MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2026, 22(9), em2908
ISSN:1305-8223 (online)
OPEN ACCESS Literature Review https://doi.org/10.29333/ejmste/19177

Does physical education and health truly belong to STEM? A meta-analysis on the effects of STEM-PA integrations

Muhammet Usak [1*], Aleksandra Mihajlović [1]

[1] Faculty of Education in Jagodina, University of Kragujevac, Kragujevac, SERBIA

Received 23 June 2026 ▪ Accepted 10 August 2026

Abstract

The integration of physical education, health, and physical activity (PA) into STEM/STEAM education has received increasing attention as a way to promote both academic learning and broader student development. However, findings across individual studies remain inconsistent. This meta-analysis synthesized evidence from 23 independent studies comprising 24 effect sizes and a total sample of 5,238 participants to examine the overall effectiveness of PA-integrated STEM/STEAM interventions and to identify factors that may explain differences in their effectiveness. Effect sizes were calculated using a random-effects model, and moderator analyses were conducted based on educational level, cultural dimension, STEM discipline, outcome type, teaching model, integration type, and intervention duration. The overall analysis indicated a large positive effect of PA-integrated STEM/STEAM interventions on student outcomes (g = 1.05, 95% confidence interval [0.84, 1.26]). Moderator analyses showed that intervention effectiveness varied across all examined moderators. Technology-enhanced and embodied cognition-based teaching models produced the largest effects, while interventions integrating technology with sports contexts were more effective than other integration approaches. Studies conducted in vertical collectivist cultural contexts yielded larger effects than those conducted in vertical individualist contexts. In addition, interventions lasting between three and eight weeks generally produced stronger effects than longer programs. Research findings provide practical guidance for educators and curriculum developers by emphasizing that the effectiveness of PA-integrated STEM/STEAM education depends on purposeful instructional design rather than the inclusion of PA alone.

Keywords: STEPAM, STEM-PEH integration, STE(A)M, meta-analysis, embodied cognition

INTRODUCTION

In the contemporary educational landscape, science, technology, engineering, and mathematics (STEM) disciplines are recognized as the bedrock of national development, economic competitiveness, and social prosperity (Abdelaziz et al., 2025; Ali et al., 2021; Zhan et al., 2022). As the demands of the 21st century shift toward complex problem-solving and innovation, education systems worldwide are prioritizing STEM literacy to prepare a skilled workforce (Zollman, 2012). The transition from STEM to STEAM represents an educational shift from focusing solely on technical competencies to a more holistic, transdisciplinary framework that integrates the arts (Lee, 2021; Wijayanti et al., 2025). This inclusion is designed to break down disciplinary boundaries by adding aesthetic elements and creative design, which foster creativity and imagination in students (Almarcha et al., 2023; Chen et al., 2023). Physical education and health (PEH) has long been positioned as a peripheral discipline in educational reform movements, largely excluded from the interdisciplinary frameworks that have reshaped modern curricula. Parallelly, there is a growing global concern regarding physical inactivity and the rise of sedentary behaviors (SBs) among children and adolescents, which negatively impacts both health and cognitive outcomes (Barbosa et al., 2020). Historically, physical education (PE) and academic instruction have been treated as separate domains, often leading to a reduction in physical activity (PA) time to accommodate rigorous academic testing (Perry et al., 2025).

✉ musaktr@gmail.com **(*Correspondence)** ✉ aleksandra.mihajlovic@pefja.kg.ac.rs

Contribution to the literature

- This study is among the first comprehensive meta-analyses to quantitatively synthesize the effects of integrating physical education, health, and PA into STEM/STEAM education, consolidating previously fragmented experimental findings into a single pooled effect size.
- It systematically examines seven distinct moderator variables (educational level, cultural dimension, STEM discipline, outcome type, teaching model, integration type, and intervention duration), demonstrating that the effect is not uniform but contingent on specific design features of the intervention.
- It provides empirical support for the STEPAM framework by showing that conceptual alignment between bodily movement and STEM content (e.g., embodied cognition-based or technology-enhanced approaches) is a stronger determinant of effectiveness than the mere inclusion of physical activity. The findings identify practical implementation parameters, such as optimal intervention duration and cultural context (vertical collectivist settings), offering actionable, evidence-based guidance for curriculum developers and teacher training programs.

To address these dual challenges, a paradigm shift toward an integrated "whole-child" approach has emerged (Bonner et al., 2024; Perry et al., 2025). Research increasingly suggests that PA is not merely a health-protective factor but a catalyst for cognitive enhancement (Donnelly et al., 2016; Hillman et al., 2008).

STE(A)M-PA and STE(A)M-PEH Integration

In this review, STE(A)M-PA refers to instructional approaches that intentionally integrate PA into STEM or STEAM learning experiences to support students' conceptual understanding through movement-based learning (Bonner et al., 2024; Perry et al., 2025). STE(A)M-PEH is used to describe a broader interdisciplinary approach that integrates PE and health with STEM/STEAM disciplines to promote academic learning alongside students' physical and cognitive development (Da Cruz et al., 2025). Although the concept of STEM emerged earlier, research explicitly integrating PE, health, and PA with STEM education began to gain momentum around 2005. Health education plays a strong integrative role between PE and STEAM disciplines. Integrating concepts from fields such as physiology, nutritional science, and biomechanics into PE classes contributes to the development of both students' scientific thinking skills and health literacy. Highlighted by Hillman et al. (2008), reveals that regular PA has a strong and consistent effect on the brain and cognition. Considering that approximately 80% of children and adolescents do not reach the minimum 60 minutes of moderate-intensity PA recommended by the World Health Organization (Guthold et al., 2020), strengthening the evidence regarding the role of PA on cognitive development and translating these findings into policy is of great importance. This situation demonstrates that PEH courses have a structure that directly contributes not only to motor development but also to cognitive and academic outcomes and therefore have significant potential for integration with STEM education. Castelli et al. (2014) indicated that regular and structured physical activities positively affect not only physical health but also cognitive functions and academic performance. Donnelly and Lambourne (2011) found that the integration of academic content in PE classes such as especially activities that blend science and mathematics concepts with PA improved both students' health indicators and cognitive engagement.

Considering the relationship between PA and cognitive performance from a developmental perspective offers critical point for the pedagogical design of STEM-PEH integration. It is known that the sub-dimensions of executive functions exhibit different developmental patterns; inhibition develops earlier than working memory and therefore shows stronger associations with PA (Best & Miller, 2010). Luo et al. (2023) empirically confirmed this situation; while no significant relationship was found between working memory and moderate-to-vigorous PA, a statistically significant link was found between cardiorespiratory Fitness and working memory. This provides important guidance on which types of physical activities should be paired with which cognitive goals in PEH lessons. For example, it can be predicted that structured physical activities requiring coordination and balance may support inhibition control, while aerobic endurance-focused activities may support working memory. Niederer et al. (2011) also reported similar findings in their longitudinal study examining the relationships between aerobic fitness and motor skills on memory and attention in preschool children. The scientific evidence strongly supports the idea that PE programs can be structured not only for the purpose of developing motor skills, but also as an integrated STEM learning context that supports cognitive and academic development.

SB refers to inactive activities during wakefulness that require low energy expenditure, such as sitting, lying down, or spending time in front of screens. The relationship between SB and cognitive performance is particularly important in discussions of STEM-PEH integration. A systematic review by Carson et al. (2015) similarly revealed mixed results, noting that different

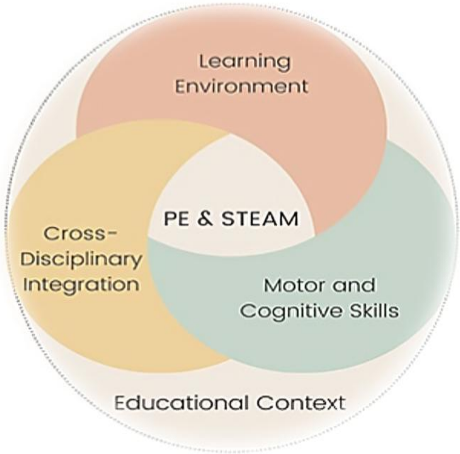

Figure 1. Adapted conceptual model bridging PE and STEAM (Da Cruz et al., 2025)

types of SB can have divergent outcomes. These findings suggest that treating physical inactivity as a homogeneous variable in STEM-PEH integration models can be misleading; rather, the cognitive content of sedentary time must be carefully considered. In this context effective STEM-PEH integration requires not merely increasing movement duration, but consciously designing the type, intensity, and cognitive demands of PA. Jylänki et al. (2022) demonstrated in their systematic review that structured motor skill and PA interventions produced meaningful gains in preschoolers' cognitive and academic outcomes, while Diamond and Ling (2016) further argued that the cognitive engagement embedded within PA, rather than movement alone, is the critical driver of executive function development. In this context STEM-PEH integration holds transformative potential for teacher training, curriculum development, and education policy only when grounded in intentional, cognitively rich pedagogical design.

STEPAM Framework

A central foundation of this analysis is the STEPAM (teacher training in physical activities and embodied learning through STEAM education) framework. Unlike multidisciplinary approaches where subjects are merely juxtaposed, STEPAM seeks a transdisciplinary synergy where physical action and conceptual understanding are inseparable (Da Cruz et al., 2025).

The proposed framework addresses the disconnect between PE and STEAM by stressing teacher training as the main method of integrating the two (Da Cruz et al., 2025) (**Figure 1**). First, it acknowledges that most PE teachers lack the skills to teach STEM-related concepts (Fullerton & Chapman, 2026; Li et al., 2019). STEPAM implements "embodied learning" by teaching educators how to develop exercises wherein learners demonstrate physical manifestations of STEM knowledge, for instance, using balancing benches to learn about dynamic systems theory or applying soccer games to learn about kinematics (Almarcha et al., 2023; Queiruga-Dios et al., 2025). Finally, the framework promotes policy normalization by requiring the allocation of minimum PA hours towards STEAM learning objectives (Da Cruz et al., 2025). Leveraging technology-based platforms like "active science," which tracks activities related to STEM tasks, STEPAM transforms the PE class into a "dynamic laboratory" of learning integration (Da Cruz et al., 2025; Rahmaningrum et al., 2025).

This meta-analytic proposal aims to fill that gap by systematically collecting, coding, and statistically synthesizing all available experimental and quasi-experimental studies that have implemented a STEAM-PA/STEM-PEH integrated teaching approach and measured its effects on student outcomes. The primary research question addresses whether the STEAM-PA/STEM-PEH integrated approach yields measurable improvements in cognitive outcomes (science and mathematics), motor skill development, PA levels, and health literacy, and whether any observed effect sizes differ according to moderator variables such as student age group, educational level, intervention duration, instructional model, or country context.

RQ1: What is the overall effect size of STEAM-PA/STEM-PEH integration on student outcomes across experimental and quasi-experimental studies?

RQ2: Does educational level moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

RQ3: Does cultural dimension moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

RQ4: Does STEM discipline (mathematics, physics/engineering, science/biology, integrated STEM, health/science) moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

RQ5: Does outcome type (academic achievement, STEM attitude, motivation/engagement, higher-order thinking, well-being) moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

RQ6: Does teaching model moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

RQ7: Does integration type (sports-context STEM, PA-embedded STEM, health-integrated STEM) moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

RQ8: Does intervention duration (≤ 2 weeks, 3-8 weeks, 9-18 weeks, > 18 weeks) moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

RQ9: Does publication year moderate the effect of STEAM-PA/STEM-PEH integration on student outcomes?

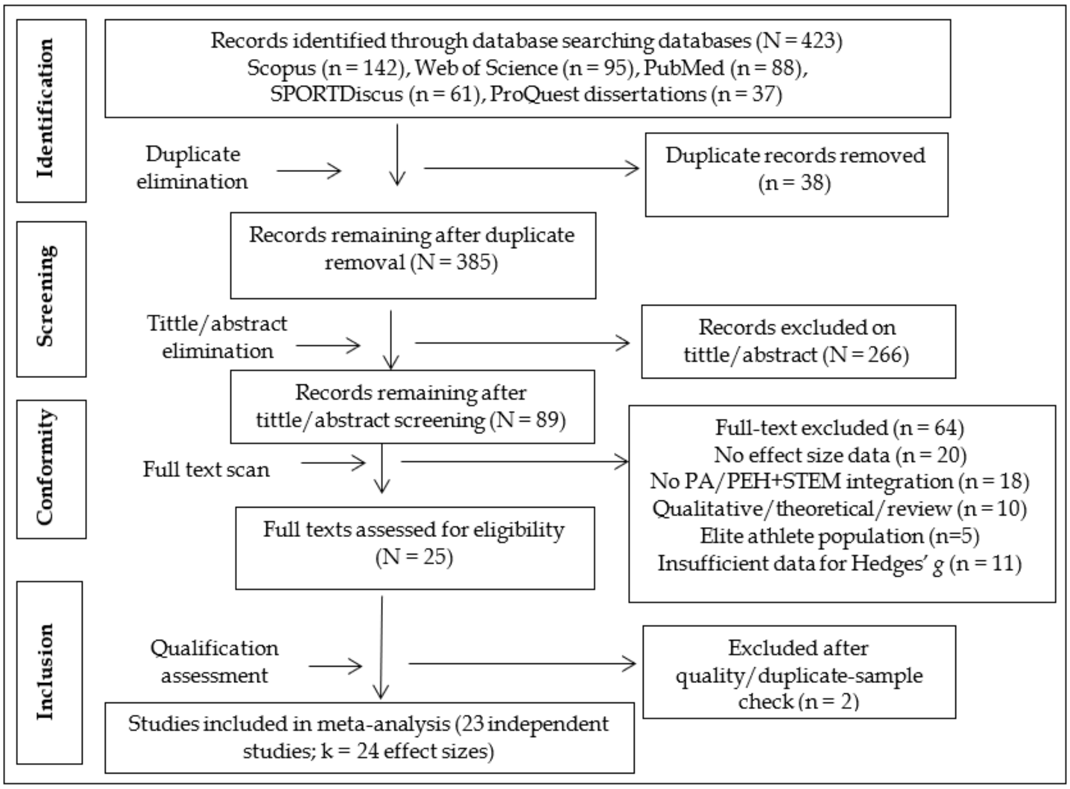

Figure 2. PRISMA flowchart showing the study selection process (the authors' own elaboration)

METHOD

Research Design

This study employed a meta-analytic research design in accordance with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) (Page et al., 2021) guidelines. Meta-analysis is a quantitative method that statistically aggregates effect sizes from independent studies to yield a cumulative estimate of a population-level effect (Borenstein et al., 2009; Glass, 1976). This approach was selected because it permits identification of both the direction and magnitude of STEAM-PA/STEM-PEH integration effects across diverse educational contexts.

Literature Search and Inclusion Criteria

A systematic literature search was conducted across Web of Science, Scopus, PubMed, SPORTDiscus, and ProQuest dissertations, covering English-language publications from 2005 to March 2026. The search was limited to studies published from 2005 onward because this period corresponds to the early development of research examining the integration of PE, health, and PA within STEM education. The Boolean search string combined three keyword clusters:

(1) integration framework: *"STEM" OR "STEAM" OR "STEAM-PA" OR "STEM-PEH" OR "integrated curriculum" OR "interdisciplinary teaching"*;

(2) content domain: *"physical education" OR "health education" OR "sport science" OR "kinesiology" OR "movement education"*; and

(3) outcomes: *"academic achievement" OR "motor skill" OR "health literacy" OR "physical activity" OR "learning outcome" OR "engagement."*

Literature searches across the five databases reached 423 records in total. Once the 38 duplicates were removed, 385 records went on to title and abstract screening, and 296 of these were dropped as clearly unrelated to the topic, leaving 89 studies to be tracked down in full text. At the full-text stage, 64 studies were excluded altogether: 20 didn't report enough statistical information to calculate an effect size, 18 didn't actually combine PA or PE with STEM/STEAM, 10 turned out to be qualitative, theoretical, or review papers rather than empirical studies, 5 were based on elite athlete samples, and 11 simply didn't have the data needed to compute Hedges' *g* (Hedges & Olkin, 1985). That left 25 studies, which were then checked against the inclusion criteria and assessed for eligibility and methodological quality, two more were dropped here because of overlapping samples or quality issues, bringing the final count to 23 independent studies. Since Duncan et al. (2024) reported two separate outcomes from the same sample (motor skills and well-being), it was coded as two distinct effect sizes rather than one. So the final dataset ends up being 24 effect sizes pulled from those 23 independent studies (see **Figure 2**).

Studies were included if they:

(a) employed an experimental or quasi-experimental design with a control or comparison condition, or used a single-group pre-/post-test design with sufficient statistical data;
(b) explicitly integrated PE, health, or sport with at least one STEM or STEAM discipline;
(c) reported sufficient quantitative data (means, standard deviations, F-statistics, eta-squared, or Cohen's *d*) to calculate Hedges' *g*; and
(d) targeted pre-school through undergraduate student populations.

Purely qualitative studies, theoretical papers, descriptive surveys, and studies conducted exclusively with elite athlete populations were excluded.

Effect Size Calculation

All effect sizes were converted to Hedges' *g*, which corrects for small-sample bias and is recommended for meta-analytic research. Effect-size conversions followed the procedures described by Borenstein et al. (2009). Specifically, studies reporting eta-squared were first converted to Cohen's *d* using the appropriate transformation and then adjusted to Hedges' *g* using the small-sample correction factor. Marzano (2003) was used for interpreting the magnitude of the resulting effect sizes. Where a single study reported multiple independent outcomes (e.g., motor skills and well-being), each outcome was treated as a separate effect size and coded independently (e.g., Duncan et al., 2024).

Statistical Analysis

Given the anticipated methodological and contextual heterogeneity of included studies, the DerSimonian and Laird (1986) random-effects model was adopted as the primary analytical framework. Heterogeneity was assessed using Cochran's Q statistic and the I^2 index, with values of 25%, 50%, and 75% indicating low, moderate, and high heterogeneity, respectively (Higgins et al., 2003). Between-study variance (τ^2) was estimated via the method-of-moments estimator. Publication bias was evaluated using the Egger weighted regression test (Egger et al., 1997) and visually inspected via a funnel plot. Moderator analyses were conducted through subgroup analyses using the mixed-effects model, with the Qb statistic used to assess between-group differences. All computations were performed in Python 3.12 using numpy and scipy libraries, replicating the analytical procedures of comprehensive meta-analysis version 4.

Coding and Moderator Variables

Coding of each suitable study included independent coding of the following variables:

(1) educational level (preschool, elementary, secondary, high school, and undergraduate education);
(2) cultural dimension, which was coded according to the four-category classification of Triandis and Gelfand (1998) as follows: vertical collectivist (VC): high power distance, group-oriented culture; examples include China, Indonesia, Malaysia, South Korea, Qatar), horizontal collectivist (HC): egalitarian group values; examples include Australia, Denmark, Germany), vertical individualist (VI): individualistic, competitive, hierarchical culture; examples include the USA, the UK), and horizontal individualist (HI): individualistic, egalitarian culture; examples include Spain) (because the primary studies did not report participants' individual cultural orientations, this classification was based on the country in which each study was conducted and was used as a proxy for the broader cultural context rather than as a measure of individual-level cultural values);
(3) STEM subject (mathematics, physics/engineering, science/biology, integrated STEM, health/science);
(4) outcome measure (academic achievement, STE(A)M attitude, motor skills, motivation/engagement, health literacy, PA, higher-order thinking skills, well-being, skills);
(5) teaching method (technology-enhanced, embodied cognition, STEM-based PE, problem-based learning, gamification, integrated curriculum, inquiry-based, active learning, classroom-based PA, integrated PA);
(6) integration approach (technology and sport analysis, embodied/kinesthetic, sports-based STEM, PA embedded STEM, health-based STEM, gamification); and
(7) intervention duration (≤ 2 weeks, 3-8 weeks, 9-18 weeks, > 18 weeks). Inter-rater reliability was assessed as a 30% random subsample using Cohen's kappa.

The descriptive characteristics of the studies included in the meta-analysis are presented in **Table 1** and **Table 2**.

Table 1. Characteristics of studies included in the meta-analytic review

Study	Country	Education level	n	Duration	STEM discipline	Teaching model	Integration type	Outcome measure
Abdelaziz et al. (2025)	Qatar	Middle school	107	5 days	Science/tech/eng	PBL	Sports-themed STEM	STEM interest survey
Bonner et al. (2024)	USA	Middle school	121	6 wks	Math/science/tech	Active learning	PA + STEM + health (THINK)	STEM attitude survey

Table 1 (Continued).

Study	Country	Education level	n	Duration	STEM discipline	Teaching model	Integration type	Outcome measure
Chen et al. (2023)	China	Middle school	40	6 wks	Physics/math	STEAM/video analysis	STEAM + sport analysis	Learning engagement scale
Duncan et al. (2024)	UK	Primary school	60	10 wks	Mathematics	Integrated curriculum	PA + academic (pedometer)	Motor skills (TGMD)
	UK	Primary school	60	10 wks	Mathematics	Integrated curriculum	PA + academic (pedometer)	Well-being (SMFQ)
Evans and Willis (2025)	Australia	High school	14	8 wks	Physics/math	F.I.T./tactical games	Fitness + biomechanics + STEM	SALG assessment
Have et al. (2018)	Denmark	Primary school	505	36 wks	Mathematics	Classroom-based PA	Active math lessons	Standardized math test
Jung et al. (2021)	USA	Primary school	167	10 wks	Computer science	Gamification (SMART)	PA + academic (SMART)	PA reporting accuracy
Kang et al. (2024)	USA	Primary school	48	6 wks	Physics	Embodied cognition	Active movements + physics	Physics knowledge test
Lee (2021)	South Korea	Middle school	178	14 wks	STEM (integrated)	STEAM-based PE	STEAM-based physical ed.	PE attitude
Lesiak et al. (2024)	USA	High school	2,122	1–2 wks	Biology	Inquiry-based	STEM + health (diabetes)	Knowledge gains
Mavilidi et al. (2016)	Australia	Preschool	90	1 wk	Science	Integrated PA	Embodied PA + STEM	Geography knowledge test
Mavilidi et al. (2017)	Australia	Preschool	90	4 wks	Science	Embodied cognition	Embodied PA + STEM	Science score (planetary)
Mullender-Wijnsma et al. (2016)	Netherlands	Primary school (grades 2-3)	499	22 weeks	Mathematics	Fit & Vaardig op school	PA + math academic ach.	Math achievement
Nazarudin et al. (2025)	Malaysia	Undergraduate	50	6 wks	Physics/math/eng	Mobile app/PBL	Mobile app + biomechanics + STEM	Biomechanical analysis skills
Noviana et al. (2025)	Indonesia	High school	101	12 wks	Physics	STEM E-module/PBL	STEM + sport (karate)	Creative thinking skills
Perry et al. (2025)	USA	Middle school	81	18 wks	STEM (integrated)	Active learning (THINK)	PA + nutrition + STEM	STEM attitude survey
Prieto Andreu (2024)	Spain	Undergraduate	9	~1 wk	Interdisciplinary	Active board game	Physical exercise + game	Motivation scale
Queiruga-Dios et al. (2025)	Spain	High school	71	18 wks	Physics/math	Collaborative PBL	STEM + sport (soccer)	Academic performance
Rahmaningrum et al. (2025)	Indonesia	Primary school	57	4 wks	Mathematics	Embodied numeracy play	Embodied play in PE	Motor skills (TGMD-2)
Sevil-Serrano et al. (2022)	Spain	Middle school	210	36 wks	Cross-curricular	SDT-based	School-based PA + PE	Motivation (SDT scales)
Strobl et al. (2020)	Germany	Middle school	233	36 wks	Health/science	Participatory planning	Health knowledge + PE	Health knowledge (HKU)
Wjayanti et al. (2025)	Indonesia	Undergraduate	84	8 wks	STEAM (integrated)	Differentiated STEAM	Differentiated STEAM + PA	Concept understanding
Yuan et al. (2022)	China	Primary school	280	Cross-sect.	STEAM	Physical exercise	STEAM + physical exercise	Positive emotions scale
Zhang et al. (2019)	China	High school	460	45 wks	Math/lang arts	Strengthened PE	Strengthened PE intervention	Standardized test scores

Note. wks: Weeks; PBL: Project-based learning; STEAM: STEM + arts; SDT: Self-determination theory; TGMD: Test of gross motor development; SALG: Student assessment of learning gains; HKU: Health knowledge and understanding; & Duncan et al. (2024) has independent effect sizes reflecting separate outcome domains for two rows

The review comprised 24 independent effect sizes published between 2016 and 2025, with the majority of studies (62.50%) published during the 2023-2025 period, reflecting the recent growth of research examining the integration of PA, sports, and movement-based approaches within STEM education. The studies were conducted across 11 countries, with the USA contributing the largest number of effect sizes, followed by Australia, China, Spain, and Indonesia. Based on the cultural framework of Triandis and Gelfand (1998), the sample represented all four cultural orientations, although studies conducted in VC contexts were most prevalent (**Table 1** and **Table 2**).

The interventions covered a broad educational spectrum, ranging from preschool to undergraduate education, with primary and middle school settings accounting for the largest proportion of the included studies. The instructional focus also reflected considerable diversity. While integrated STEM/STEAM approaches were the most frequently represented discipline, mathematics, physics/engineering, science, biology, and health-related STEM applications were also well represented. Likewise, a variety of pedagogical models including technology-enhanced learning, embodied cognition, inquiry-based instruction, project-based learning, active learning, gamification, and integrated PA approaches were employed, illustrating

Table 2. Descriptive profile of studies included in the meta-analytic review

Variable	Category	k	P (%)
Year of publication	2015-2019	4	16.67
	2020-2022	5	20.83
	2023-2025	15	62.50
Educational level	Preschool	2	8.33
	Primary school	7	29.17
	Middle school	7	29.17
	High school	5	20.83
	Undergraduate	3	12.50
STEM discipline	Mathematics	5	20.83
	Physics/engineering	6	25.00
	Science/biology	3	12.50
	Integrated STEM	8	33.33
	Health/science	2	8.33
Cultural dimension	VC	9	37.50
	HC	5	20.83
	VI	5	20.83
	HI	5	20.83
Teaching model	Technology-enhanced	3	12.50
	Embodied cognition	3	12.50
	STEAM-based PE	2	8.33
	PBL	2	8.33
	Gamification	2	8.33
	Integrated curriculum	2	8.33
	Inquiry-based	3	12.50
	Classroom-based PA	2	8.33
	Active learning	2	8.33
	Integrated PA	2	8.33
Integration type	Technology + sport analysis	3	12.50
	Embodied/kinesthetic	7	29.17
	Sports-context STEM	3	12.50
	PA-embedded STEM	7	29.17
	Health-integrated STEM	3	12.50
	Gamification-based STEM	1	4.17
Country	USA	5	20.83
	Australia	3	12.50
	China	3	12.50
	Spain	3	12.50
	Indonesia	3	12.50
	South Korea	1	4.17
	Malaysia	1	4.17
	Qatar	1	4.17
	UK	1	4.17
	Denmark	1	4.17
	Germany	1	4.17
Total		24	100

Note. k: Number of effect sizes; P: Percentage; Percentages may not sum to 100% due to rounding or multi-category coding; & cultural dimension classification follows Triandis and Gelfand (1998)

the absence of a single dominant instructional strategy across the literature.

Table 1 shows that intervention times ranged from approximately one week to 45 weeks, and sample sizes varied. Although all studies combined PA with STEM education, they did so in different ways. Such as including movement-based learning, sports-related activities, health-focused programs, and technology-supported instruction. The studies also measured different outcomes, covering both academic indicators (e.g., achievement and conceptual understanding) outcomes such as attitudes, engagement, motor skills, well-being, and health knowledge.

RESULTS

Overall Effect Size

Table 3 presents the overall meta-analytic findings. Based on the synthesis of k = 24 effect sizes from 23 independent studies (N = 5,238), the pooled effect size under the random-effects model was Hedges' g = 1.050 (standard error [SE] = 0.110, 95% confidence interval [CI] [0.835, 1.264], Z = 9.59, p < .001). According to Cohen's (1988) benchmarks, this constitutes a large effect, indicating that STEPAM-integrated instruction produces meaningfully superior student outcomes compared with control conditions.

Substantial heterogeneity was detected (Q [23] = 867.36, p < .001; I^2 = 97.3%; τ^2 = 0.252, τ = 0.502), confirming that approximately 97% of the total variance in effect sizes reflects genuine between-study differences rather than sampling error. This level of heterogeneity strongly supports the use of a random-effects model and necessitated the moderator analyses reported below.

Publication Bias

Publication bias was assessed using Egger's weighted regression test (Egger et al., 1997). The result was statistically non-significant (intercept = 5.15, p = .129), providing no evidence of systematic publication bias.

A funnel plot was constructed by plotting the individual effect sizes (Hedges' g) against their corresponding SEs. Under conditions of no publication bias, study estimates should cluster symmetrically around the pooled effect within the inverted funnel, with greater dispersion for smaller studies.

As seen in **Figure 3**, the distribution of the 24 effect sizes shows approximate symmetry around the pooled estimate of g = 1.050. The x-axis was restricted to the range [-0.60, 3.60] to enhance readability. Chen et al. (2023) produced an extreme effect size (g = 4.63) that falls beyond this range; this value is indicated with an annotated arrow at the right boundary of the plot. This

Table 3. Overall effect size and heterogeneity analysis results

Model	k	N	g	SE	95% CI lower	95% CI upper	Z	p	Q (I^2)
Random effects	24	5,238	1.050	0.110	0.835	1.264	9.59	< .001	867.36 (97.3%)

Note. g: Hedges' g; Z: z-statistic; Q: Cochran's Q; & I^2: Proportion of variance due to heterogeneity

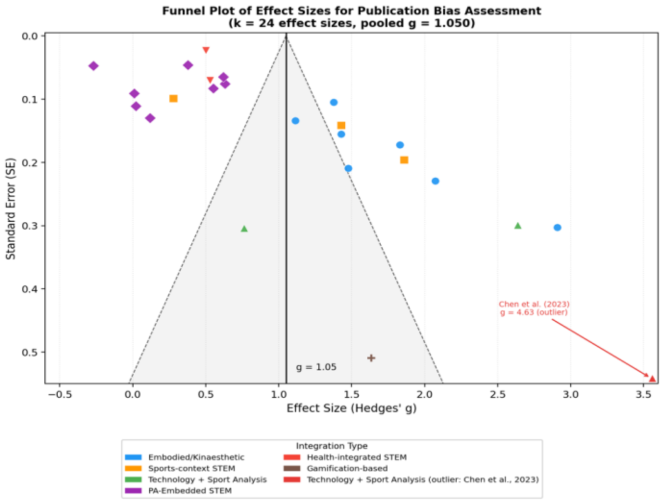

Figure 3. Funnel plot of effect sizes for publication bias assessment (the authors' own elaboration)

outlying estimate reflects the unique design of that study (video-analysis-based STEAM instruction applied to a single 40 student classroom over six weeks) and is not considered indicative of selective reporting. No systematic asymmetry suggesting missing small-sample studies was detected in the main body of the funnel.

Duval and Tweedie's (2000) trim-and-fill method was applied as a third independent check. As reported in **Table 4**, the procedure required no trimming (studies trimmed = 0), yielding an adjusted pooled effect identical to the observed estimate (g = 1.050, 95% CI [0.835, 1.264]). This indicates that the distribution of studies around the pooled estimate is sufficiently symmetric to require no correction. Taken together, the three publication bias indicators consistently suggest that the present synthesis is unlikely to be substantially distorted by selective reporting.

Moderator Analyses

To examine potential sources of variability, seven moderator analyses were conducted using the mixed-effects subgroup approach. The between-group heterogeneity statistic (Q*b*) was used to evaluate statistical significance of between-group differences. **Table 5** presents unified moderator analysis results.

Level of education proved to be a statistically significant moderator variable (Qb [4] = 798.69, p < .001). The most pronounced pooled effect was found for undergraduate-level studies (g = 1.89, 95% CI [1.22, 2.57]) and preschool research (g = 1.46). Other levels of education, such as primary school (g = 1.07), high school (g = 0.83), and middle school (g = 0.78) also exhibited a substantial impact on students' performance but were relatively less pronounced than at the undergraduate level. It can be explained by relatively small sample sizes in these studies (k = 3, N = 143).

Cultural dimension appeared to be a statistically significant moderator (Qb [3] = 799.10, p < .001). The largest effect sizes were found in the case of VC environments (g = 1.41, 95% CI [1.05, 1.76]). Then came HI (g = 1.19), HC (g = 0.91), and VI (g = 0.49). STEM field proved to be a statistically significant moderator (Qb [4] = 808.93, p < .001). Integration of physics/engineering had the largest effect size (g = 1.93, 95% CI [1.47, 2.39]), which is probably due to physical movements providing a direct kinesthetic analogy for the underlying biomechanical and mechanical phenomena (force, velocity, trajectory). After that came science/biology (g = 1.12), mathematics (g = 0.91), and finally, integrated STEM (g = 0.68). The health/science category produced the smallest effect (g = 0.58) with a CI crossing zero, suggesting greater variability and less consistent benefit in health-focused integration.

Outcome type was a statistically significant moderator (Qb [8] = 810.50, p < .001). Biomechanical skills (g = 2.64), well-being (g = 2.07), motivation/engagement (g = 1.43), higher-order thinking (g = 1.43), and motor skills (g = 1.34) all demonstrated large effects. Academic achievement (g = 0.98) remained substantial. STEM attitude (g = 0.42), PA (g = 0.55), and health literacy (g = 0.53) produced smaller effects whose CIs crossed zero, indicating greater outcome uncertainty in these domains. Teaching model was a statistically significant moderator (Qb [10] = 834.85, p < .001). Technology-enhanced models (mobile apps, e-modules, video analysis; g = 2.55) and embodied cognition approaches (g = 2.03) produced the highest effects, followed by STEAM-based PE (g = 1.40). PBL (g = 1.03), and gamification (g = 0.92) demonstrated moderate-to-large effects. Active learning approaches (THINK program) demonstrated essentially null effects (g = 0.02), driven largely by two null-finding studies (Bonner et al., 2024; Perry et al., 2025).

Integration type was a statistically significant moderator (Qb [5] = 822.93, p < .001). Technology + sport analysis integrations (e.g., biomechanical video analysis, mobile app-based sports analytics; g = 2.41) and embodied/kinesthetic approaches (g = 1.48) produced the strongest effects. Sports-context STEM (using sport as a conceptual framework for science or math instruction; g = 1.16) yielded a large effect. PA-

Table 4. Duval and Tweedie's (2000) trim-and-fill test results

Model	Studies trimmed	Point estimate (g)	95% CI lower	95% CI upper	Q
Observed values	-	1.050	0.835	1.264	867.36
Adjusted values	0	1.050	0.835	1.264	867.36

Note. g: Hedges' g weighted effect size; Q: Cochran's Q heterogeneity statistic; & adjusted values are identical to observed values because no studies were trimmed

Table 5. Unified moderator analysis results for STEPAM integration outcomes

Moderator	Category	k	N	*g*	95% CI lower	95% CI upper	Qb (p)
Educational level	Preschool	2	180	1.46	0.74	2.19	Qb (4) = 798.69, p < .001
	Primary school	7	1,177	1.07	0.68	1.46	
	Middle school	7	970	0.78	0.39	1.18	
	High school	5	2,768	0.83	0.36	1.29	
	Undergraduate	3	143	1.89	1.22	2.57	
Cultural dimension	VC	9	1,357	1.41	1.05	1.76	Qb (3) = 799.10, p < .001
	HC	5	932	0.91	0.45	1.37	
	HI	5	410	1.19	0.70	1.67	
	VI	5	2,539	0.49	0.04	0.94	
STEM discipline	Physics/engineering	6	324	1.93	1.47	2.39	Qb (4) = 808.93, p < .001
	Science/biology	3	2,302	1.12	0.54	1.71	
	Mathematics	5	1,142	0.91	0.44	1.37	
	Integrated STEM	8	1,027	0.68	0.31	1.04	
	Health/science	2	443	0.58	-0.12	1.28	
Outcome type	Skills (biomechanical)	1	50	2.64	1.49	3.78	Qb (8) = 810.50, p < .001
	Well-being	1	60	2.07	0.99	3.15	
	Motivation/engagement	4	539	1.43	0.86	2.01	
	Higher-order thinking	1	101	1.43	0.41	2.45	
	Motor skills	2	117	1.34	0.58	2.11	
	Academic achievement	9	3,484	0.98	0.64	1.33	
	Physical activity	1	167	0.55	-0.45	1.55	
	Health literacy	1	233	0.53	-0.46	1.52	
	STEM attitude	4	487	0.42	-0.08	0.92	
Teaching model	Technology-enhanced	3	191	2.55	1.87	3.22	Qb (10) = 834.85, p < .001
	Embodied cognition	3	195	2.03	1.40	2.65	
	STEAM-based PE	2	262	1.40	0.68	2.12	
	Integrated curriculum	2	120	1.04	0.29	1.78	
	PBL	2	178	1.03	0.30	1.76	
	Gamification	2	176	0.92	0.10	1.73	
	Tactical/games approach	1	14	0.76	-0.39	1.91	
	Inquiry-based	3	2,565	0.55	-0.02	1.13	
	Classroom-based PA	2	785	0.50	-0.20	1.20	
	Integrated PA	2	550	0.40	-0.31	1.11	
	Active learning	2	202	0.02	-0.70	0.73	
Integration type	Technology + sport analysis	3	104	2.41	1.70	3.11	Qb (5) = 822.93, p < .001
	Gamification-based	1	9	1.63	0.23	3.04	
	Embodied/kinesthetic	7	827	1.48	1.09	1.88	
	Sports-context STEM	3	279	1.16	0.57	1.76	
	PA-embedded STEM	7	1,583	0.47	0.08	0.85	
	Health-integrated STEM	3	2,436	0.35	-0.22	0.93	
Intervention duration	≤ 2 weeks (short)	4	2,328	0.77	0.24	1.31	Qb (3) = 811.00, p < .001
	3-8 weeks (medium)	8	504	1.75	1.35	2.14	
	9-18 weeks (long)	7	718	1.03	0.64	1.42	
	> 18 weeks (extended)	4	1,408	0.32	-0.18	0.81	
Publication year	2015-2019	4	1,145	0.73	0.23	1.24	Qb (2) = 794.50, p < .001
	2020-2022	5	1,068	0.74	0.29	1.19	
	2023-2025	15	3,025	1.27	0.99	1.55	

Note. k: Number of effect sizes; N:Total sample size; *g*: Hedges' *g*; Qb: Between-group heterogeneity statistic; cultural dimension classification follows Triandis and Gelfand's (1998) four-category typology: VC high power distance + group priority (China, Indonesia, Malaysia, South Korea, Qatar); HC egalitarian group norms (Australia, Denmark, Germany); VI competitive, hierarchical, autonomous (USA, UK); HI autonomous, egalitarian (Spain); & all Qb statistics are statistically significant (p < .001)

embedded STEM (PA added to STEM lessons without conceptual alignment; $g = 0.47$) and health-integrated STEM ($g = 0.35$) showed smaller, less consistent effects, with the latter's CI crossing zero. This pattern underscores a critical design distinction: integrations where PA is conceptually tethered to STEM content (embodied or technology-mediated) consistently outperform those where PA serves only as a generalized attention or motivation boost. Intervention duration was a statistically significant moderator (Qb [3] = 811.00, $p <$ .001). Medium-duration interventions (3-8 weeks) yielded the highest pooled effect ($g = 1.75$, 95% CI [1.35, 2.14]), followed by long interventions (9-18 weeks; $g = 1.03$). Short interventions (≤ 2 weeks; $g = 0.77$) and extended interventions (> 18 weeks; $g = 0.32$) produced smaller effects. The diminishing returns observed at extended durations may reflect novelty decay, participant fatigue, or methodological attenuation effects in longitudinal designs.

DISCUSSION

This meta-analysis examined the effect of integrating STEM/STEAM education with PE, health, and PA on student outcomes. The random-effects model, based on 24 effect sizes and a combined sample of 5,238 participants, produced a large, pooled effect. The integration of PEH into STEM education generates considerable improvements in academic performance, motor skills, motivation, and high-level thinking when the integration process is coherent, technology-based, or physically engaging, and occurs for a sufficient period. The fact that the resulting effect size sits above what classic PA-cognitive outputs meta-analyses typically report (usually small to moderate) (e.g., Fedewa & Ahn, 2011; He et al., 2025; Norris et al., 2020) suggests that STEM-PA integration is not simply a "movement break" model, but a pedagogical design that creates a direct conceptual link between the intervention and the targeted outcome. According to meta-analysis by Fedewa and Ahn (2011) encompassing 59 studies revealed that PA significantly and positively impacted children's academic and cognitive performance, although this effect was medium. Similarly, a more recent meta-analysis of school-based PA studies by He et al. (2025) found that PA had a small but significant effect on overall academic and a similarly small effect on math achievement. Given that in most of these studies PA was linked to academic content only in a temporal sense (for example, movement breaks inserted between lessons) rather than conceptually, the large effect observed here can reasonably be attributed to the STEPAM framework's core principle of treating the body as a vehicle that directly carries STEM concepts. This indicated that the need to reposition PE not merely as a domain of motor development, but as a functional learning context for building STEM literacy.

The high level of heterogeneity and the fact that all seven moderator variables reached statistical significance show that a blanket statement such as "STEM-PA integration works" falls short, and that the size of the effect shifts depending on how the intervention is designed. The strongest effects came from technology-supported teaching models and approaches grounded in embodied cognition, which indicates that the theoretical grounding of the instructional design plays a decisive role in outcomes. Embodied cognition theory holds that cognitive processes are shaped not only by the brain but also by the body's interaction with its environment, and the fact that interventions built on this theoretical basis consistently produced stronger effects offers empirical support for that claim. Likewise, the integration type combining technology and sport analysis produced the highest effect, which suggests that digital tools (wearable technologies, biomechanical measurement systems) act as a bridge that converts concrete movement into cognitive STEM performance.

Looking at the findings related to intervention duration, medium-length interventions (3-8 weeks) produced the strongest effect, while the effect declined somewhat for longer interventions (> 18 weeks; $g = 0.32$). This pattern may also be partly explained by methodological factors, such as the larger sample sizes typically found in longer studies (for example, $N = 460$ in Zhang et al., 2019) and the gradual "contamination" of control conditions over time (for instance, control-group teachers picking up similar practices). Turning to the cultural dimension moderator, VC cultures showed a significantly higher effect than VI cultures. This gap can be interpreted through Hofstede's (2001) cultural dimensions framework: in educational settings built around group harmony and acceptance of authority, structured, teacher-led STEM-PA programs may be implemented more consistently, which in turn raises intervention fidelity. In individualistic and competitive educational cultures, by contrast, students' concern with individual performance may have limited the potential benefit of collaborative, integrated learning activities.

The findings of research support earlier meta-analysis examining the link between PA and cognitive performance. Donnelly et al. (2016) review on PA, fitness, and academic achievement in children reported that structured PA interventions consistently produced positive effects on academic outcomes, though the size of that effect depended on how well the activity was integrated with academic content. The fact that the present meta-analysis found a weaker effect in the PA-embedded STEM category (where PA was added without conceptual linkage; $g = 0.47$) but a stronger one in the embodied/kinesthetic category (where movement was directly tied to conceptual content; $g = 1.48$) supports Donnelly et al.'s (2016) emphasis on the quality of integration as the deciding factor. Becker and Park (2011) indicated moderate positive effects from interdisciplinary STEM approaches, though that work left PA and PE outside its scope. The large effect size obtained here suggests that bringing PE into STEM integration models can yield additional gains beyond what purely academic interdisciplinary integration achieves on its own, a result that fits with the idea that kinesthetic learning has the potential to make abstract concepts concrete (Glenberg, 2010).

The advantage shown by embodied cognition approaches also fits well with Wilson's (2002) theoretical work outlining six views of embodied cognition. Wilson (2002) argues that cognition is situated and action-based, meaning that learning tasks blended with physical experience produce deeper conceptual understanding than abstract or symbolic instruction alone. The strongest effect found here in the physics/engineering field can be read through the same lens: because concepts like force, velocity, and acceleration can be experienced directly through bodily movement, it is unsurprising that kinesthetic integration would strongly support conceptual learning in these disciplines.

The moderator analysis by publication year showed that studies published between 2023 and 2025 seemed larger effects than those from earlier. The reason for this can be with more recent interventions, probably aided by greater use of technology, becoming better grounded theoretically and more effective in design. This is also consistent with the teacher-training-centered, technology-supported integration model proposed by the STEPAM framework (Da Cruz et al., 2025), suggesting that it lines up with where the field appears to be heading.

The results of this meta-analysis should be read with several limitations in mind. The number of studies included was quite low in some moderator categories, which limits the reliability of those subgroup estimates; effect sizes drawn from a small handful of studies should be interpreted with caution. Although the high heterogeneity was addressed through seven moderator variables, the possibility that unmeasured sources, such as differences in sample size, measurement tools, and intervention fidelity across studies, also contributed cannot be ruled out. In addition, a substantial portion of the included studies relied on small samples (for example, N = 9 in Prieto Andreu, 2024; N = 14 in Evans & Willis, 2025), which increases the share of sampling error in the effect size estimates. The cultural dimension classification, based on Triandis and Gelfand's (1998) four-category typology, involves country-level generalizations and may overlook regional or institutional differences within the same country. Finally, although publication bias indicators (Egger's weighted regression test [Egger et al., 1997], the funnel plot, the trim-and-fill method) showed no sign of systematic bias, the search strategy was limited to English-language publications and a restricted set of databases, which may have left out relevant studies published in other languages.

CONCLUSION AND RECOMMENDATIONS

This meta-analysis demonstrates, on the basis of the data, that integrating STEM/STEAM education with PE, health, and PA produces a large positive effect on student outcomes overall. The findings show that this effect is not homogeneous and varies meaningfully depending on the theoretical grounding of the integration, the technology involved, the duration of the intervention, the educational level, and the cultural context. Important point is the finding that approaches grounded in embodied cognition, supported by technology, and built around a direct conceptual link between physical movement and STEM content produce the greatest gains. Based on these findings, several recommendations can be offered to stakeholders.

Curriculum developers and school administrators should redesign PE not as a field separate from the STEM curriculum, but as a functional learning environment where conceptual learning can be made concrete. When designing interventions, PA should not be treated merely as an attention-boosting break but should be structured to create a direct conceptual link with the targeted STEM concept (for example, biomechanical analysis, kinematics-based sport applications). Teacher training programs should incorporate modules, in line with the STEPAM framework, that build teachers' ability to translate STEM concepts into embodied learning activities. It is also worth bearing in mind that designing interventions within a 3-8 week window may help sustain the effect at its peak by preventing the decay of the novelty effect.

Areas represented by relatively few studies, such as health-STEM integration and gamification-based integration types, require further experimental research. Cross-cultural comparative studies can be designed to cultural factors not just at the country level, but at the institutional and classroom level as well. Long-term can help studies clarify how durable the observed gains are and whether any decay effects emerge over time. Broader analyses that include studies published in other languages and indexed in additional databases would likely improve the generalizability of findings in this area.

Author contributions: MU & AM: conceptualization, formal analysis, methodology, writing – original draft, writing – review & editing. Both authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: This article is a literature review and does not contain any original empirical studies conducted by the authors with human participants or animals. The synthesis and presentation of the existing research were conducted in accordance with standard academic integrity and publishing ethics.

AI statement: The authors declare that AI-assisted technologies were only used to check and improve the grammar and language of select sentences in this manuscript. The authors wrote the entire manuscript themselves and took full responsibility for the content of the publication.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

*The studies were included in the meta-analysis.

*Abdelaziz, M., Muneer, S., Ali, R., Jlassi, K., Al-Thani, N., & Bhadra, J. (2025). Igniting STEM passion: Building STEM interest and 21st-century skills through science in sports for preparatory male students in Qatar. *School Science and Mathematics*. https://doi.org/10.1111/ssm.18390

Ali, R., Bhadra, J., Siby, N., Ahmad, Z., & Al-Thani, N. J. (2021). A STEM model to engage students in sustainable science education through sports: A case study in Qatar. *Sustainability, 13*(6), Article 3483. https://doi.org/10.3390/su13063483

Almarcha, M., Vázquez, P., Hristovski, R., & Balagué, N. (2023). Transdisciplinary embodied education in elementary school: A real integrative approach for the science, technology, engineering, arts, and mathematics teaching. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1134823

Barbosa, A., Whiting, S., Simmonds, P., Scotini Moreno, R., Mendes, R., & Breda, J. (2020). Physical activity and academic achievement: An umbrella review. *International Journal of Environmental Research and Public Health, 17*(16), Article 5972. https://doi.org/10.3390/ijerph17165972

Becker, K., & Park, K. (2011). Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. *Journal of STEM Education, 12*(5-6), 23-37.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641-1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

*Bonner, J., Xiong, W., Velasquez, C., Nienhuis, N., Wallace, B., Friedman, A., Lee, D., & Perry, A. (2024). A novel summer camp integrating physical, psychological, and educational health in youth: The THINK program. *Nutrients, 16*(12), Article 1838. https://doi.org/10.3390/nu16121838

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386

Carson, V., Kuzik, N., Hunter, S., Wiebe, S. A., Spence, J. C., Friedman, A., Tremblay, M. S., Slater, L. G., & Hinkley, T. (2015). Systematic review of sedentary behavior and cognitive development in early childhood. *Preventive Medicine, 78*, 115-122. https://doi.org/10.1016/j.ypmed.2015.07.016

Castelli, D. M., Centeio, E. E., Hwang, J., Barcelona, J. M., Glowacki, E. M., Calvert, H. G., & Nicksic, H. M. (2014). VII. The history of physical activity and academic performance research: Informing the future. *Monographs of the Society for Research in Child Development, 79*(4), 119-148. https://doi.org/10.1111/mono.12133

*Chen, H., Rao, F., Chen, R., & Lin, Z. (2023). Teaching STEAM in the shaolin staff program: Ways to stimulate student engagement in learning. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1264985

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Da Cruz, M., Lau, J., Mihajlović, A., Fenyvesi, K., Lavicza, Z., & Ignjatović, A. (2025). Comparative perspectives on integrating physical activity and STEAM in Austria, Serbia, and Finland: A policy framework for interdisciplinary learning and the STEPAM initiative. In G. Luptakova, B. Antala, G. Blint, M. Lazar, & D. Novak (Eds.), *Bridging theory and practice in physical education. Researches-curriculum-teacher training* (pp. 223-240). Publishing House Bacău.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177-188. https://doi.org/10.1016/0197-2456(86)90046-2

Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience, 18*, 34-48. https://doi.org/10.1016/j.dcn.2015.11.005

Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine, 52*(Suppl. 1), S36-S42. https://doi.org/10.1016/j.ypmed.2011.01.021

Donnelly, J. E., Hillman, C. H., Castelli, D., Etnier, J. L., Lee, S., Tomporowski, P., Lambourne, K., & Szabo-Reed, A. N. (2016). Physical activity, fitness, cognitive function, and academic achievement in children: A systematic review. *Medicine & Science in Sports & Exercise, 48*(6), 1197-1222. https://doi.org/10.1249/MSS.0000000000000901

*Duncan, M. J., Fitton Davies, K., Okwose, N., Harwood, A. E., & Jakovljevic, D. G. (2024). Effects of a 10-week integrated curriculum intervention on physical activity, resting blood pressure, motor skills, and well-being in 6- to 7-year-olds. *Journal of Physical Activity and Health, 21*(6), 595-605. https://doi.org/10.1123/jpah.2023-0239

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89-98. https://doi.org/10.2307/2669529

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

*Evans, S., & Willis, C. (2025). Fitness integrated with technology approach to teaching biomechanics and STEM in a high school setting: A case report. *Frontiers in Sports and Active Living, 7*. https://doi.org/10.3389/fspor.2025.1681868

Fedewa, A. L., & Ahn, S. (2011). The effects of physical activity and physical fitness on children's achievement and cognitive outcomes: A meta-analysis. *Research Quarterly for Exercise and Sport, 82*(3), 521-535. https://doi.org/10.1080/02701367.2011.10599785

Fullerton, S. A., & Chapman, I. (2026). Integrating STEM education within strength & conditioning-based physical education classes. *Strength & Conditioning Journal, 48*(2), 191-200. https://doi.org/10.1519/SSC.0000000000000947

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3-8. https://doi.org/10.3102/0013189X005010003

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *WIREs Cognitive Science, 1*(4), 586-596. https://doi.org/10.1002/wcs.55

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: A pooled analysis of 298 population-based surveys with 1·6 million participants. *The Lancet Child & Adolescent Health, 4*(1), 23-35. https://doi.org/10.1016/S2352-4642(19)30323-2

*Have, M., Nielsen, J. H., Ernst, M. T., Gejl, A. K., Fredens, K., Grøntved, A., & Kristensen, P. L. (2018). Classroom-based physical activity improves children's math achievement – A randomized controlled trial. *PLoS ONE, 13*(12), Article e0208787. https://doi.org/10.1371/journal.pone.0208787

He, H., Yang, Y., Sun, J., Wang, F., Zhang, W., & Zhu, F. (2025). Effects of school-based physical activity on academic achievement in children and adolescents: A systematic review and meta-analysis. *Frontiers in Public Health, 13*. https://doi.org/10.3389/fpubh.2025.1651883

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press. https://doi.org/10.1016/C2009-0-03396-0

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560. https://doi.org/10.1136/bmj.327.7414.557

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. *Nature Reviews Neuroscience, 9*(1), 58-65. https://doi.org/10.1038/nrn2298

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). SAGE.

*Jung, Y., Burson, S. L., Julien, C., Bray, D. F., & Castelli, D. M. (2021). Development of a school-based physical activity intervention using an integrated approach: Project SMART. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.648625

Jylänki, P., Mbay, T., Hakkarainen, A., Sääkslahti, A., & Aunio, P. (2022). The effects of motor skill and physical activity interventions on preschoolers' cognitive and academic skills: A systematic review. *Preventive Medicine, 155*, Article 106948. https://doi.org/10.1016/j.ypmed.2021.106948

*Kang, S., Lu, M., Black, J. B., & Kim, S. (2024). Mindful movements matter: Differentiating active body movements in underprivileged students' learning of physics concepts. *Research in Science & Technological Education, 42*(2), 412-430. https://doi.org/10.1080/02635143.2022.2093344

*Lee, D.-J. (2021). The effect of STEAM-based physical Education classes on middle school students' attitudes toward physical education classes and self-directed learning abilities. *Iranian Journal of Public Health. 50*(5), 938-948. https://doi.org/10.18502/ijph.v50i5.6111

*Lesiak, A., Griswold, J. C., Moylan, A., & Starks, H. (2024). Development and evaluation of integrated diabetes curricula for teaching gene by environment concepts to high school health and biology students. *Journal of STEM Outreach, 7*(1), 1-22. https://doi.org/10.15695/jstem/v7i1.04

Li, C., Kam, W. K. K., & Zhang, M. (2019). Physical education teachers' behaviors and intentions of integrating STEM education in teaching. *Physical Educator, 76*(4), 1086-1101. https://doi.org/10.18666/TPE-2019-V76-I4-9104

Luo, X., Herold, F., Ludyga, S., Gerber, M., Kamijo, K., Pontifex, M. B., Hillman, C. H., Alderman, B. L., Müller, N. G., Kramer, A. F., Ishihara, T., Song, W., & Zou, L. (2023). Association of physical activity and fitness with executive function among preschoolers. *International Journal of Clinical and Health Psychology, 23*(4), Article 100400. https://doi.org/10.1016/j.ijchp.2023.100400

Marzano, R. J. (2003). *What works in schools: Translating research into action*. ASCD.

*Mavilidi, M., Okely, A. D., Chandler, P., & Paas, F. (2016). Infusing physical activities into the classroom: Effects on preschool children's geography learning. *Mind, Brain, and Education, 10*(4), 256-263. https://doi.org/10.1111/mbe.12131

*Mavilidi, M., Okely, A. D., Chandler, P., & Paas, F. (2017). Effects of integrating physical activities into a science lesson on preschool children's learning and enjoyment. *Applied Cognitive Psychology, 31*(3), 281-290. https://doi.org/10.1002/acp.3325

*Mullender-Wijnsma, M. J., Hartman, E., De Greeff, J. W., Doolaard, S., Bosker, R. J., & Visscher, C. (2016). Physically active math and language lessons improve academic achievement: A cluster randomized controlled trial. *Pediatrics, 137*(3), Article e20152743. https://doi.org/10.1542/peds.2015-2743

*Nazarudin, M. N., Pa, W. A. M. W., Zainal, M. S., Musa, R. M., Ismail, S. I., & Mu, Y. (2025). The transformative role of mobile applications in biomechanics education: Enhancing learning, engagement, and STEM integration. In *Proceedings of the 2025 10th International STEM Education Conference* (pp. 1-6). https://doi.org/10.1109/iSTEM-Ed65612.2025.11129333

Niederer, I., Kriemler, S., Gut, J., Hartmann, T., Schindler, C., Barral, J., & Puder, J. J. (2011). Relationship of aerobic fitness and motor skills with memory and attention in preschoolers (Ballabeina): A cross-sectional and longitudinal study. *BMC Pediatrics, 11*, Article 34. https://doi.org/10.1186/1471-2431-11-34

Norris, E., van Steen, T., Direito, A., & Stamatakis, E. (2020). Physically active lessons in schools and their impact on physical activity, educational, health and cognition outcomes: A systematic review and meta-analysis. *British Journal of Sports Medicine, 54*(14), 826-838. https://doi.org/10.1136/bjsports-2018-100502

*Noviana, R. S., Sriyanti, I., Bayu, W. I., & Rahman, N. F. B. A. (2025). STEM integration in karate sport as physics learning media: Development, validation, and analysis of student learning outcomes on impulse and momentum materials. *Jurnal Penelitian Pendidikan IPA, 11*(9), 72-82. https://doi.org/10.29303/jppipa.v11i9.11972

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

*Perry, A., Bonner, J., Williams, S., Xiong, W., Garcia, A., Velasquez, C., Friedman, A., Lee, D. L., Hernandes, I. D. L., Shen, J., Meyer, M., & Fernandez, L. (2025). Integrating physical activity into a nutrition and exercise science middle school curriculum: The THINK program. *Nutrients, 17*(9), Article 1538. https://doi.org/10.3390/nu17091538

*Prieto Andreu, J. M. (2024). Paleocontacto GT: Fusionando ejercicio físico con un juego de mesa (Paleocontacto GT: Merging physical exercise with a board game]. *Retos, 59*, 892-902. https://doi.org/10.47197/retos.v59.107734

*Queiruga-Dios, M. Á., Vázquez Dorrío, J. B., Sáiz-Manzanares, M. C., López-Iñesta, E., & Diez-Ojeda, M. (2025). STEM approach using soccer: Improving academic performance in physics and mathematics in a real-world context. *Frontiers in Psychology, 16*. https://doi.org/10.3389/fpsyg.2025.1503397

*Rahmaningrum, T. J., Kristiyandaru, A., & Indahwati, N. (2025). Embodied numeracy play in physical education: Effects on fundamental motor skills and critical thinking in primary school children. *AL-ISHLAH: Jurnal Pendidikan, 17*(4). https://doi.org/10.35445/alishlah.v17i4.8963

*Sevil-Serrano, J., Aibar, A., Abós, Á., Generelo, E., & García-González, L. (2022). Improving motivation for physical activity and physical education through a school-based intervention. *The Journal of Experimental Education, 90*(2), 383-403. https://doi.org/10.1080/00220973.2020.1764466

*Strobl, H., Ptack, K., Töpfer, C., Sygusch, R., & Tittlbach, S. (2020). Effects of a participatory school-based intervention on students' health-related knowledge and understanding. *Frontiers in Public Health, 8*. https://doi.org/10.3389/fpubh.2020.00122

Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology, 74*(1), 118-128. https://doi.org/10.1037/0022-3514.74.1.118

*Wijayanti, A., Wiyanto, W., Ridlo, S., & Parmin, P. (2025). El efecto de la actividad física en proyectos STEAM diferenciados en el aprendizaje de ciencias y el desarrollo de habilidades del siglo XXI [The effect of physical activity in differentiated STEAM projects on science learning and the development of 21st-century skills]. *Retos, 70*, 1079-1096. https://doi.org/10.47197/retos.v70.116789

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625-636. https://doi.org/10.3758/BF03196322

*Yuan, Y., Ji, X., Yang, X., Wang, C., Samsudin, S., & Omar Dev, R. D. (2022). The effect of persistence of physical exercise on the positive psychological emotions of primary school students under the STEAM education concept. *International Journal of Environmental Research and Public Health, 19*(18), 11451. https://doi.org/10.3390/ijerph191811451

Zhan, Z., Shen, W., Xu, Z., Niu, S., & You, G. (2022). A bibliometric analysis of the global landscape on STEM education (2004-2021): Towards global distribution, subject integration, and research trends. *Asia Pacific Journal of Innovation and Entrepreneurship, 16*(2), 171-203. https://doi.org/10.1108/APJIE-08-2022-0090

*Zhang, Y., Ma, X., Zhao, J., Shen, H., & Jiang, F. (2019). The effect of strengthened physical education on academic achievements in high school students: A quasi-experiment in China. *International Journal of Environmental Research and Public Health, 16*(23), 4688. https://doi.org/10.3390/ijerph16234688

Zollman, A. (2012). Learning for STEM literacy: STEM literacy for learning. *School Science and Mathematics, 112*(1), 12-19. https://doi.org/10.1111/j.1949-8594.2012.00101.x