrieved on June 17, 2020: <https://everfi.com/insights/blog/5-reasons-stem-career-training-start-middle-school/>
- D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E. & Haertel, G. (2014). Simulations for STEM learning: systematic review, and meta-analysis. Menlo Park, CA: SRI International. <https://www.sri.com/wp-content/uploads/2021/12/simulations-for-stem-learning-brief.pdf>
- Dewey, J. (1897). My Pedagogic Creed. *School Journal*, 54(3), 77–80.
- Dugger, W. E. (2010). Evolution of STEM in the United States. In *6th Biennial International Conference on Technology Education Research* (Vol. 10). <https://www.academia.edu/download/47244343/AustraliaPaper.pdf>
- Estrada, M., Burnett, M., Campbell, A. G., Campbell, P. B., Denetclaw, W. F., Gutiérrez, C. G., Hurtado, S., John, G. H., Matsui, J., McGee, R., Okpodu, C. M., Robinson, T. J., Summers, M. F., Werner-Washburne, M., & Zavala, M. (2016). Improving underrepresented minority student persistence in STEM. *CBE Life Sciences Education*. <https://doi.org/10.1187/cbe.16-01-0038>
- Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education*, 75(3), 203–220. <https://doi.org/10.3200/JEXE.75.3.203-220>
- Fullan, M. (2007). *Leading in a culture of change*. John Wiley & Sons. <https://eric.ed.gov/?id=ED467449>
- Gazibeyoglu, T., & Aydin, A. (2019). The effect of STEM-based activities on 7th grade students' academic achievement in force and energy unit and students' opinions about these activities. *Universal Journal of Educational Research*, 7(5), 1275–1285. <https://doi.org/10.13189/ujer.2019.070513>
- Gelbgiser, D., & Alon, S. (2016). Math-oriented fields of study and the race gap in graduation likelihoods at elite colleges. *Social Science Research*, 58, 150–164. <https://doi.org/10.1016/j.ssresearch.2016.03.005>
- Giasi, T. (2018). *Post-Program Implementation of Integrated STEM Instruction: A Qualitative Case Study* [Doctoral Dissertation, Ohio State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=osu1525688091215031
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8. <https://doi.org/10.3102/0013189X00501003>
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Sage Publications, Incorporated.
- Gonzalez, H., & Kuenzi, J. (2012). Science, technology, engineering, and mathematics (STEM) education: A primer. Congressional Research Service, 7(5700), 1–34. <https://fas.org/spp/crs/misc/R42642.pdf>
- Han, S., Capraro, R., & Capraro, M. M. (2015). How science, technology, engineering, and mathematics (STEM) project-based learning (PBL) affects high, middle, and low achievers differently: The impact of student factors on achievement. *International Journal of Science and Mathematics Education*, 13(1), 1089–1113. <https://doi.org/10.1007/s10763-014-9526-0>
- 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>
- Harlan, J. M., Pruet, S. A., Van Haneghan, J., & Dean, M. D. (2014). Using curriculum-integrated engineering modules to improve understanding of math and science content and STEM attitudes in middle grade students. In *2014 ASEE Annual Conference & Exposition* (pp. 24–1325). <https://doi.org/10.18260/1-2-23258>
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press. <https://doi.org/10.1016/c2009-0-03396-0>
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). "Robust variance estimation in meta-regression with dependent effect size estimates": Erratum. *Research Synthesis Methods*, 1(2), 164–165. <https://doi.org/10.1002/jrsm.17>
- Hiğde, E., & Aktamış, H. (2022). The effects of STEM activities on students' STEM career interests, motivation, science process skills, science achievement and views. *Thinking Skills and Creativity*, 43, 101000.
- Honey, M., Pearson, G., & Schweingruber, H. (Eds.). (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research*. Committee on Integrated STEM Education; National Academy of Engineering; National Research Council. <https://doi.org/10.17226/18612>
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage Publications. <https://doi.org/10.4135/9781483398105>
- Irvine, J. J. (2010). Foreword to *Culture, Curriculum, and Identity in Education*. ed. H. R. Milner. Palgrave Macmillan, 12. https://doi.org/10.1057/9780230105669_1
- Ito, T. A., & McPherson, E. (2018a). Factors influencing high school students' interest in pSTEM. *Frontiers in Psychology*, 9, 1535. <https://doi.org/10.3389/fpsyg.2018.01535>
- Izgi, S., & Kalaycı, S. (2020). The effect of the STEM approach based on the 5E model on academic achievement and scientific process skills: the transformation of electrical energy. *International Journal of Education Technology and Scientific Research*, 5(13), 1578–1628. <https://doi.org/10.35826/ijetsar.259>
- Jimenez, M. F., Laverty, T. M., Bombaci, S. P., Wilkins, K., Bennett, D. E., & Pejchar, L. (2019). Underrepresented faculty play a disproportionate role in advancing diversity and inclusion. *Nature Ecology & Evolution*, 3(7), 1030–1033. <https://doi.org/10.1038/s41559-019-0911-5>
- Johnson, B. T. (1989). *DSTAT: Software for the meta-analytic review of research literatures*. Lawrence Erlbaum Associates.
- Kazu, I. Y., & Kurtoglu Yalcin, C. (2021). The effect of STEM education on academic performance: A meta-analysis study. *Turkish Online Journal of Educational Technology-TOJET*, 20(4), 101–116. <https://eric.ed.gov/?id=EJ1313488>
- Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*. <https://doi.org/10.1186/s40594-016-0046-z>
- Kim, K. H., & Zabelina, D. (2015). Cultural bias in assessment: Can creativity assessment help?. *The International Journal of Critical Pedagogy*, 6(2). <https://core.ac.uk/outputs/234819766>
- Kricorian, K., Seu, M., Lopez, D., Ureta, E., & Equils, O. (2020). Factors influencing participation of underrepresented students in STEM fields: Matched mentors and mindsets. *International Journal of STEM Education*, 7(1), 1–9. <https://doi.org/10.1186/s40594-020-00219-2>
- 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*. [Doctoral dissertation, Wilmington University]. <https://www.proquest.com/docview/873247689/fulltextPDF/ED787F7D1AC04FBCPQ/1?accountid=9902>
- Le Thi Thu, H., Tran, T., Trinh Thi Phuong, T., Le Thi Tuyet, T., Le Huy, H., & Vu Thi, T. (2021). Two decades of STEM education research in middle school: A bibliometrics analysis in scopus database (2000–2020). *Education Sciences*, 11(7), 353. <https://doi.org/10.3390/educsci11070353>
- Lee, M. J., Collins, J. D., Harwood, S. A., Mendenhall, R., & Hunt, M. B. (2020). "If you aren't White, Asian or Indian, you aren't an engineer": Racial micro-aggressions in STEM education. *International Journal of STEM Education*, 7(1), 1–16. <https://doi.org/10.1186/s40594-020-00241-4>
- Lesseig, K., Slavit, D., & Nelson, T. H. (2017). Jumping on the STEM bandwagon: How middle grades students and teachers can benefit from STEM experiences. *Middle School Journal*, 48(3), 15–24. <https://doi.org/10.1080/00940771.2017.1297663>
- Lin, L., & Chu, H. (2018). Quantifying publications bias in meta-analysis. *Biometrics*, 74(3), 785–794. <https://doi.org/10.1111/biom.12817>
- Lounsbury, J. H. (2010). This we believe: Keys to educating young adolescents. *Middle School Journal*, 41(3), 52–53. <https://doi.org/10.1080/00940771.2010.11461722>
- Lynch, S., Burton, E., Behrend, T., House, A., Ford, M., Spillane, N., Matray, S., Han, E., & Means, B. (2017). Understanding inclusive STEM high schools as opportunity structures for underrepresented students: Critical

- components. *Journal of Research in Science in Science Teaching*, 55(5), 712–748. <https://doi.org/10.1002/tea.21437>
- Major, C., & McDonald, E. (2021). Developing instructor TPACK: A research review and narrative synthesis. *Journal of Higher Education Policy and Leadership Studies*, 2(2), 51–67. <https://doi.org/10.52547/johepal.2.2.51>
- Maltese, A. V., Melki, C. S., & Wiebke, H. L. (2014). The nature of experiences responsible for the generation and maintenance of interest in STEM. *Science Education*, 98(6), 937–962. <https://doi.org/10.1002/sce.21132>
- Master, A., & Meltzoff, A. N. (2020). Cultural stereotypes and sense of belonging contribute to gender gaps in STEM. *International Journal of Gender, Science and Technology*, 12(1), 152–198. <https://genderandset.open.ac.uk/index.php/genderandset/article/view/674>
- Masters of Gravity. (n.d.) Teacher Guide. Western Reserve Public Media. <https://westernreservepublicmedia.org/gravity/gravity.pdf>
- McClure, E. (2017). More than a foundation: Young children are capable of STEM learners. *National Association for the Education of Young Children*, 72(5), 83–89. <https://doi.org/10.2307/90015862>
- McClure, E., Guernsey, D., Clements, S., Bales, J., Nichols, N., Kendall-Taylor, & Levine, M. (2017). STEM Starts Early: Grounding Science, Technology, Engineering, and Math Education in Early Childhood. The Joan Ganz Cooney. Center at Sesame Workshop. www.joanganzcooneycenter.org/wp-content/uploads/2017/01/jgcc_stemstartsearly_final.pdf
- McCroskey, L. L., McCroskey, J. C., & Richmond, V. P. (2006). Analysis and improvement of the measurement of interpersonal attraction and homophily. *Communication Quarterly*, 54(1), 1–31. <https://doi.org/10.1080/01463370500270322>
- McGee, E. O. (2021). *Black, brown, bruised: How racialized STEM education stifles innovation*. Harvard Education Press. <https://doi.org/10.1177/15210251211018178>
- McGee, E. O., Thakore, B. K., & LaBlance, S. S. (2017). The burden of being “model”: Racialized experiences of Asian STEM college students. *Journal of Diversity in Higher Education*, 10(3), 253. <https://doi.org/10.1037/dhe0000022>
- McKoy, T. L. (2019). *A Qualitative Study of African American Female Engineering College Students' Intersecting Identities, Sense of Belonging, and Intent to Persist* (Doctoral dissertation, Tennessee State University). <https://eric.ed.gov/?id=ED601803>
- Milner, H. R. (2020). *Start where you are, but don't stay there: Understanding diversity, opportunity gaps, and teaching in today's classrooms* (2nd ed.). Harvard Education Press. <https://doi.org/10.5860/choice.48-7061>
- Miriti, M. N. (2020). The elephant in the room: Race and STEM diversity. *BioScience*, 70(3), 237–242. <https://doi.org/10.1093/biosci/biz167>
- Moore, T. J., Johnston, A. C., & Glancy, A. W. (2020). STEM integration: A synthesis of conceptual frameworks and definitions. In *Handbook of Research on STEM Education* (pp. 3–16). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780429021381-2/stem-integration-tamara-moore-amanda-johnston-aran-glancy>
- Morales, M. P. E., Avilla, R. A., Sarmiento, C. P., Anito, J. C., Jr., Elipane, L. E., Palisoc, C. P., Palomar, B. C., Ayuste, T. O., & Ramos-Butron, B. (2022). Experiences and practices of STEM teachers through the lens of TPACK. *Journal of Turkish Science Education*, 19(1), 233–252. <https://doi.org/10.36681/tused.2022.1120>
- Moreno, N., Tharp, B., Vogt, G., Newell, A., & Burnett, C. (2016). Preparing students for middle school through after-school STEM activities. *Journal of Science Education and Technology*, 25(6), 889–898. <https://doi.org/10.1007/s10956-016-9643-3>
- Morrison, J. (2006). TIES STEM education monograph series: Attributes of STEM education. Baltimore, MD: Teaching Institute for Essential Science. http://www.wytheexcellence.org/media/STEM_Articles.pdf
- Myers, S. A., Horan, S. M., Kennedy-Lightsey, C. D., Madlock, P. E., Sidelinger, R. J., Byrnes, K., Frisby, B., & Mansson, D. H. (2009). The relationship between college students' self-reports of class participation and perceived instructor impressions. *Communication Research Reports*, 26(2), 123–133. <https://doi.org/10.1080/08824090902861580>
- NAEP. (2019). National Assessment for Academic Progress Report Card. National Center for Education Statistics. (2022). Characteristics of Public School Teachers. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved June 17, 2022, from <https://nces.ed.gov/programs/coe/indicator/clr>
- NSF. (2020). *STEM Education for the Future - 2020 Visioning Report*. National Science Foundation. <https://www.nsf.gov/ehf/Materials/STEM%20Education%20for%20the%20Future%20-%202020%20Visioning%20Report.pdf>
- O'Brien, L. T., Blodorn, A., Adams, G., Garcia, D. M., & Hammer, E. (2015). Ethnic variation in gender-STEM stereotypes and STEM participation: An intersectional approach. *Cultural Diversity and Ethnic Minority Psychology*, 21, 169–180. <https://doi.org/10.1037/a0037944>
- Pew Research Center. (2021). “STEM Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity.” https://www.pewresearch.org/science/wp-content/uploads/sites/16/2021/03/PS_2021.04.01_diversity-in-STEM_REPORT.pdf
- Premraj, D., Rudi Thompson, R., Hughes, L., & Adams, J. (2021). Key factors influencing retention rates among historically underrepresented student groups in STEM fields. *Journal of College Student Retention Research, Theory & Practice*, 23(2), 457–478. <https://doi.org/10.1177/1521025119848763>
- Pusey, T. S. (2020). An equal education: Reducing gender bias in STEM education. *UC Merced Undergraduate Research Journal*. <https://doi.org/10.5070/M4122047579>
- Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International Journal of STEM Education*, 5(1), 1–14. <https://doi.org/10.1186/s40594-018-0115-6>
- Roehrig, G. H., Dare, E. A., Ellis, J. A., & Ring-Whalen, E. (2021). Beyond the basics: A detailed conceptual framework of integrated STEM. *Disciplinary and Interdisciplinary Science Education Research*, 3(1), 1–18. <https://doi.org/10.1186/s43031-021-00041-y>
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Sage. <https://doi.org/10.4135/9781412984997>
- Saraç, H. (2018). The effect of science, technology, engineering, and mathematics-stem educational practices on students' learning outcomes: A meta-analysis study. *TOJET: The Turkish Online Journal of Educational Technology*, 17(2), 125–142. <https://eric.ed.gov/?id=EJ1176176>
- Schaldenbrand, P., Lobczowski, N. G., Richey, J. E., Gupta, S., McLaughlin, E. A., Adeniran, A., & Koedinger, K. R. (2021). Computer-supported human mentoring for personalized and equitable math learning. In *International Conference on Artificial Intelligence in Education* (pp. 308–313). Springer, Cham. https://doi.org/10.1007/978-3-030-78270-2_55
- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological pedagogical content knowledge (TPACK) the development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education*, 42(2), 123–149. <https://files.eric.ed.gov/fulltext/EJ868626.pdf>
- Selcen Guzey, S., Harwell, M., Moreno, M., Peralta, Y., & Moore, T. J. (2017). 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>
- Snyder, M. (2018). A century of perspectives that influenced the consideration of technology as a critical component of STEM education in the United States. *Journal of Technology Studies*, 44(2), 42–57. <https://doi.org/10.2307/26730730>
- Spears, J. L. (2016). Experiences of Credibility: Female Instructors of Color at Faith Based Universities. <https://digitalcommons.acu.edu/etd/43>
- Spillane, N., Lynch, S., & Ford, M. (2016). Inclusive STEM high schools increase opportunities for underrepresented students. *The Phi Delta Kappan*, 97(8), 54–59. <https://www.jstor.org/stable/24893337>
- Suran, M. (2021). Keeping Black students in STEM. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.2108401118>
- Thevenot, Y. (2022). Culturally responsive and sustaining STEM curriculum as a problem-based science approach to supporting student achievement for Black and Latinx students. *Research Perspectives in Urban Education*. <https://doi.org/10.33682/bhr8-6z56>
- Tytler, R., Williams, G., Hobbs, L., & Anderson, J. (2019). Challenges and opportunities for a STEM interdisciplinary agenda. *Interdisciplinary Mathematics Education*. https://doi.org/10.1007/978-3-030-11066-6_5
- U.S. Department of Education (2021). *Spring 2019 through Spring 2021 Human Resources component, Fall Staff section*. National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS). Retrieved June 17, 2022, from https://nces.ed.gov/programs/digest/d21/tables/dt21_315.20.asp

- Van Sickle, J., Schuler, K. R., Holcomb, J. P., Carver, S. D., Resnick, A., Quinn, C., Jackson, D.K., Duffy, S.F., Sridhar, N. (2020). Closing the Achievement Gap for Underrepresented Minority Students in STEM: A Deep Look at a Comprehensive Intervention. *Journal of STEM Education: Innovations and Research*, 21(2). <https://jstem.org/jstem/index.php/JSTEM/article/view/2452/2160>
- Verschelden, C. (2017). Bandwidth recovery: Helping students reclaim cognitive resources lost to poverty, racism, and social marginalisation. *Stylus Publishing*. <https://doi.org/10.1080/19496591.2018.1470007>
- Villegas, A. M., & Lucas, T. (2002). Preparing culturally responsive teachers: Rethinking the curriculum. *Journal of Teacher Education*, 53(1), 20–32.
- Wade-Shepherd, A. A. (2016). *The effect of middle school STEM curriculum on science and math achievement scores*. <https://ui.adsabs.harvard.edu/abs/2016PhDT.....172W/abstract>
- Wheless, V. E., Witt, P. L., Maresh, M., Bryand, M. C., & Schrodt, P. (2011). Instructor credibility as a mediator of instructor communication and students' intent to persist in college. *Communication Education*, 60(3), 314–339. <https://doi.org/10.1080/03634523.2011.555917>
- White, D. (2014). What is STEM education and why is it important? *Florida Association of Teacher Educators Journal*, 1(14), 1–9. <http://www.fate1.org/journals/2014/white.pdf>
- Wiliam, D. (2010). Standardized testing and school accountability. *Educational Psychologist*, 45(2), 107–122. <https://doi.org/10.1080/00461521003703060>
- Williams, M. J., George-Jones, J., & Hebl, M. (2019). The face of STEM: Racial phenotypic stereotypicality predicts STEM persistence by—and ability attributions about—students of color. *Journal of Personality and Social Psychology*, 116(3), 416–443. <https://doi.org/10.1037/pspi0000153>
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Sage Publications. <https://doi.org/10.4135/9781412984980>
- Yüceliyigit, S., & Toker, Z. (2021). A meta-analysis on STEM studies in early childhood education. *Turkish Journal of Education*, 10(1), 23–36. <https://doi.org/10.19128/turje.783724>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
