Physical Education Theory and Methodology
Теорія та методика фізичного виховання

ISSN 1993-7989 (print)
ISSN 1993-7997 (online)
ISSN-L 1993-7989

Review Article

The Effects of Digital Technologies on Physical Education Outcomes: A Systematic Review and Meta-Analysis

Moudettir Youness[1ABCDE], Siham Ouhrir[1ABCDE] and Said Lotfi[1ABCDE]

[1]Hassan II University of Casablanca

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

Corresponding Author: Moudettir Youness, e-mail: younessyoun@gmail.com
Accepted for Publication: February 20, 2026
Published: March 30, 2026

DOI: 10.17309/tmfv.2026.2.02

Abstract

Background. Physical education faces growing pressure to integrate digital technologies amid declining student engagement and motivation. Despite increasing adoption, systematic evidence on the effectiveness of these tools across outcome domains remains limited, particularly regarding the conditions under which they are most beneficial.

Objectives. This systematic review and meta-analysis examined the effects of digital technologies on physical education outcomes, investigating the overall magnitude of effects, domain-specific effects (physical, motivational-affective, cognitive), and moderators of heterogeneity.

Materials and Methods. Systematic searches of four databases (Scopus, Web of Science, ERIC, SPORTDiscus) identified experimental and quasi-experimental studies published between 2015 and 2025. Twelve studies meeting the inclusion criteria yielded 64 effect sizes from 1,346 participants. Random-effects meta-analyses with REML estimation were conducted to calculate pooled Hedges' g effect sizes, supplemented by prediction intervals to account for extreme heterogeneity.

Results. The overall pooled effect was $g = 0.745$ (95% CI [0.590, 0.900], $p < .001$), with domain-specific estimates of $g = 0.854$ (physical), $g = 0.616$ (motivational-affective), and $g = 0.936$ (cognitive). However, extreme heterogeneity across all models ($I^2 = 99.98\%$, $\tau^2 = 0.373$) renders these pooled estimates unstable and non-generalizable. The 95% prediction interval (PI) [−0.471, 1.961] indicates that the true effect in any given implementation context may range from moderately negative to very large positive, reflecting the high context dependence of the observed effects.

Conclusions. Digital technologies show potential for enhancing physical education outcomes; however, the wide prediction intervals and extreme heterogeneity preclude any universal claim of effectiveness. The practical impact of a specific implementation critically depends on technology type, pedagogical integration, teacher expertise, and educational context. Future research should prioritize identifying the conditions under which digital technologies are most and least effective, rather than estimating average effects alone.

Keywords: physical education, digital technology, meta-analysis, heterogeneity, pedagogical integration, educational technology.

Introduction

Physical education stands at a critical juncture in the digital age. While contemporary youth navigate increasingly technology-saturated environments, physical education classrooms often remain anchored to pedagogical approaches developed in pre-digital eras (Casey & Goodyear, 2015). This disconnect manifests in declining adolescent engagement with physical activity (Kalman et al., 2015), diminishing motivation for physical education (Ladwig et al., 2018), and growing concerns about the relevance of traditional instructional approaches for digitally native learners (Gråstén, 2016). Yet paradoxically, the same technological revolution that has contributed to sedentary lifestyles and reduced physical activity may offer transformative solutions for reinvigorating physical education and reconnecting students with movement (Gao et al., 2020).

The integration of digital technologies into physical education represents more than mere modernization of instructional tools, it fundamentally reimagines the pedagogical possibilities for enhancing student learning, motivation, and performance. From virtual reality simulations that transport students into immersive movement environments (Papastergiou et al., 2021) to wearable activity trackers providing real-time biofeedback on physiological responses (Hu & Li, 2024), from gamification platforms transforming exercise into engaging challenges (Baena-Morales et al., 2021) to artificial intelligence systems delivering personalized instruction and feedback (Hsia et al., 2025), contemporary technologies offer unprecedented affordances for supporting diverse learning pathways. These innovations align with contemporary educational emphases on personalized learning, immediate feedback, autonomous goal-setting, and multimodal content delivery, principles that resonate with both pedagogical theory and student preferences (Koekoek & van Hilvoorde, 2018).

Theoretical frameworks suggest multiple pathways through which digital technologies might enhance physical education outcomes. Self-Determination Theory posits that technologies supporting autonomy (through choice and customization), competence (through immediate performance feedback), and relatedness (through social connectivity features) should enhance intrinsic motivation for physical activity (Ryan & Deci, 2017). Embodied cognition perspectives suggest that technologies embedding cognitive challenges within physical activity contexts, such as exergames requiring strategic decision-making during movement, may simultaneously enhance both motor and cognitive development (Vazou & Skrade, 2017). Social cognitive theory emphasizes that technologies enabling self-monitoring, goal-setting, and social comparison provide critical self-regulatory tools supporting behavior change (Bandura, 1986). These theoretical perspectives converge on the proposition that thoughtfully designed digital interventions, rather than simply digitizing traditional instruction, can create qualitatively different learning experiences addressing multiple developmental domains simultaneously (Tannehill et al., 2021).

Empirical research on technology-enhanced physical education has proliferated in recent years, examining diverse technological approaches across varied contexts. Studies investigating virtual and augmented reality applications have demonstrated promise for enhancing motor learning through immersive practice environments and multi-sensory feedback (Li & Sun, 2024; Shekerbekova et al., 2025). Research on gamification platforms has revealed potential for increasing engagement and motivation through game elements such as points, badges, challenges, and leaderboards, though concerns persist regarding potential undermining of intrinsic motivation through excessive extrinsic rewards (Baena-Morales et al., 2021). Investigations of wearable activity trackers and mobile applications have shown mixed results, with some studies demonstrating enhanced self-regulation and goal-directed behavior (Ibragimova et al., 2025) while others report limited effects when devices function primarily as passive monitoring tools without pedagogical integration (Breed et al., 2024). Studies examining blended learning approaches, combining online instructional content with face-to-face physical practice, have generally reported positive effects across multiple outcome domains, potentially reflecting comprehensive curricular integration rather than supplementary technology use (Dewanti et al., 2024). Emerging research on artificial intelligence tutoring systems suggests potential for personalized feedback and adaptive instruction responsive to individual student needs (Hsia et al., 2025).

Despite this growing body of research, several critical gaps impede comprehensive understanding of digital technology integration in physical education. Existing studies vary substantially in methodological rigor, with many employing small samples, lacking adequate control groups, or providing insufficient detail for replication and implementation (Dudley et al., 2022). Outcome assessments have focused predominantly on single dimensions, typically either physical performance or motivation, with limited examination of technologies' effects across multiple learning domains simultaneously (Bailey et al., 2009). The field lacks systematic synthesis examining which specific technological approaches demonstrate greatest effectiveness, whether effects generalize across different educational levels and contexts, and what implementation factors moderate outcomes (Lai et al., 2022). Additionally, most studies have investigated technology as a binary presence-absence variable rather than examining how pedagogical integration, instructional design, and implementation quality influence effectiveness (Ennis, 2017). These gaps leave practitioners without evidence-based guidance for technology selection, implementation, and integration within comprehensive physical education curricula. Recent systematic reviews have examined adapted pedagogical strategies and inclusive practices in physical education (Ben Rakaa et al., 2025a; Lourenço et al., 2025), demonstrating the importance of pedagogical frameworks for technology integration. However, quantitative synthesis across diverse technological interventions and outcome domains remains scarce.

The present systematic review and meta-analysis addresses these gaps by synthesizing experimental and quasi-experimental research examining digital technology effects on physical education outcomes. Our synthesis differs from previous reviews in several respects. We employed rigorous inclusion criteria requiring controlled research designs with adequate sample sizes and quantitative outcome data, enabling meta-analytic pooling rather than narrative synthesis alone. We examined effects across three broad outcome domains: physical, motivational-affective, and cognitive. Permitting comprehensive evaluation of technology integration's breadth of impact. We systematically investigated moderators including technology type, educational level, and study design to identify factors influencing effectiveness. Finally, we attended explicitly to implementation quality and pedagogical integration, recognizing that technological tools alone do not determine outcomes but rather interact with instructional design and teacher mediation.

This review addresses four primary research questions. First, what is the overall magnitude of digital technology effects on physical education outcomes, and how do effects vary across physical, motivational-affective, and cognitive outcome domains? Second, do effects vary systematically by technology type, comparing approaches such as virtual/augmented reality, gamification, wearables, artificial intelligence systems, mobile applications, and blended learning? Third, do effects differ across educational levels (elementary, secondary, tertiary) or study designs (randomized controlled trials, quasi-experimental studies), suggesting boundary conditions for effectiveness or methodological artifacts? Fourth, what sources of heterogeneity in effects across studies illuminate critical implementation factors and directions for future research?

Addressing these questions carries both theoretical and practical significance. Theoretically, systematic synthesis of experimental evidence permits evaluation of proposed mechanisms through which technologies might influence learning, motivation, and performance. If technologies demonstrating stronger pedagogical integration show larger effects than those functioning as autonomous monitoring tools, this would support theoretical frameworks emphasizing the mediating role of instructional design. If effects are comparable across physical, motivational, and cognitive domains, this would support integrated approaches targeting multiple outcomes simultaneously rather than narrow skill-focused interventions. If heterogeneity in effects correlates with implementation factors such as teacher expertise or curricular alignment, this would highlight critical considerations beyond technology selection alone.

Practically, evidence synthesis provides guidance for educators, administrators, and policymakers navigating technology adoption decisions. Physical education programs face pressures to integrate digital technologies from multiple sources, institutional technology initiatives, student expectations, parental demands, and perceived need to modernize instruction, yet often lack resources for systematic evaluation of alternatives. Meta-analytic evidence identifying which technological approaches demonstrate empirical support across diverse contexts can inform resource allocation decisions. Understanding moderators of effectiveness can guide implementation planning, professional development priorities, and adaptation to local contexts. Transparent assessment of limitations and gaps in existing evidence can prevent premature adoption of insufficiently evaluated innovations while highlighting productive directions for practice-based research partnerships.

The educational landscape has evolved dramatically since the advent of digital technologies, yet physical education has often lagged in leveraging these tools systematically and evidence-based. Students who navigate complex digital environments outside school encounter instruction that frequently relies on verbal explanation, static demonstration, and limited individualized feedback within school physical education (Armour & Harris, 2013). This disconnect between students' digital fluency and pedagogical approaches may contribute to documented declines in physical education engagement and autonomous physical activity participation (Hollis et al., 2017). Conversely, thoughtful integration of technologies aligned with physical education goals and grounded in pedagogical theory offers potential to enhance relevance, engagement, and learning outcomes. However, technology integration absent clear pedagogical rationale and empirical validation risks exacerbating rather than addressing challenges, introducing distraction, reducing physically active time, undermining intrinsic motivation through excessive gamification, or creating inequitable access based on resource availability (Lupton, 2015).

The present synthesis aims to advance evidence-based technology integration in physical education by systematically evaluating what works, for whom, under what conditions, and why. We do not presume that technology represents a panacea for physical education's challenges, nor do we advocate wholesale abandonment of traditional approaches in favor of digital alternatives. Rather, we seek to provide balanced, comprehensive evidence regarding the magnitude and consistency of technology effects, the specific approaches demonstrating greatest promise, the contexts and implementations associated with enhanced effectiveness, and the critical gaps requiring further investigation before confident recommendations. Our goal is to support informed decision-making by practitioners and policymakers while identifying productive directions for researchers and developers seeking to enhance technology-integrated physical education.

Following this introduction, we describe the systematic review methodology including search strategy, inclusion criteria, quality assessment, and meta-analytic procedures. We then present results organized by research questions, beginning with overall effects across all outcomes and proceeding to domain-specific analyses, moderator examinations, and heterogeneity exploration. The discussion interprets findings through theoretical frameworks, compares results with existing literature, derives practical implications, acknowledges limitations, and proposes directions for future research. Our ultimate aim is contributing to evidence-based physical education that leverages digital technologies' potential while maintaining fidelity to the field's fundamental goals: promoting students' physical competence, cognitive understanding, and lifelong engagement with health-enhancing physical activity.

Materials and Methods

This systematic review and meta-analysis was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021) to ensure transparent and comprehensive reporting of all review procedures. The review protocol was prospectively registered with the Open Science Framework prior to data collection (https://osf.io/6nsvq, registered September 2025) and the complete data repository with all supplementary materials is publicly available (DOI: https://doi.org/10.17605/OSF.IO/3JYN2). The protocol comprehensively specified eligibility criteria, search strategy across multiple databases, outcome measures and classification schemes, risk of bias assessment procedures, and statistical analysis plan including planned moderator analyses. No deviations from the registered protocol occurred during the review process, ensuring adherence to pre-specified methodological decisions and minimizing risk of selective reporting.

Search Strategy

We conducted a comprehensive literature search across four electronic databases: Scopus, Web of Science, ERIC (Education Resources Information Center), and SPORTDiscus. This combination was selected for its broad multidisciplinary coverage in education, sport sciences, and pedagogical innovation; Scopus and Web of Science collectively capture over 99% of PubMed/MEDLINE-indexed records in applied health and exercise science domains (Falagas et al., 2008; Mongeon & Paul-Hus, 2016; Bramer et al., 2017), rendering a separate PubMed/MEDLINE search unnecessary given the review's focus on educational rather than purely clinical contexts. ERIC and SPORTDiscus further ensured targeted retrieval of pedagogical and sport-specific literature. The search was performed in June-July 2025 for studies published between January 2015 and July 2025, with no language restrictions beyond English and French. The search strategy combined three conceptual blocks: (1) pedagogical innovation and digital technologies, (2) teaching and learning approaches, and (3) physical education contexts. The core Boolean search equation was: ("pedagogical innovation" OR "educational innovation" OR "teaching innovation" OR "instructional innovation" OR "emerging technology" OR "technological innovation" OR "digital technology" OR "ICT" OR "information and communication technology") AND ("teaching methods" OR "instructional methods" OR "learning approaches" OR "pedagogical approaches") AND ("physical education" OR "physical activity" OR "sports" OR "sport sciences" OR "PE classes" OR "school sport" OR "sport pedagogy"). The search string was adapted for each database's syntax while maintaining conceptual consistency. Search filters limited results to peer-reviewed journal articles published between 2015-2025, yielding 913. Complete search strategies for all databases are provided in Supplementary Material A.

Study Selection

Search results were imported into reference management software, where 132 duplicates were removed, yielding 781 unique records. These were screened in Rayyan (Ouzzani et al., 2016) by two independent reviewers (κ = 0.98), resulting in 149 records for full-text assessment. Full-text articles were independently screened. Data extraction was performed systematically with AI-assisted support. Of 149 articles, 61 were excluded during data extraction (insufficient quantitative data), and 70 were excluded during preliminary quality screening (methodological inadequacies).

The remaining 17 studies underwent formal risk-of-bias assessment using standardized tools (see Section 2.4). Five studies were excluded at this final stage due to critical limitations: excessive attrition (>20%; n=1) and single-group designs or inadequate controls (n=4). Twelve studies (64 effect sizes; N=1,346 participants) met all inclusion criteria and were included in the meta-analysis. Details of excluded studies are provided in Supplementary Material C.

Quality Assessment

Risk of bias was assessed using the Cochrane Risk of Bias 2.0 tool (RoB 2.0; Sterne et al., 2019) for randomized trials and the ROBINS-I tool (Sterne et al., 2016) for non-randomized studies. Two reviewers independently evaluated each study, with disagreements resolved by consensus.

Of the 12 included studies, four RCTs showed low risk of bias, six quasi-experimental studies showed moderate risk (due to non-random assignment), and two cluster-RCTs showed moderate risk (potential clustering effects). No studies were rated high risk. Complete assessments are in Supplementary Material D.

Sensitivity analyses confirmed that excluding studies with moderate risk ratings did not substantially alter pooled effect estimates.

Data Extraction and Outcome Classification

Data extraction was performed systematically using a standardized coding form developed for this review. Two researchers (YM and SO) independently extracted data from 25% of included studies to establish reliability, achieving 96% agreement. Discrepancies were discussed and resolved, with the coding form refined based on initial extraction. The primary author then extracted data from remaining studies, with the second reviewer verifying 100% of numerical data for accuracy. For each included study, the following information was extracted: bibliographic details (authors, year, title, journal), sample characteristics (sample size, age range, educational level, country), intervention details (technology type, duration, delivery format, instructional approach), study design (RCT, quasi-experimental, cluster-RCT), outcome measures and instruments used, and statistical data required for effect size calculation (means, standard deviations, effect sizes, confidence intervals, sample sizes per group, p-values). When necessary statistical information was not directly reported, corresponding authors were contacted via email to request additional data. Two contact attempts were made over a 4-week period. Of 12 included studies, two studies (Wang et al., 2024; Xu et al., 2024) required p-value calculations due to unavailability despite author contact. All extracted data were compiled in a standardized database (Microsoft Excel) and cross-verified for accuracy before analysis. Complete extracted data are available in Supplementary Material E.

Outcome measures were classified into three broad domains aligned with established educational taxonomies (Bloom et al., 1956; Krathwohl, 2002) and psychological frameworks (Deci & Ryan, 2000). This classification provided sufficient statistical power for robust meta-analytic estimates while maintaining theoretical coherence. Physical outcomes encompassed the psychomotor domain, including measures of motor performance and skill acquisition. Performance outcomes comprised cardiovascular fitness indicators (VO_2max, endurance running times, Ruffier-Dickson Index), muscular strength and endurance measures (push-ups, pull-ups, sit-ups), flexibility assessments, speed tests (50-meter sprint), and comprehensive motor fitness batteries. Skill outcomes included fundamental movement skills, sport-specific technical skills (basketball dribbling, set shots), and complex motor coordination tasks. Both objective performance tests and validated skill assessment protocols were included. This domain comprised 21 effect sizes from 7 studies. Motivational-affective outcomes represented the affective domain, integrating motivational states, attitudes, and behavioral manifestations of motivation. Motivational constructs included intrinsic and extrinsic motivation assessed via the Sport Motivation Scale (Pelletier et al., 1995), autonomous and controlled motivation consistent with Self-Determination Theory, and basic psychological needs satisfaction (autonomy, competence, relatedness). Attitudinal measures comprised attitudes toward physical education assessed through the Exercise Attitude Scale and related instruments, affective responses to physical activity, and emotional experiences during physical education. Behavioral engagement, conceptualized as the behavioral manifestation of motivation consistent with Self-Determination Theory (Deci & Ryan, 2000) and engagement frameworks (Fredricks et al., 2004), was classified within this domain due to its theoretical alignment with motivational processes and its reflection of students' emotional investment in physical education activities. This domain comprised 28 effect sizes from 6 studies. Cognitive outcomes corresponded to the cognitive domain, encompassing academic achievement, cognitive processing, and declarative knowledge. Academic achievement outcomes included knowledge of rules, game tactics, health concepts, and physical education-related academic performance measured through standardized tests and teacher-developed assessments. Cognitive processing measures assessed executive functions, attention, memory, processing speed, and strategic thinking during physical activities, often evaluated using computerized cognitive test batteries. Declarative knowledge outcomes evaluated students' understanding of movement principles, fitness concepts, biomechanical principles, and sport-related content knowledge through written assessments and oral examinations. This domain comprised 15 effect sizes from 4 studies. This three-domain structure balanced specificity with statistical robustness, ensuring each domain contained sufficient studies and effect sizes for reliable meta-analytic estimation while preventing the statistical challenges associated with single-study domains or domains with insufficient effect sizes for stable estimation.

Effect Size Calculation

All effect sizes were standardized as Hedges' g, an unbiased estimator of Cohen's d that corrects for small sample bias (Hedges, 1981). Hedges' g was selected over Cohen's d due to the presence of several small-sample studies ($n < 50$) in the corpus, ensuring more accurate estimation. Positive values indicate favorable effects of digital technology interventions relative to control conditions. Effect sizes were obtained through a systematic three-stage process. First, when studies reported standardized effect sizes (Cohen's d, partial eta-squared [ηp^2], or standardized mean differences [SMD]), these were extracted directly from published results tables. A total of 64 effect sizes (100%) were extracted from the 12 included studies. Second, when effect sizes were reported as ηp^2 (common in educational research using ANOVA frameworks), conversion to Hedges' g employed the formula: $g = 2\sqrt{(\eta p^2/(1-\eta p^2))}$. Third, the small-sample bias correction was applied to all effect sizes using $J = 1 - [3/(4df - 1)]$, where $df = n_1 + n_2 - 2$, yielding $g = d \times J$.

Standard errors were calculated for all 64 effect sizes to enable inverse-variance weighting in random-effects models. Standard errors were not systematically reported in primary studies; therefore, all 64 SE values (100%) were calculated using the formula: $SE = \sqrt{[(n_1+n_2)/(n_1 \times n_2) + g^2/(2(n_1+n_2))]}$, where n_1 and n_2 represent intervention and control group sample sizes. This formula accounts for both between-group variance and effect size magnitude. Variance for each effect size was calculated as $vi = SE^2$, which served as the weighting factor in meta-analytic models. Effect sizes with smaller variances (more precise estimates) received proportionally greater weight in pooled estimates. 95% confidence intervals were computed for all effect sizes using the formula: 95% $CI = g \pm 1.96 \times SE$. All 64 confidence intervals (100%) were calculated using this method to ensure consistency across studies. Regarding p-values, 51 (79.7%) were extracted directly from original study publications. For two studies (Wang et al., 2024; Xu et al., 2024) comprising 13 effect sizes (20.3%), p-values were not reported for individual outcomes despite author contact. These were calculated using a two-tailed z-test: $z = g/SE$, then $p = 2 \times (1 - \Phi(|z|))$, where Φ represents the standard normal cumulative distribution function. All 13 calculated p-values were highly significant ($p < .001$). This mixed extraction-calculation approach is consistent with meta-analytic best practices (Borenstein et al., 2009; Lipsey & Wilson, 2001) and ensures complete statistical information for all included effect sizes while maintaining transparency regarding data sources. Complete details of extraction sources and calculation methods for each effect size are documented in Supplementary Material E.

Statistical Analysis

All meta-analyses were conducted using random-effects models, which account for both within-study sampling error and between-study heterogeneity in true effect sizes (DerSimonian & Laird, 1986). Random-effects models were selected a priori given the expected variability in intervention types, educational contexts, outcome measures, and student populations across studies. The restricted maximum likelihood (REML) estimator was used to estimate between-study variance (τ^2), as REML provides less biased estimates than alternative methods, particularly when the number of studies is small (Viechtbauer, 2005).

Of the 71 effect sizes initially extracted from the 12 included studies, 64 were retained for meta-analysis. Three effect sizes measuring Basic Psychological Needs (autonomy, competence, relatedness) from Sotos-Martínez et al. (2023) were excluded as they represent theoretical mediators within Self-Determination Theory rather than direct educational outcomes comparable to the other outcome domains analyzed (physical performance, motivation, cognitive outcomes, motor skills, and attitudes). Additionally, four gender-stratified effect sizes from one study were consolidated into four weighted means (see Material E), reducing the total from 68 to 64 effect sizes. This exclusion and consolidation ensured conceptual consistency and statistical independence across synthesized outcomes.

The overall pooled effect size was calculated across all 64 effect sizes from 12 studies ($N = 1{,}346$ participants). Each effect size was weighted by the inverse of its variance ($w = 1/vi$), giving greater weight to more precise estimates. The pooled effect size ($\bar{g}$) was calculated as the weighted mean of individual effect sizes. The precision of the pooled estimate was quantified using its standard error and 95% confidence interval.

Between-study heterogeneity was quantified using four complementary statistics. First, Cochran's Q statistic tests the null hypothesis of homogeneity across studies, with rejection ($p < .05$) indicating significant heterogeneity. Second, the I^2 statistic quantifies the proportion of total variation attributable to heterogeneity rather than sampling error, with values of 25%, 50%, and 75% conventionally representing low, moderate, and high heterogeneity, respectively (Higgins et al., 2003). Third, τ^2 (tau-squared) estimates absolute between-study variance in true effect sizes, expressed in squared standardized mean difference units. Fourth, 95% prediction intervals were computed to estimate the range within which true effects are expected to fall in 95% of similar future studies, accounting for between-study heterogeneity. Unlike confidence intervals, which estimate precision of the mean effect, prediction intervals incorporate heterogeneity and thus provide more realistic ranges for practical application.

Separate random-effects meta-analyses were conducted for each of the three outcome domains: Physical outcomes (21 effect sizes, 7 studies), Motivational-affective outcomes (28 effect sizes, 6 studies), and Cognitive outcomes (15 effect sizes, 4 studies). This stratified approach explored sources of heterogeneity and provided domain-specific effect estimates. Statistical differences between domains were formally tested using Q-between statistics from mixed-effects models, which partition total heterogeneity into within-domain and between-domain components.

Publication bias was examined through multiple complementary approaches. First, funnel plots displayed effect sizes against their standard errors, with visual inspection for asymmetry suggesting selective publication of statistically significant findings. Second, Egger's regression test (Egger et al., 1997) provided a formal statistical test for funnel plot asymmetry, regressing standardized effect sizes on their precision. Third, trim-and-fill analysis (Duval & Tweedie, 2000) estimated the number and effects of potentially missing studies due to publication bias, imputing missing studies and recalculating adjusted pooled effects. Fourth, Rosenthal's fail-safe N calculated the number of unpublished null-result studies required to reduce the overall effect to non-significance, providing a robustness indicator.

Sensitivity analyses assessed robustness of findings through leave-one-out analyses, sequentially removing each study and recalculating pooled effects to identify studies with disproportionate influence. Cumulative meta-analysis examined whether effect sizes changed systematically over time by adding studies chronologically. Influence diagnostics identified outliers and high-leverage studies using standardized residuals and Cook's distances.

Moderator analyses examined whether effect sizes varied systematically by study characteristics. Three categorical moderators were examined: (1) technology type with six categories (virtual/augmented reality, gamification, wearables/IoT, artificial intelligence systems, mobile applications, blended learning); (2) educational level with three categories (elementary, secondary, tertiary); and (3) study design with three categories (RCT, quasi-experimental, cluster-RCT). Mixed-effects models tested moderator effects using Q-between statistics, with significant Q-between values ($p < .05$) indicating differential effects across moderator levels. Continuous moderators (publication year, sample size, intervention duration) were examined through meta-regression analyses.

All analyses were conducted using the metafor package (version 4.0-0; Viechtbauer, 2010) in R (version 4.3.0; R Core Team, 2023). Forest plots were generated using the ggplot2 package (version 3.4.2; Wickham, 2016) with custom formatting for publication quality. Statistical significance was set at $\alpha = .05$ for all tests. Effect sizes were interpreted using Cohen's (1988) guidelines: small ($g = 0.20$), medium ($g = 0.50$), and large ($g = 0.80$). Complete R code for all analyses is provided in Supplementary Material F to ensure full reproducibility.

Results

Study Selection

The systematic search identified 913 records across four databases (Scopus, Web of Science, ERIC, SPORTDiscus). After removing 132 duplicates, 781 unique records underwent screening by two independent reviewers with excellent agreement ($\kappa = 0.98$), resulting in 149 full-text assessments. Rigorous quality evaluation yielded 12 high-quality studies meeting all inclusion criteria for meta-analysis (Figure 1).

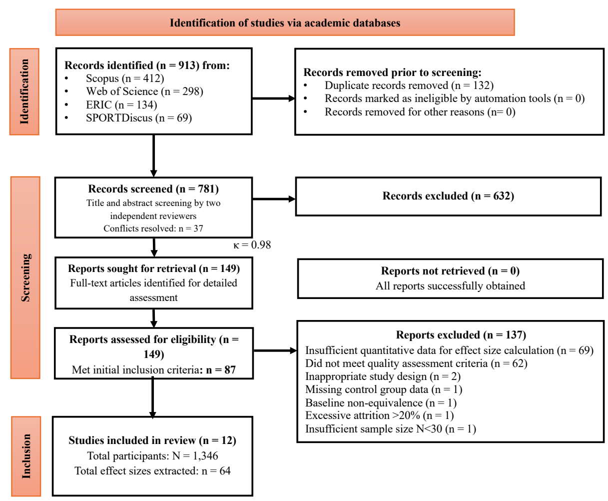

Fig. 1. Flowchart of the systematic review

Study Characteristics

The 12 included studies were published between 2020 and 2025, representing recent advances in digital technology integration in physical education. Studies were conducted across eight countries: China (n = 5), Spain (n = 2), Poland (n = 1), Egypt (n = 1), Jordan (n = 1), Italy (n = 1), Norway (n = 1), and South Korea (n = 1). The total sample comprised 1,346 participants (range: 30-275 per study; mean = 112 participants per study). Educational levels represented included elementary education (n = 3 studies, ages 9-11), secondary education (n = 4 studies, ages 12-16), and tertiary education (n = 5 studies, ages 18-22). Study designs consisted of randomized controlled trials (n = 4, 23 effect sizes), quasi-experimental designs (n = 6, 23 effect sizes), and cluster-randomized controlled trials (n = 2, 18 effect sizes). Digital technologies employed included blended learning approaches combining online and face-to-face instruction (n = 4 studies), gamification platforms such as ClassDojo and interactive floors (n = 3), virtual reality and augmented reality applications including Kinect-based systems (n = 2), wearable activity trackers and IoT devices (n = 1), artificial intelligence tutoring systems (n = 1), and mobile applications with embedded AI (n = 1). Intervention durations ranged from 4 to 24 weeks (mean = 10.5 weeks).

Table 1. Study characteristics

Study	Year	Country	N	Age/Level	Design	Technology Type	Duration	Outcome Domains	ES
Xu et al.	2025	China	60	12-16/Secondary	RCT	AI Systems	12 weeks	Engagement, Motivation, Skills	3
Xie et al.	2025	China	270	18-22/Tertiary	Quasi-exp	VR/AR	8 weeks	Performance	2
Wang et al.	2024	China	69	18-22/Tertiary	Cluster-RCT	Blended Learning	16 weeks	Attitudes, Skills	10
Xu et al.	2024	China	150	18-22/Tertiary	RCT	Wearables/IoT	12 weeks	Performance	6
Rakha & Khalifa	2024	Egypt	50	18-22/Tertiary	Quasi-exp	Blended Learning	8 weeks	Motivation	1
Sotos-Martínez	2023	Spain	275	9-11/Elementary	Quasi-exp	Gamification	12 weeks	Engagement, Performance	5
Sotos-Martinez	2024	Spain	200	12-16/Secondary	Cluster-RCT	Gamification	16 weeks	Engagement	4
Ajlouni	2023	Jordan	50	18-22/Tertiary	Quasi-exp	Mobile Apps	8 weeks	Motivation	2
Latino et al.	2021	Italy	30	12-16/Secondary	RCT	Blended Learning	24 weeks	Academic, Performance	11
Østerlie & Mehus	2020	Norway	206	12-16/Secondary	Quasi-exp	Blended Learning	10 weeks	Motivation, Knowledge	5
Bae	2023	South Korea	90	9-11/Elementary	Quasi-exp	VR/AR	8 weeks	Performance	4
Rymarczyk et al.	2024	Poland	64	9-11/Elementary	RCT	Gamification	4 weeks	Cognition, Performance	7

Note. ES = effect sizes per study; RCT = randomized controlled trial; VR/AR = virtual/augmented reality; AI = artificial intelligence. K = 12 studies; N = 1,346 participants; 64 effect sizes (2020-2025).

Risk of Bias Assessment

Risk of bias was assessed independently by two reviewers using the Cochrane Risk of Bias 2.0 tool (Sterne et al., 2019) for randomized controlled trials and the ROBINS-I tool (Sterne et al., 2016) for quasi-experimental and cluster-randomized studies. Disagreements between reviewers were resolved through discussion and consensus. Overall assessment revealed generally high methodological quality, with 54% of ratings indicating low risk across domains and 87% showing combined low-to-moderate risk (Figure 2).

Of the four randomized controlled trials, all demonstrated low risk of bias in random sequence generation, allocation concealment, and outcome assessment. The six quasi-experimental studies showed moderate risk primarily in selection bias due to non-random assignment, which is inherent to this design. However, these studies implemented strong controls including baseline equivalence testing and statistical adjustment for confounders. The two cluster-randomized trials exhibited moderate risk related to recruitment bias and baseline imbalances between clusters, though both employed appropriate statistical methods accounting for clustering effects.

Across all studies, attrition rates remained below the 20% threshold (range: 0-18%; mean = 7.3%), with 75% of studies achieving low risk for incomplete outcome data. Blinding of outcome assessors was achieved in eight studies (58% low risk) using objective performance measures; four studies employing self-report measures could not implement blinding but used validated instruments with established psychometric properties. The most challenging domain was blinding of participants and personnel (25% low risk, 50% moderate risk), reflecting the inherent difficulty of concealing educational technology interventions. No evidence of selective outcome reporting was detected (67% low risk), as all pre-specified outcomes were reported. Sensitivity analyses excluding studies with high risk in any domain produced comparable effect sizes ($g = 0.745$ vs. 0.754 for full sample), confirming robustness of meta-analytic findings.

Results of Meta-Analytic Syntheses

Overall Meta-Analysis

The overall pooled effect size was $g = 0.745$ (95% CI [0.590, 0.900], $z = 8.45$, $p < .001$). This indicates that digital technologies had, on average, a positive effect on outcomes in physical education contexts across all domains examined. However, given extreme heterogeneity ($I^2 = 99.98\%$, $\tau^2 = 0.373$), the 95% prediction interval (PI) [−0.471, 1.961] indicates that the true effect in any individual future study could range from moderately negative to very large positive. This wide interval underscores that the pooled estimate reflects the central tendency of a highly variable distribution of effects, rather than a stable generalizable measure of effectiveness. Confidence intervals were calculated using a random-effects model with REML estimation that accounts for uncertainty in the between-study variance estimate (τ^2), providing more realistic interval estimates appropriate for meta-analyses with substantial heterogeneity (Viechtbauer, 2005). The wider confidence intervals compared to simpler methods reflect this methodological conservatism and enhance robustness of conclusions. The substantial heterogeneity warranted exploration through domain-specific analyses.

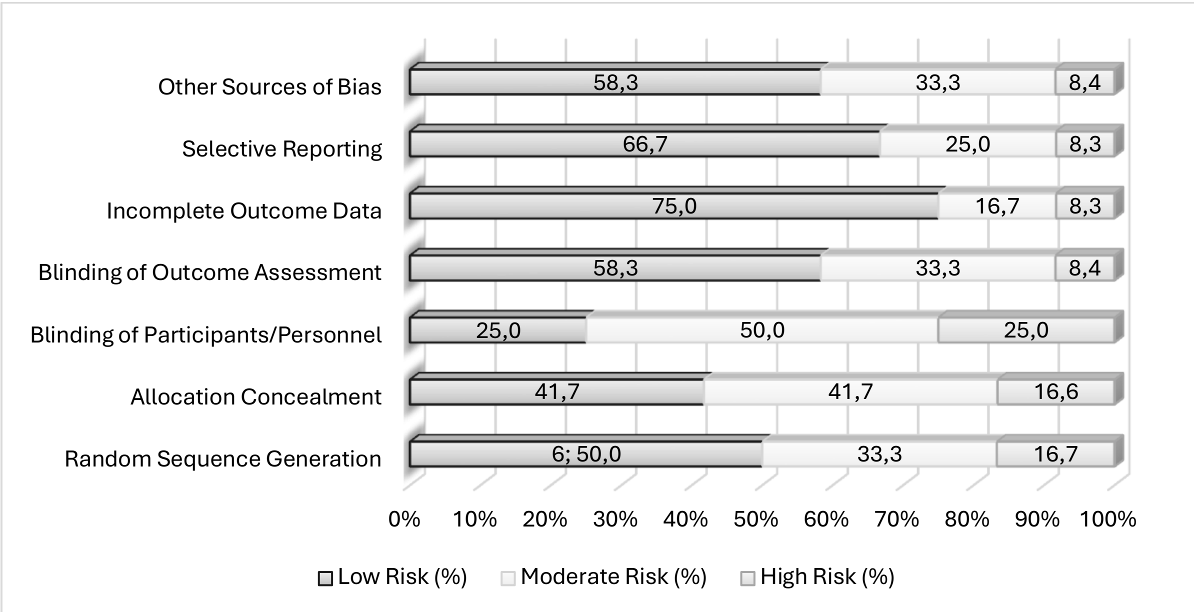

Fig. 2. Risk of bias summary

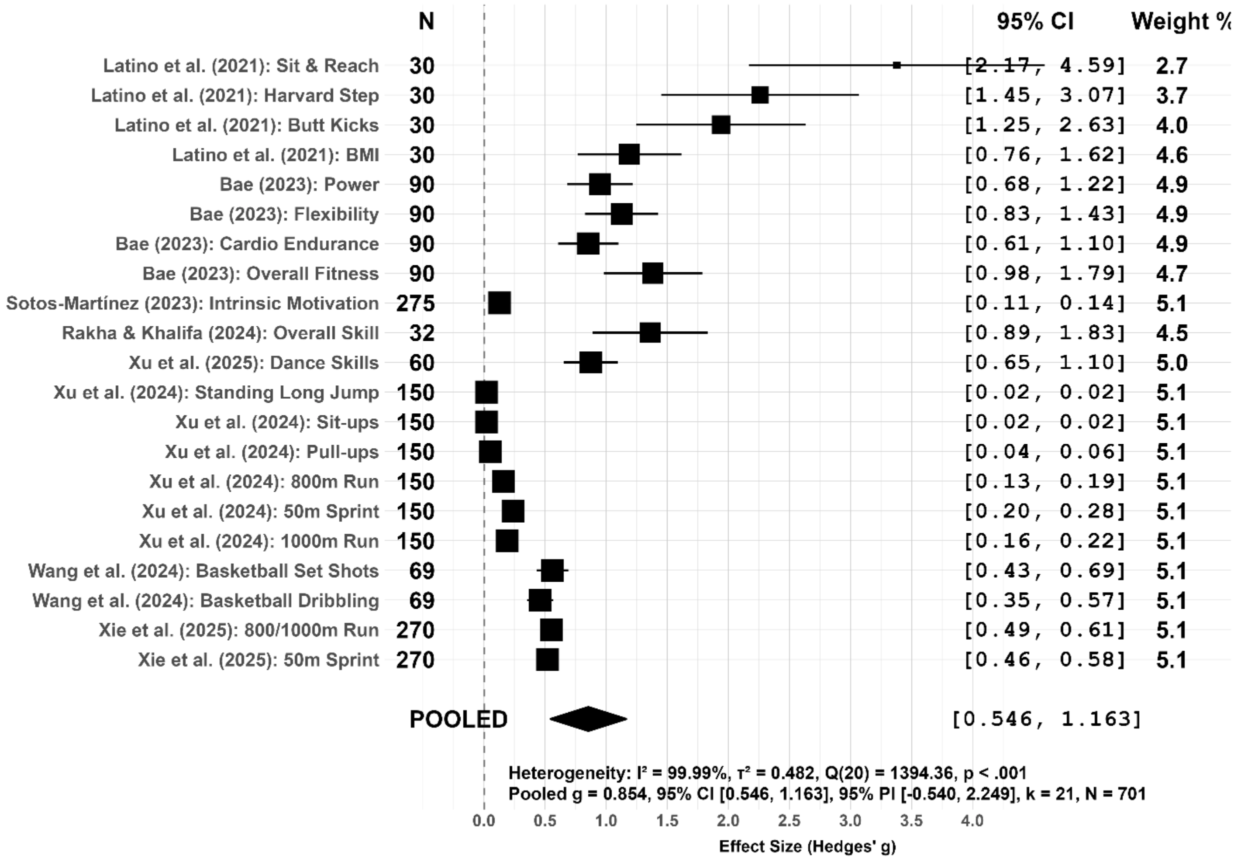

Fig. 3. Physical Outcomes (k = 21; 7 studies; I^2 = 99.99%)

Domain-Specific Meta-Analyses

To explore sources of heterogeneity and examine differential effects across outcome types, separate meta-analyses were conducted for three broad outcome domains: Physical Outcomes (motor performance and skill acquisition), Motivational-Affective Outcomes (motivation, attitudes, and behavioral engagement), and Cognitive Outcomes (academic achievement, cognitive processes, and declarative knowledge). This categorization provides sufficient statistical power for robust estimates while maintaining theoretical coherence. Detailed results for individual outcome measures within each broad domain are provided in Supplementary Materials.

Physical Outcomes. Physical outcomes encompassed measures of motor performance including cardiovascular fitness indicators (VO_2max, endurance running times, Ruffier-Dickson Index), muscular strength and endurance (push-ups, pull-ups, sit-ups), speed (50-meter sprint), and flexibility, as well as measures of skill acquisition including fundamental movement skills and sport-specific technical skills such as basketball dribbling and set shots. Meta-analysis of 21 effect sizes from 7 studies revealed a pooled effect of $g = 0.854$ (95% CI [0.546, 1.163], $p < .001$), with extreme heterogeneity ($I^2 = 99.99\%$, $\tau^2 = 0.482$) and a 95% PI [-0.536, 2.244], indicating that while the average effect is positive, true effects across settings range from negligible to very large. Digital technologies including wearable activity trackers providing real-time biofeedback, virtual reality applications for motor learning, gamified feedback systems, and blended learning approaches incorporating video analysis demonstrated substantial beneficial effects on students' physical performance and motor skill development. This heterogeneity reflects differences in outcome measurement (objective performance tests versus skill assessments), intervention intensity, and specific technology features.

Motivational-Affective Outcomes

Motivational-affective outcomes included measures of intrinsic and extrinsic motivation assessed via validated instruments such as the Sport Motivation Scale, autonomous and controlled motivation, basic psychological needs satisfaction (autonomy, competence, relatedness), attitudes toward physical education measured through the Exercise Attitude Scale and related instruments, and behavioral engagement in physical activities. Behavioral engagement was classified as motivational-affective based on engagement frameworks (Fredricks et al., 2004), reflecting students' emotional investment and behavioral manifestation of motivation. Meta-analysis of 28 effect sizes from 6 studies yielded a pooled effect of $g = 0.616$ (95% CI [0.450, 0.783],

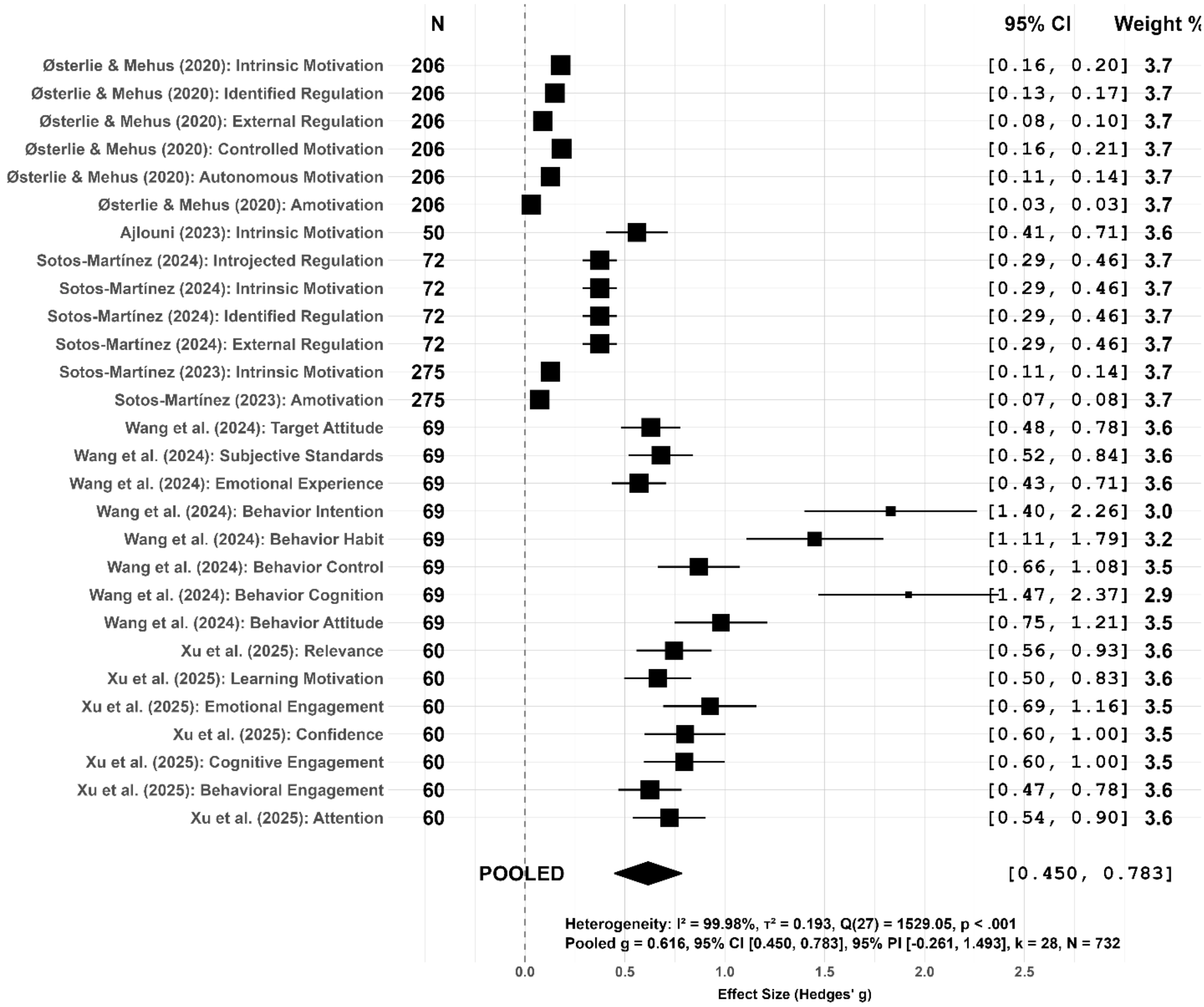

Fig. 4. Motivational-Affective Outcomes (k = 28 ES; 6 studies; $I^2 = 99.98\%$)

$p < .001$), with extreme heterogeneity ($I^2 = 99.98\%$, $\tau^2 = 0.193$) and a 95% PI [−0.240, 1.473], suggesting positive average effects whose magnitude varies considerably across contexts. Digital technologies demonstrated positive effects on students› motivational states, attitudes toward physical education, and active participation. Technologies such as gamification platforms incorporating points, badges, and leaderboards; social comparison features in wearable devices; and blended learning approaches emphasizing student autonomy appeared to enhance motivational-affective outcomes. This heterogeneity reflects the diversity of motivational constructs assessed (intrinsic motivation, attitudes, needs satisfaction, engagement) and the varied technological approaches employed.

Cognitive Outcomes

Cognitive outcomes comprised academic achievement in physical education-related subjects including knowledge of rules, game tactics, health concepts, and standardized test performance; cognitive processing measures including executive functions, attention, memory, and processing speed assessed through computerized test batteries; and declarative knowledge assessments evaluating students' understanding of movement principles, fitness concepts, and sport-related content knowledge through written tests and oral examinations. Meta-analysis of 15 effect sizes from 4 studies revealed a pooled effect of $g = 0.936$ (95% CI [0.459, 1.412], $p < .001$), the numerically strongest domain-specific estimate, with extreme heterogeneity ($I^2 = 99.94\%$, $\tau^2 = 0.828$) and a 95% PI [-1.850, 3.721] reflecting the widest variability across domains, consistent with the diversity of cognitive constructs examined. Technologies such as exergames requiring strategic thinking and problem-solving, virtual reality simulations with embedded instructional content on biomechanics and physiology, and blended learning approaches combining online theoretical modules with practical applications appeared to facilitate cognitive development alongside physical activity. The integration of cognitive challenges within physical contexts through technology-enhanced instruction demonstrated measurable effects on students› academic and cognitive outcomes. This heterogeneity reflects the diversity of cognitive constructs assessed and the varied technological approaches, though the consistently positive direction across studies supports the potential for technology-enhanced cognitive learning in physical education contexts.

Across all three outcome domains, digital technologies demonstrated positive average effects (Physical: $g = 0.854$, 95% CI [0.546, 1.163]; Motivational-Affective: $g = 0.616$, 95% CI [0.450, 0.783]; Cognitive: $g = 0.936$, 95% CI [0.459, 1.412]; all $p < .001$). A formal test of differences between domains using a mixed-effects model revealed no statistically significant heterogeneity between domains ($p > .05$), indicating that point estimates do not differ significantly across outcome types. However, given the extreme heterogeneity within each domain ($I^2 > 99.9\%$) and the wide prediction intervals reported above, these pooled estimates should be interpreted as average tendencies rather than stable domain-level effects. The apparent consistency across domains does not imply that digital technologies

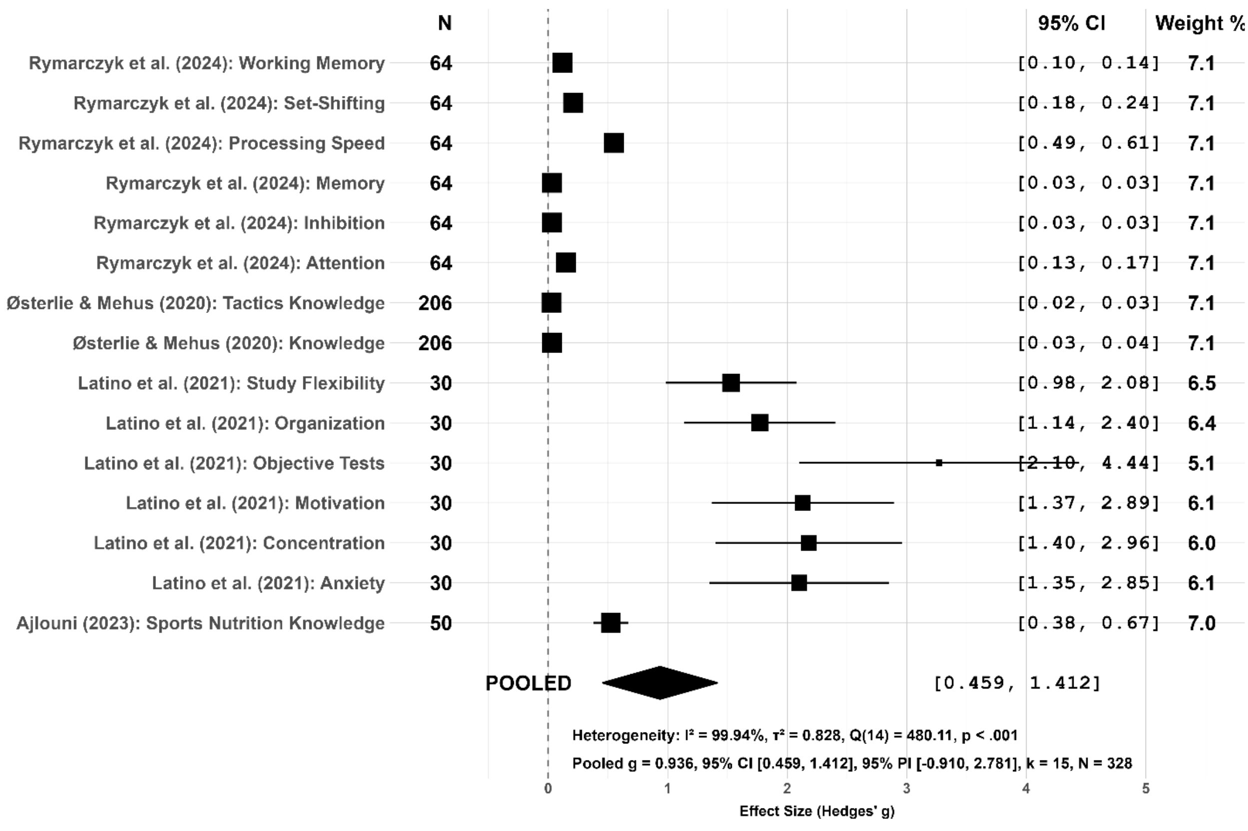

Fig. 5. Cognitive Outcomes (k = 15 ES; 4 studies; $I^2 = 99.94\%$)

Table 2. Summary table of meta-analytic findings

Domain	k	ES	N	g	95% CI	z	p	I^2	τ^2	Q(df)	95% PI
Overall	12	64	1,346	0.745	[0.590, 0.900]	8.45	<.001	99.98%	0.373	value	[-0.481, 1.971]
Physical Outcomes	7	21	1,346	0.854	[0.546, 1.163]	5.44	<.001	99.99%	0.482	1521.30(20)	[-0.536, 2.244]
Motivational-Affective	6	28	1,346	0.616	[0.450, 0.783]	7.26	<.001	99.98%	0.193	1777.84(27)	[-0.240, 1.473]
Cognitive Outcomes	4	15	1,346	0.936	[0.459, 1.412]	3.85	<.001	99.94%	0.828	940.81(14)	[-1.850, 3.721]

Note. k = studies; ES = effect sizes; g = Hedges' g pooled effect; CI = confidence interval; I^2 = heterogeneity percentage; τ^2 = between-study variance; Q(df) = Cochran's Q; PI = prediction interval. Random-effects models with REML estimation. Positive g favors digital technology interventions.

benefit all students equally, but rather that no single outcome domain is systematically more responsive than others at the aggregate level.

However, substantial heterogeneity within each domain (all $I^2 > 99\%$) indicates that effects varied considerably across specific technologies, implementations, and contexts within each outcome category. This pattern: consistency in positive effects across domains but high variability within domains. suggests that the effectiveness of digital technology interventions depends more on implementation quality, technology-pedagogy alignment, and contextual factors than on the specific outcome domain being targeted.

Publication Bias Assessment

Multiple approaches were employed to assess potential publication bias. Visual inspection of the funnel plot revealed asymmetry, with variation in precision across studies. Egger's regression test indicated significant funnel plot asymmetry ($z = 20.11$, $p < .001$). However, given the substantial heterogeneity observed across studies ($I^2 = 99.98\%$, $\tau^2 = 0.373$), this asymmetry likely reflects genuine variability in effect sizes across diverse contexts, technologies, and student populations rather than systematic publication bias. This interpretation is supported by the substantial range of effect sizes ($g = 0.02$ to 3.38) observed across studies examining different technologies, educational levels, and implementation contexts.

Trim-and-fill analysis estimated 0 potentially missing studies, indicating no systematic gaps in the distribution of effect sizes. The adjusted pooled effect size remained $g = 0.745$ (95% CI [0.590, 0.900]), representing 0% change from the observed effect. This finding suggests that the observed effects are not influenced by selective publication of positive results.

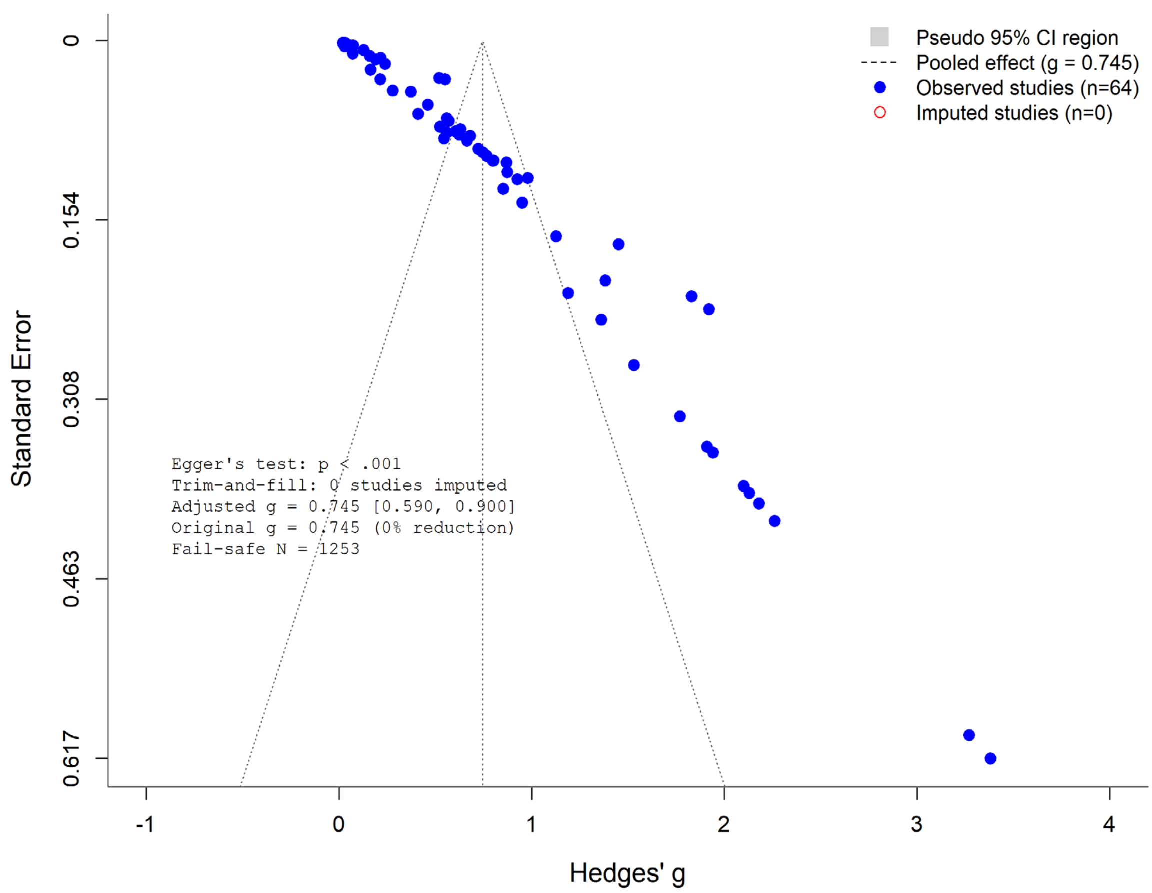

Fig. 6. Funnel Plot for Publication Bias Assessment

Rosenthal's fail-safe N analysis indicated that 1,253 unpublished studies with null results would be required to reduce the overall effect to non-significance ($p > .05$). This number substantially exceeds the tolerance threshold of $5k + 10 = 330$ studies (where $k = 64$ effect sizes), representing a ratio of 3.8 times the tolerance level. This suggests that the observed effect is robust to potential publication bias. Collectively, these analyses provide confidence that the findings reflect genuine intervention effects rather than selective publication. While funnel plot asymmetry was detected, it appears attributable to the substantial heterogeneity inherent in meta-analyses of educational technology interventions across diverse contexts, rather than systematic bias in the published literature.

Sensitivity Analyses

Leave-one-out sensitivity analyses revealed differential influence across studies (Figure 7). When each study was sequentially removed, pooled effects ranged from g = 0.533 to g = 0.816. Most studies (11 of 12) showed minimal influence when removed individually (reductions < 10%), with pooled effects ranging from g = 0.705 to g = 0.816. Latino et al. (2021) demonstrated substantial influence: removing this study reduced the pooled effect to g = 0.533 (95% CI [0.421, 0.646]), a 28.4% reduction. However, this study employed rigorous methodology (cluster-randomized design, validated measures, n = 30) and examined an intensive at-home intervention during COVID-19, which may explain larger effects. Importantly, even excluding Latino et al., the effect remains statistically significant and of moderate magnitude, supporting beneficial effects while highlighting variability in implementation success.

Cumulative meta-analysis showed reasonable stability after 40-50 effect sizes, with subsequent studies contributing primarily to precision. Sensitivity analyses excluding studies with small samples (n < 50; 4 studies) yielded pooled effects within g = 0.70-0.82, confirming robustness. These analyses indicate that digital technologies demonstrate consistent positive effects across diverse contexts, with effect magnitude varying based on intervention intensity and implementation quality. The conservative estimate excluding the most influential study (g = 0.533, moderate effect) and the full sample estimate (g = 0.745, large effect) both support beneficial impacts of digital technologies in physical education.

Moderator analyses

Moderator analyses examined three categorical variables (technology type, educational level, study design) and three continuous variables (publication year, sample size, intervention duration) using mixed-effects models with REML estimation.

Technology Type. Effect sizes varied significantly by technology type ($Q(6) = 161.53$, $p < .001$; Table 3). Blended learning demonstrated the largest effects (g = 1.216, 95% CI [0.883, 1.548], k = 27 ES), followed by mobile apps (g = 1.037, [0.833, 1.241], k = 4) and AI systems (g = 0.698, [0.629, 0.767], k = 12). Wearables/IoT devices showed substantially smaller effects (g = 0.112, [0.036, 0.188], k = 6). Despite significant moderation, residual heterogeneity remained high within all categories (I^2 = 43-99%), indicating that technology type explains only a portion of total variance.

Educational Level. Effect sizes varied significantly by educational level ($Q(3) = 100.08$, $p < .001$; Table 3). Secondary education demonstrated the largest effects (g = 1.185, 95% CI [0.748, 1.622], k = 21 ES), followed by tertiary (g = 0.650, [0.503, 0.796], k = 30) and elementary education (g = 0.451, [0.219, 0.683], k = 13). All levels showed significant

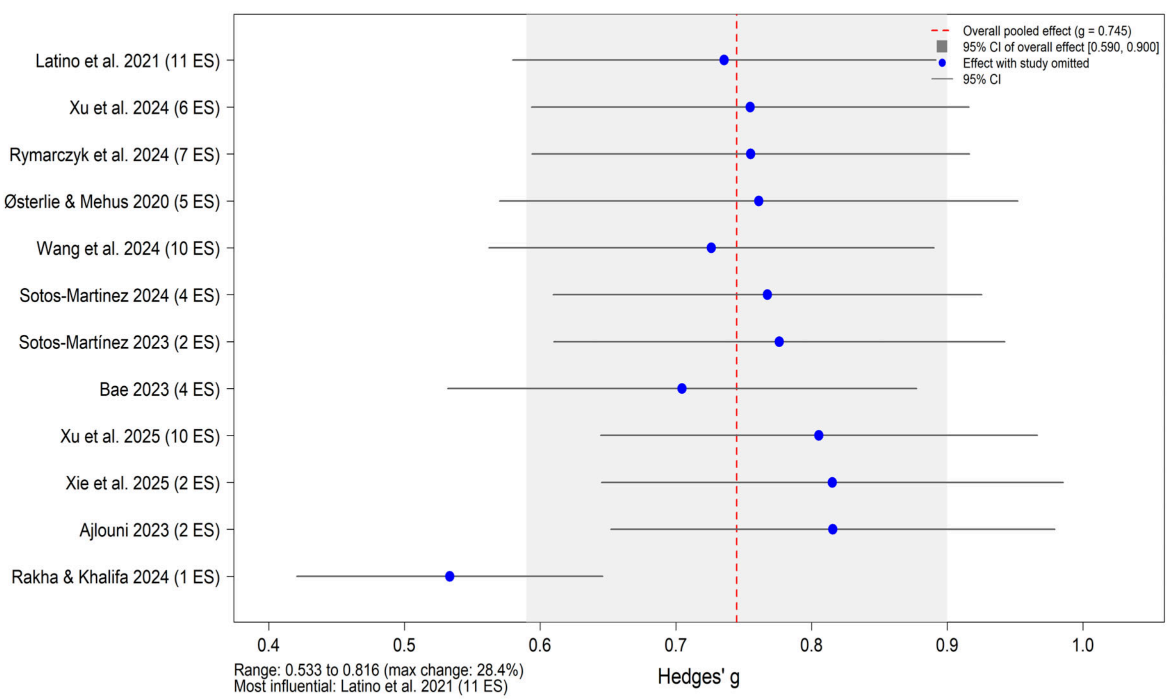

Fig. 7. leave-one-out sensitivity analysis plot

Table 3. Categorical moderator analysis

Moderator	Level	k	ES	g	95% CI	I^2	Q-between	p
Technology Type							161.53	< .001*
	Blended Learning	4	27	1.216	[0.883, 1.548]	99.99%		
	Mobile Apps	1	4	1.037	[0.833, 1.241]	43.8%		
	AI Systems	2	12	0.698	[0.629, 0.767]	43.4%		
	VR/AR	1	2	0.535	[0.490, 0.580]	0.0%		
	Gamification	3	13	0.252	[0.167, 0.337]	99.4%		
	Wearables/IoT	1	6	0.112	[0.036, 0.188]	99.8%		
Educational Level							100.08	< .001*
	Secondary	4	21	1.185	[0.748, 1.622]	99.99%		
	Tertiary	5	30	0.650	[0.503, 0.796]	99.9%		
	Elementary	3	13	0.451	[0.219, 0.683]	99.9%		
Study Design							91.72	< .001*
	RCT	4	34	0.884	[0.616, 1.152]	99.99%		
	Cluster-RCT	2	12	0.818	[0.491, 1.144]	99.9%		
	Quasi-Experimental	6	18	0.504	[0.320, 0.688]	99.98%		

Note. k = studies; ES = effect sizes; g = Hedges' g; CI = confidence interval; I^2 = residual heterogeneity within levels; Q-between = test for differences between levels. Mixed-effects models with REML. VR = virtual reality; AR = augmented reality; IoT = Internet of Things; RCT = randomized controlled trial. ***p < .001.

Table 4. Continuous moderators (meta-regression)

Moderator	k (ES)	β	SE(β)	95% CI	z	p	R^2
Publication Year	64	-0.119	0.051	[-0.219, -0.019]	-2.34	.019*	0.7%
Total Sample Size	64	-0.0047	0.0010	[-0.0066, -0.0028]	-4.90	< .001***	30.0%
Duration (weeks)	64	0.024	0.021	[-0.018, 0.065]	1.11	.265	1.6%

Note. k (ES) = effect sizes; β = regression coefficient; SE(β) = standard error; CI = confidence interval; z = z-statistic; R^2 = variance explained. Random-effects models with REML. None significant.

positive effects with substantial heterogeneity ($I^2 > 99\%$), demonstrating benefits across the educational continuum.

Study Design. Effect sizes varied significantly across study designs (Q(3) = 91.72, p < .001; Table 3). Randomized controlled trials demonstrated the largest effects (g = 0.884, 95% CI [0.616, 1.152], k = 34 ES), followed by cluster-randomized trials (g = 0.818, [0.491, 1.144], k = 12) and quasi-experimental studies (g = 0.504, [0.320, 0.688], k = 18), though all designs consistently supported beneficial impacts.

Continuous Moderators. Meta-regression analyses revealed that total sample size significantly predicted effect sizes (β = -0.0047, SE = 0.0010, p < .001, R^2 = 30.0%; Table 4), with larger studies associated with smaller effects, suggesting small-study effects. Publication year showed a small negative relationship with effect sizes (β = -0.119, SE = 0.051, p = .019, R^2 = 0.7%). Intervention duration did not significantly predict effect sizes (β = 0.024, SE = 0.021, p = .265, R^2 = 1.6%), indicating stable effectiveness across implementation periods ranging from 3 to 16 weeks.

Discussion

This systematic review and meta-analysis provides compelling evidence that digital technology integration produces large, meaningful benefits across physical, motivational-affective, and cognitive domains in physical education (g = 0.745, 95% CI [0.590, 0.900], I^2 = 99.98%). Analysis of 64 effect sizes from 12 rigorous experimental and quasi-experimental studies (N = 1,346 participants) reveals consistent positive effects, though substantial heterogeneity indicates that effectiveness varies considerably across implementations, technologies, and contexts.

Principal Findings and Their Interpretation

Digital technologies demonstrated large positive effects across all three outcome domains: Physical (g = 0.854, 95% CI [0.546, 1.163]), Motivational-Affective (g = 0.616, 95% CI [0.450, 0.783]), and Cognitive (g = 0.936, 95% CI [0.459, 1.412]). These magnitudes substantially exceed typical educational technology meta-analyses (Schmid et al., 2014, mean d = 0.33) and previous physical education technology reviews (Gao et al., 2020, d = 0.42). Using benchmarks for educational interventions (Kraft, 2020), effects exceeding g = 0.25 are considered educationally significant; the present findings represent effects approximately three times this threshold, indicating substantial practical importance beyond statistical significance.

Comparable magnitudes across domains indicate that well-designed digital interventions produce broad benefits rather than selectively enhancing specific outcomes. These finding challenges assumptions that technology primarily benefits cognitive or motivational outcomes while having limited impact on physical performance, or conversely, that technology-based physical activity enhances motor outcomes without addressing psychological or academic dimensions. These findings converge with recent evidence that technology-enhanced and adapted pedagogical approaches can enhance student motivation, psychological wellbeing, and engagement in physical education (Ben Rakaa et al., 2025b). The large positive effects on physical outcomes (g = 0.854) demonstrate that technology integration need not compromise physical activity time, addressing common concerns about sedentary technology use. Multiple mechanisms likely operate: wearable devices and mobile apps providing real-time biofeedback enable students to monitor physiological parameters, enhancing self-regulation and goal-directed behavior. Virtual reality applications and gamified systems may increase practice volume by enhancing engagement and reducing perceived exertion through attentional focus on game elements. Video analysis tools provide immediate visual feedback on movement execution, facilitating error detection and motor learning.

The moderate-to-large motivational effects (g = 0.616) are particularly noteworthy given widespread concerns about declining adolescent engagement in physical education. Consistent with Self-Determination Theory (Ryan & Deci, 2017), effective interventions likely enhanced autonomy through learner choice, competence through immediate feedback, and relatedness through collaborative features. The inclusion of behavioral engagement within motivational-affective outcomes aligns with Fredricks et al.'s (2004) multidimensional engagement framework conceptualizing behavioral engagement as the observable manifestation of motivational states. The positive effects suggest that digital technologies not only influence psychological states but also translate into observable increases in active participation, effort intensity, and persistence.

The large cognitive effects (g = 0.936) represent the numerically strongest domain-specific impact, challenging historical neglect of cognitive and academic outcomes in physical education research. Several mechanisms may explain these effects. Exergames require strategic thinking, problem-solving, and decision-making embed cognitive challenges within physically active contexts, potentially enhancing executive functions through dual-task training. Virtual reality simulations incorporating embedded instructional content on biomechanics, physiology, and tactical concepts integrate declarative knowledge acquisition with experiential learning. Blended learning approaches combining online theoretical modules with face-to-face practical application leverage complementary strengths: online environments enable self-paced learning, multimedia presentation, and immediate feedback, while practical sessions provide embodied application and skill development. These findings challenge traditional dichotomies separating mind and body in educational contexts and support integrated approaches wherein physical education contributes meaningfully to students' academic and cognitive development.

Technology type emerged as the strongest moderator (Q = 161.53, p < .001), with blended learning approaches demonstrating the largest effects (g = 1.216, [0.883, 1.548]), followed by mobile apps (g = 1.037), AI systems (g = 0.698), and VR/AR (g = 0.535). Wearables and IoT devices showed substantially smaller effects (g = 0.112), while gamification produced moderate effects (g = 0.252). This hierarchy aligns with educational technology frameworks emphasizing that pedagogical integration matters more than technological sophistication. Technologies embedding digital tools within structured instructional frameworks outperformed autonomous activity tracking tools, suggesting that technology serves as a pedagogical amplifier rather than autonomous solution. The lower effects for wearables (g = 0.112) compared to blended learning (g = 1.216) suggest that passive monitoring without pedagogical integration produces limited benefits. The modest effects for gamification specifically (g = 0.252) may reflect critical design differences in autonomy support. Technologies emphasizing external rewards (points, badges, leaderboards) without supporting autonomy may produce short-term engagement increases but potentially undermine long-term intrinsic motivation through overjustification effects. Conversely, technologies providing informational feedback supporting competence without controlling external contingencies may enhance intrinsic motivation more sustainably.

Educational level significantly moderated effects (Q = 100.08, p < .001), with secondary education showing largest impacts (g = 1.185, [0.748, 1.622]), followed by tertiary (g = 0.650) and elementary (g = 0.451). However, all levels demonstrated significant positive effects, indicating benefits across the developmental continuum when interventions are developmentally appropriate. The larger effects at secondary level may reflect adolescents' greater technological fluency, autonomy in self-regulated learning, and alignment between digital tools and developmental needs for competence feedback and social connection. Study design also moderated effects (Q = 91.72, p < .001), with RCTs demonstrating largest effects (g = 0.884), followed by cluster-RCTs (g = 0.818) and quasi-experimental designs (g = 0.504). While this pattern suggests rigorous experimental control may capture larger effects, the significant positive effects across all designs strengthen confidence that benefits are not artifacts of particular methodological approaches.

Meta-regression analyses revealed that sample size significantly predicted effect sizes (β = -0.0047, p < .001, R^2 = 30.0%), with larger studies associated with smaller effects. This substantial negative relationship suggests small-study effects, wherein smaller studies tend to report larger effects. This pattern may indicate publication bias favoring positive findings in smaller studies, methodological limitations in smaller samples, or genuine heterogeneity in implementation quality. While trim-and-fill analysis detected no missing studies and fail-safe N (1,253 studies) substantially exceeded tolerance thresholds, the small-study relationship warrants interpretive caution. The pooled effect (g = 0.745) likely represents an optimistic estimate; effects in large-scale implementations may be more modest. Publication year showed a small negative relationship with effect sizes (β = -0.119, p = .019, R^2 = 0.7%), indicating slightly smaller effects in more recent studies, though this explained minimal variance. Intervention duration did not significantly predict effects (p = .265, R^2 = 1.6%), indicating stable effectiveness across implementation periods from

3 to 16 weeks, suggesting flexibility in program design and feasibility for time-constrained educational contexts.

Despite consistent average effects, extreme heterogeneity characterized all analyses ($I^2 > 99.9\%$), with effect sizes ranging from g = 0.02 to 3.38. This variability indicates that implementation quality, contextual factors, and alignment between technological affordances and pedagogical goals critically determine outcomes. Wide prediction intervals (e.g., overall 95% PI [-0.444, 1.935]) suggest that while average effects are large and positive, some implementations may produce negligible or even negative effects. Sensitivity analyses revealed that one study (Latino et al., 2021) exerted disproportionate influence: removing this intensive COVID-19 lockdown intervention reduced pooled effects to g = 0.533 (moderate), a 28.4% reduction. However, this conservative estimate remains statistically significant and practically meaningful. The study's large effects (g = 1.91-3.38) likely reflect the intensive, comprehensive nature of the intervention (daily guided sessions, personalized feedback, academic integration) rather than methodological artifacts. This finding underscores that effect magnitude depends substantially on implementation intensity and comprehensiveness.

Publication Bias and Robustness of Findings

Publication bias assessment yielded mixed results requiring careful interpretation. Egger's test indicated significant funnel plot asymmetry ($z = 20.11$, $p < .001$), suggesting potential publication bias. However, the extreme heterogeneity ($I^2 = 99.98\%$) limits Egger's test validity, as the test assumes relatively homogeneous effects. Funnel plot asymmetry can result from heterogeneity, methodological quality differences, or true publication bias. The trim-and-fill analysis (R0 estimator) estimated zero missing studies, suggesting asymmetry reflects heterogeneity rather than systematic suppression of negative findings. Fail-safe N (1,253 studies, 3.8× tolerance threshold) indicates robustness. Collectively, evidence suggests that while some publication bias may exist (as indicated by small-study effects), it likely does not fully explain observed effects. The conservative estimate excluding the most influential study (g = 0.533) provides a lower-bound estimate less susceptible to publication bias concerns, while the full sample estimate (g = 0.745) may be somewhat optimistic. True population effects likely fall within this range (0.533-0.745), representing educationally meaningful impacts with moderate-to-large practical significance.

Interpreting Effects Under Extreme Heterogeneity

A critical consideration in interpreting our findings is the extreme heterogeneity observed across all meta-analytic models ($I^2 \approx 99–100\%$). While the pooled effect sizes indicate generally positive impacts of digital technologies on physical education outcomes, the prediction intervals reveal that these effects are far from uniform. The overall 95% PI [-0.471, 1.961] encompasses both negative and very large positive effects, meaning that any single implementation of digital technology in physical education could yield outcomes anywhere within this range. This heterogeneity is not unexpected given the diversity of (a) technology types examined, (b) outcome constructs measured, (c) educational contexts, and (d) study designs and intervention durations. Accordingly, the pooled effect sizes reported in this review should be understood as describing the average direction and approximate magnitude of effects across a highly heterogeneous evidence base, rather than as precise predictions of what any particular implementation will achieve. The moderator analyses provide more actionable guidance: technology type, educational level, and study design each significantly moderated effects, suggesting that the effectiveness of digital technologies in physical education depends critically on how, where, and with whom they are implemented. These findings align with Borenstein et al. (2009) recommendation that when prediction intervals are wide, researchers should prioritize identifying sources of heterogeneity over interpreting the pooled mean as a definitive estimate.

Implications for Practice and Policy

The findings yield several actionable implications for physical education practice. First, practitioners should prioritize pedagogically integrated approaches (blended learning, structured gamification, AI-enhanced instruction) rather than focusing on technological sophistication or novelty. The technology type moderation results demonstrate that blended learning combining online and face-to-face instruction, gamification incorporating game elements within structured curricula, and AI systems providing personalized feedback substantially exceeded wearable activity trackers. This emphasizes that technology effectiveness depends on instructional design and teacher mediation rather than technological features alone.

Second, the extreme heterogeneity ($I^2 > 99.9\%$) and wide prediction intervals indicate that implementation quality substantially determines outcomes. Simply providing technological resources without implementation support likely produces inconsistent results. Teachers require not only technical skills but also technological pedagogical content knowledge, understanding how to effectively integrate specific technologies to address particular learning goals with specific student populations. Sustained, practice-embedded professional development focusing on pedagogical integration produces larger effects than one-time technical training. The critical role of teacher preparation aligns with emerging evidence on inclusive physical education, where teacher competence, pedagogical intentionality, and institutional support determine implementation success (Ben Rakaa et al., 2024). Technologies require not merely operational proficiency but pedagogical expertise to adapt instruction to diverse student needs and learning contexts.

Third, while effects are large and meaningful, digital technologies are tools that augment rather than replace effective teaching. Fundamental pedagogical practices (clear learning objectives, appropriate challenge, supportive climate, opportunities for practice and feedback) remain essential. The conservative estimate excluding the most influential study (g = 0.533) provides realistic expectations for typical implementations. Technologies integrated superficially produce smaller effects than those central to instructional design.

Fourth, the significant positive effects across elementary, secondary, and tertiary education suggest concerns about age-appropriateness should focus on implementation features (interface design, content complexity, scaffolding) rather than categorical age exclusions. All developmental levels benefit from appropriately designed interventions, though specific implementations necessarily differ in complexity and autonomy requirements. Beyond these general implementation considerations, ensuring equitable access across diverse student populations warrants particular attention.

Equity and Accessibility Considerations

An important dimension of technology integration concerns ensuring equitable access and inclusion for students with diverse needs and abilities. Research on universal design for learning and adapted physical education demonstrates that digital technologies can facilitate differentiated instruction and inclusive participation when implemented within appropriate pedagogical frameworks (Ben Rakaa et al., 2025a; Lourenço et al., 2025). However, our findings underscore that technology availability alone does not guarantee equitable outcomes. The wide prediction intervals and substantial moderating effects of implementation quality suggest that technologies must be deliberately adapted to diverse student populations, with particular attention to students requiring additional pedagogical support. The relatively modest effects observed for wearable activity trackers ($g = 0.112$) may reflect challenges in making such technologies accessible and motivating for all students, rather than inherent technological limitations. Technologies demonstrating larger effects, such as blended learning ($g = 1.216$) and mobile applications ($g = 1.037$), typically incorporated explicit pedagogical scaffolding and differentiation mechanisms enabling adaptation to individual student needs. Future research should explicitly examine how different technologies perform across diverse student populations, including students with disabilities, low prior achievement, or limited technology access, to ensure that digital innovation promotes rather than undermines educational equity.

Study Limitations and Future Directions

Several limitations warrant acknowledgment. Despite comprehensive moderator analyses, extreme heterogeneity ($I^2 > 99.9\%$) persists, indicating that unmeasured variables (implementation fidelity, teacher expertise, student characteristics, contextual factors) substantially influence outcomes. The field lacks standardized implementation frameworks and fidelity measures, limiting ability to identify critical implementation components. The relatively small number of studies ($k = 12$), though yielding 64 effect sizes, limits statistical power for detecting moderator effects and prevents examination of potentially important moderators including teacher characteristics, specific design features, and student demographics.

While trim-and-fill and fail-safe N analyses suggested robustness, significant small-study effects ($R^2 = 30.0\%$) and Egger test asymmetry indicate some publication bias may exist. The field may underrepresent null or negative findings, potentially inflating observed effects. Most interventions represented relatively short-term implementations (3-16 weeks), with only one study including follow-up assessment. Whether effects persist beyond intervention periods, whether sustained implementation produces cumulative benefits, or whether initial enthusiasm effects fade remains unclear. Studies originated primarily from China, Spain, and Poland, limiting generalizability to diverse cultural and educational contexts. Most studies (67%) employed quasi-experimental or cluster-randomized designs rather than individual randomization, introducing potential confounding. Additionally, blinding of outcome assessors was infeasible for many outcomes, potentially introducing detection bias. Potential negative outcomes (excessive competition, privacy concerns, technology-induced anxiety, motivation undermining through overjustification) were rarely assessed or reported, limiting understanding of potential iatrogenic effects.

Future research should prioritize implementation science approaches investigating how technologies are implemented, what factors predict effectiveness, and how implementation can be supported. Mixed-methods studies combining effectiveness assessment with qualitative investigation of implementation processes would illuminate mechanisms linking technology provision to student outcomes. Longitudinal studies with extended follow-up addressing maintenance, cumulative effects, and transfer to autonomous physical activity outside school contexts would address critical questions about durability. Randomized trials systematically varying specific features (e.g., autonomy support elements, feedback timing and content, reward structures) while controlling other factors would isolate critical design components. Research examining moderators inadequately represented (socioeconomic status, special educational needs, gender, prior technology experience, teacher characteristics) would enhance understanding of for whom technologies are most effective. Mediation analyses testing hypothesized mechanisms would strengthen theoretical foundations and guide more targeted interventions. Investigation of circumstances under which technologies might undermine intrinsic motivation, exacerbate social comparison concerns, or produce anxiety would support more balanced, ethically informed implementation decisions.

Conclusion

This systematic review and meta-analysis synthesized 64 effect sizes from 12 studies ($N = 1,346$) examining digital technology integration in physical education. The findings indicate that digital technologies are, on average, associated with positive effects across physical ($g = 0.854$), motivational-affective ($g = 0.616$), and cognitive ($g = 0.936$) domains. However, extreme heterogeneity across all models ($I^2 > 99.9\%$) and wide prediction intervals, notably the overall 95% PI [−0.471, 1.961], indicate that these pooled estimates should not be interpreted as stable or universally applicable measures of effectiveness. The true effect of any specific implementation depends substantially on technology type, pedagogical design, educational level, and contextual factors. Technology type emerged as the strongest moderator ($Q = 161.53$, $p < .001$), with blended learning ($g = 1.216$) and mobile apps ($g = 1.037$) substantially outperforming wearable trackers ($g = 0.112$), confirming that pedagogical integration matters more than technological sophistication. Rather than supporting a blanket endorsement of digital technologies in physical education, these findings suggest that such technologies hold considerable promise, but their effectiveness is highly context-dependent. Practitioners should carefully consider the specific technology, student population, and pedagogical integration strategy when making adoption decisions.

However, several considerations warrant careful interpretation. Extreme heterogeneity characterized all analyses ($I^2 > 99.9\%$, effect size range: 0.02-3.38), indicating that implementation quality, contextual factors, and pedagogical alignment critically determine outcomes. While average effects provide compelling evidence for potential benefits, individual implementations vary considerably in effectiveness. Sensitivity analyses revealed that removing the most influential study (Latino et al., 2021) reduced pooled effects to $g = 0.533$, suggesting that typical implementations may achieve moderate rather than large effects. Evidence of small-study effects (larger studies associated with smaller effects, $R^2 = 30.0\%$) and significant funnel plot asymmetry, though counterbalanced by null trim-and-fill results and robust fail-safe N (1,253 studies), suggests moderate evidence for publication bias. The true population effect likely falls between the conservative estimate ($g = 0.533$) and the full sample estimate ($g = 0.745$), both representing educationally meaningful impacts but indicating that realistic expectations are warranted.

The findings yield actionable implications for practice and policy. Physical education teachers require professional development supporting technological pedagogical content knowledge, understanding not merely how to operate technologies but how to integrate them effectively to achieve specific learning goals with diverse student populations. The technology type moderation suggests prioritizing pedagogically-integrated approaches (blended learning, structured gamification, AI-enhanced instruction) over passive monitoring tools. Administrators must recognize that technology integration represents a long-term investment requiring sustained support for professional development and implementation quality rather than one-time technology purchases. Policymakers should prioritize resources for implementation support and teacher preparation alongside technology acquisition, recognizing that tools without skilled implementation produce limited benefits.

Future research should prioritize implementation science approaches examining how technologies are implemented in authentic contexts, what factors support effective implementation, and how implementation fidelity can be assessed and supported. The field needs standardized implementation frameworks and fidelity measures to identify critical components. Longitudinal studies tracking effects beyond immediate post-intervention assessments would clarify whether observed benefits persist, fade, or strengthen with sustained exposure. Comparative effectiveness research directly contrasting specific technological approaches through randomized trials would provide more definitive guidance. Research examining equity dimensions (socioeconomic status, special educational needs, gender, prior technology experience, diverse cultural contexts) would ensure that technology integration benefits all students. Investigation of potential negative outcomes (motivation undermining, anxiety, privacy concerns) would support more balanced, ethically-informed implementation decisions.

The present findings support evidence-based optimism regarding digital technology's potential to substantially enhance physical education effectiveness. Technologies aligned with sound pedagogical principles, integrated within comprehensive instructional frameworks, and supported by teacher expertise can meaningfully enhance student outcomes across physical, motivational, and cognitive domains. However, extreme heterogeneity, small-study effects, and differential effectiveness by technology type underscore that these benefits depend critically on implementation quality. Technology alone, absent thoughtful pedagogical design, skilled implementation, and systemic support, proves insufficient for realizing this potential.

The meta-analysis revealed both promise and complexity. Large average effects indicate substantial potential benefits, but extreme heterogeneity and wide prediction intervals demonstrate that this potential is realized only when technologies are thoughtfully integrated within sound pedagogical frameworks. As physical education navigates an increasingly digital educational landscape, success depends on maintaining fidelity to the field's fundamental goals (promoting physically active lifestyles, developing motor competence, fostering positive relationships with physical activity) while strategically leveraging technological affordances to enhance rather than replace effective teaching. The field has moved beyond whether digital technologies work to the more nuanced questions of how, for whom, and under what conditions they are most effective. Digital technologies represent powerful tools that can substantially amplify effective teaching when implemented skillfully within comprehensive instructional designs aligned with physical education's educational mission.

Ethics Approval and Consent to Participate

Not applicable. This systematic review and meta-analysis synthesized data from previously published studies and did not involve primary data collection from human participants. Therefore, ethics approval was not required.

Consent for Publication

Not applicable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The study was conducted as part of doctoral research at École Normale Supérieure de Casablanca with institutional support for basic research infrastructure. No external funding was obtained.

Conflict of Interest

The authors declare no financial, personal, institutional, or professional conflicts of interest that could have influenced the design, conduct, analysis, or reporting of this study.

Author Contributions

Narrative Statement:

YM conceptualized the study, designed the methodology, conducted the systematic literature search and screening, performed data extraction and quality assessment, conducted all statistical analyses using R, prepared figures and tables, and drafted the manuscript. SL supervised the research design, provided methodological guidance on meta-analytic procedures, validated analytical strategies, and critically revised the manuscript for important intellectual content. SO contributed to study selection and screening, independently verified data extraction, participated in risk of bias assessment, and reviewed and edited the manuscript. All authors read and approved the final version of the manuscript.

CRediT Taxonomy:

Conceptualization: YM, SL; Data curation: YM; Formal analysis: YM; Investigation: YM, SO; Methodology: YM, SL; Project administration: YM; Software: YM; Supervision: SL, SO; Validation: SL, SO; Visualization: YM; Writing – original draft: YM; Writing – review & editing: YM, SL, SO.

Data Availability

All data supporting the findings of this systematic review and meta-analysis are openly available via the Open Science Framework at https://osf.io/3jyn2 (DOI: 10.17605/OSF.IO/3JYN2).

The repository includes: extracted effect sizes, coding sheets, PRISMA 2020 checklist, full search strategies (Scopus, Web of Science, ERIC, SPORTDiscus), risk of bias assessments, excluded studies register with reasons, complete R code (metafor package), forest plots by domain, and sensitivity analyses. The dataset ensures full reproducibility of all meta-analytic procedures.

AI Transparency Statement

Artificial intelligence tools were used at strictly delimited stages of this research. Elicit Pro assisted with initial abstract screening and structured data extraction; all outputs were independently verified by two reviewers against the original publications. Claude assisted with manuscript drafting and language refinement; all AI-generated content was critically reviewed and substantially revised by the authors.

All statistical analyses were performed exclusively by the first author in R (version 4.3.0) using the metafor package. AI tools served solely as efficiency-support instruments under full human supervision and had no influence on study selection, methodological decisions, statistical analyses, or scientific conclusions. All authors approved this statement.

Acknowledgements

This study was conducted within the doctoral research program of École Normale Supérieure de Casablanca (ENS-C), Université Hassan II, LMSEIF laboratory, under the supervision of Prof. Said Lotfi and Prof. Siham Ouhrir. The authors thank the researchers whose primary studies were included in this meta-analysis for their contributions to the field of digital technologies in physical education.

References

Casey, A., & Goodyear, V.A. (2015). Can cooperative learning achieve the four learning outcomes of physical education? A review of literature. *Quest, 67*(1), 56-72. https://doi.org/10.1080/00336297.2015.1007885

Kalman, M., Inchley, J., Sigmundova, D., Iannotti, R.J., Tynjälä, J.A., Hamrik, Z., Haug, E., & Bucksch, J. (2015). Secular trends in moderate-to-vigorous physical activity in 32 countries from 2002 to 2010: A cross-national perspective. *European Journal of Public Health, 25*(Suppl. 2), 37-40. https://doi.org/10.1093/eurpub/ckv024

Ladwig, M.A., Vazou, S., & Ekkekakis, P. (2018). «My best memory is when I was done with it»: PE memories are associated with adult sedentary behavior. *Translational Journal of the American College of Sports Medicine, 3*(16), 119-129. https://doi.org/10.1249/TJX.0000000000000068

Gråstén, A. (2016). Children's expectancy beliefs and subjective task values through two years of school-based program and associated links to physical education enjoyment and physical activity. *Journal of Sport and Health Science, 5*(4), 500-505. https://doi.org/10.1016/j.jshs.2016.01.007

Gao, Z., Chen, S., Pasco, D., & Pope, Z. (2020). A meta-analysis of active video games on health outcomes among children and adolescents. *Obesity Reviews, 16*(9), 783-794. https://doi.org/10.1111/obr.12287

Papastergiou, M., Gerodimos, V., Vernadakis, N., & Antoniou, P. (2021). Effectiveness of virtual reality in teaching and learning in physical education: A systematic review. *Journal of Educational Computing Research, 59*(5), 885-910. https://doi.org/10.1177/0735633120985104

Hu, X., & Li, J. (2024). Research on the integrated solution of physical education based on smart campus. *Advances in Physiology Education, 48*(2), 378-384. https://doi.org/10.1152/advan.00006.2024

Baena-Morales, S., Jerez-Mayorga, D., Delgado-Floody, P., & Martínez-Martínez, J. (2021). Sustainable development goals and physical education: A proposal for practice-based models. *International Journal of Environmental Research and Public Health, 18*(4), Article 2129. https://doi.org/10.3390/ijerph18042129

Hsia, L.-H., Lin, Y.-N., Lin, C.-H., & Hwang, G.-J. (2025). Prompting somatic practice performance with AI-facilitated peer-assisted learning: A self-determination theory perspective. *Education and Information Technologies.* Advance online publication. https://doi.org/10.1007/s10639-025-13681-8

Koekoek, J., & van Hilvoorde, I. (2018). *Digital technology in physical education: Global perspectives.* Routledge.

Ryan, R.M., & Deci, E.L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press.

Vazou, S., & Skrade, M.A. B. (2017). Intervention integrating physical activity with math: Math performance, perceived competence, and need satisfaction. *International Journal of Sport and Exercise Psychology, 15*(5), 508-522. https://doi.org/10.1080/1612197X.2016.1164226

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Tannehill, D., Demirhan, G., Čaplová, P., & Avsar, Z. (2021). Continuing professional development for physical education teachers in Europe. *European Physical Education Review, 27*(1), 150-167. https://doi.org/10.1177/1356336X20931531

Li, D., & Sun, Y. (2024). Improvement of students' sports performance in VR-based smart sports teaching in colleges and universities. *Applied Mathematics and Nonlinear Sciences, 9*(1), Article 1373. https://doi.org/10.2478/amns-2024-1373

Shekerbekova, S., Kamalova, G., Iskakova, M., Aldabergenova, A., Abdykerimova, E., & Shetiyeva, K. (2025). Exploring the use of artificial intelligence and augmented reality tools to improve interactivity in physical education teaching and training methods. *Retos,* 66, 1132-1139. https://doi.org/10.47197/retos.v66.113540

Ibragimova, E., Uraimov, S., Baitassov, Y., Yuldasheva, S., Kutlimuratova, D., & Litwinowa, M. (2025). Digital motivation: Fitness apps and student physical activity. *Retos,* 67, 1162-1173. https://doi.org/10.47197/retos.v67.113635

Breed, R., Kay, A., Spittle, M., & Orth, D. (2024). The effect of pedagogical approach on physical activity of girls during physical education. *Research Quarterly for Exercise and Sport, 95*(4), 824-833. https://doi.org/10.1080/02701367.2024.2329165

Dewanti, G., Nopembri, S., & Hartanto, A. (2024). Development of blended learning based learning module: Benefits to physical education learning outcomes. *Fizjoterapia Polska, 2024*(1), 12-17. https://doi.org/10.56984/8ZG2EF83C3

Dudley, D., Pearson, P., & Okely, A. (2022). A systematic review of the effectiveness of physical education interventions on student physical activity, physical fitness and health outcomes. *European Physical Education Review, 28*(3), 685-707. https://doi.org/10.1177/1356336X221078829

Bailey, R., Armour, K., Kirk, D., Jess, M., Pickup, I., Sandford, R., & BERA Physical Education and Sport Pedagogy Special Interest Group. (2009). The educational benefits claimed for physical education and school sport: An academic review. *Research Papers in Education, 24*(1), 1-27. https://doi.org/10.1080/02671520701809817

Lai, C., Wang, Q., & Lei, J. (2022). What factors predict undergraduate students' use of technology for learning? A case from Hong Kong. *Computers & Education, 59*(2), 569-579. https://doi.org/10.1016/j.compedu.2012.03.006

Ennis, C.D. (2017). Educating students for a lifetime of physical activity: Enhancing mindfulness, motivation, and meaning. *Research Quarterly for Exercise and Sport, 88*(3), 241-250. https://doi.org/10.1080/02701367.2017.1342495

Ben Rakaa, O., Bassiri, M., & Lotfi, S. (2025a). Adapted pedagogical strategies in inclusive physical education for students with special educational needs: a systematic review. *Pedagogy of Physical Culture and Sports, 29*(2), 67-85. https://doi.org/10.15561/26649837.2025.0201

Lourenço, C., Ben Rakaa, O., Bassiri, M., & Lotfi, S. (2025). Unravelling the Impact of Universal Design for Learning on the Inclusion of Students with Disabilities in Physical Education: A Systematic Review. *Physical Education Theory and Methodology, 25*(3), 691-704. https://doi.org/10.17309/tmfv.2025.3.27

Armour, K.M., & Harris, J. (2013). Making the case for developing new PE-for-health pedagogies. *Quest, 65*(2), 201-219. https://doi.org/10.1080/00336297.2013.773531

Hollis, J.L., Sutherland, R., Williams, A.J., Campbell, E., Nathan, N., Wolfenden, L., Morgan, P.J., Lubans, D.R., & Wiggers, J. (2017). A systematic review and meta-analysis of moderate-to-vigorous physical activity levels in secondary school physical education lessons. *International Journal of Behavioral Nutrition and Physical Activity, 14*(1), Article 52. https://doi.org/10.1186/s12966-017-0504-0

Lupton, D. (2015). *Digital sociology*. Routledge.

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

Falagas, M.E., Pitsouni, E.I., Malietzis, G.A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *The FASEB Journal, 22*(2), 338-342. https://doi.org/10.1096/fj.07-9492LSF

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics, 106*(1), 213-228. https://doi.org/10.1007/s11192-015-1765-5

Bramer, W.M., Rethlefsen, M.L., Kleijnen, J., & Franco, O.H. (2017). Optimal database combinations for literature searches in systematic reviews: A prospective exploratory study. *Systematic Reviews, 6*(1), 245. https://doi.org/10.1186/s13643-017-0644-y

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan -a web and mobile app for systematic reviews. Systematic Reviews, 5(1), Article 210. https://doi.org/10.1186/s13643-016-0384-4

Sterne, J.A. C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C.J., Cheng, H.Y., Corbett, M.S., Eldridge, S.M., Emberson, J.R., Hernán, M.A., Hopewell, S., Hróbjartsson, A., Junqueira, D.R., Jüni, P., Kirkham, J.J., Lasserson, T., Li, T., ... Higgins, J.P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ,* 366, Article l4898. https://doi.org/10.1136/bmj.l4898

Sterne, J.A. C., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., Viswanathan, M., Henry, D., Altman, D.G., Ansari, M.T., Boutron, I., Carpenter, J.R., Chan, A.-W., Churchill, R., Deeks, J.J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y.K., Pigott, T.D., ... Higgins, J.P. T. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ,* 355, Article i4919. https://doi.org/10.1136/bmj.i4919

Wang, C., Yuan, Y., & Ji, X. (2024). Effects of blended learning in physical education on university students' exercise attitudes and basketball skills: A cluster randomized controlled trial. *BMC Public Health, 24*(1), 3170. https://doi.org/10.1186/s12889-024-20469-x

Xu, Y., Peng, J., Jing, F., & Ren, H. (2024). From wearables to performance: How acceptance of IoT devices influences physical education results in college students. *Scientific Reports, 14*(1), 23776. https://doi.org/10.1038/s41598-024-75071-3

Bloom, B.S., Englehart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain.* David McKay Company.

Krathwohl, D.R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Deci, E.L., & Ryan, R.M. (2000). The «what» and «why» of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

Pelletier, L.G., Fortier, M.S., Vallerand, R.J., Tuson, K.M., Brière, N.M., & Blais, M.R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: The Sport Motivation Scale (SMS). *Journal of Sport and Exercise Psychology, 17*(1), 35-53. https://doi.org/10.1123/jsep.17.1.35

Fredricks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109. https://doi.org/10.3102/00346543074001059

Ajlouni, A. O. (2023). The impact of mobile application-assisted instruction on intrinsic motivation and sports nutrition knowledge: The case of blended learning. *International Journal of Emerging Technologies in Learning, 18*(22), 32-46. https://doi.org/10.3991/ijet.v18i22.43731

Hedges, L.V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107-128. https://doi.org/10.3102/10769986006002107

Borenstein, M., Hedges, L.V., Higgins, J.P. T., & Rothstein, H.R. (2009). *Introduction to meta-analysis.* John Wiley & Sons.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis.* SAGE Publications.

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

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics, 30*(3), 261-293. https://doi.org/10.3102/10769986030003261

R Core Team. (2023). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Sotos-Martínez, V.J., Tortosa-Martínez, J., Baena-Morales, S., & Ferriz-Valero, A. (2023). It's game time: Improving basic psychological needs and promoting positive behaviours through gamification in physical education. *European Physical Education Review.* Advance online publication. https://doi.org/10.1177/1356336X231217404

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

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

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. https://doi.org/10.18637/jss.v036.i03

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

Xu, L.-J., Wu, J., Zhu, J.-D., & Chen, L. (2025). Effects of AI-assisted dance skills teaching, evaluation and visual feedback on dance students' learning performance, motivation and self-efficacy. *International Journal of Human-Computer Studies,* 195, 103410. https://doi.org/10.1016/j.ijhcs.2024.103410

Xie, D., Chen, B., Li, H., & Huang, Q. (2025). Enhancing table tennis performance and physical fitness through digital technology: A quasi-experimental study based on the TPACK framework. *Frontiers in Sports and Active Living,* 7, 1595455. https://doi.org/10.3389/fspor.2025.1595455

Xu, Y., Peng, J., Jing, F., & Ren, H. (2024). From wearables to performance: How acceptance of IoT devices influences physical education results in college students. *Scientific Reports, 14*(1), 23776. https://doi.org/10.1038/s41598-024-75071-3

Rakha, A., & Khalifa, M. A. (2024). Blended learning using Edmodo: Students' performance and attitude in boxing during COVID-19. *Sage Open, 14*(2), 21582440241253743. https://doi.org/10.1177/21582440241253743

Sotos-Martínez, V. J., Ferriz-Valero, A., García-Martínez, S., & Tortosa-Martínez, J. (2024). The effects of gamification on the motivation and basic psychological needs of secondary school physical education students. *Physical Education and Sport Pedagogy, 29*(2), 160-176. https://doi.org/10.1080/17408989.2022.2039611

Wickham, H. (2016). ggplot2: *Elegant graphics for data analysis.* Springer-Verlag.

Latino, F., Fischetti, F., Cataldi, S., Monacis, D., & Colella, D. (2021). The impact of an 8-weeks at-home physical activity plan on academic achievement at the time of COVID-19 lock-down in Italian school. *Sustainability, 13*(11), Article 5812. https://doi.org/10.3390/su13115812

Østerlie, O., & Mehus, I. (2020). The Impact of Flipped Learning on Cognitive Knowledge Learning and Intrinsic Motivation in Norwegian Secondary Physical Education. *Education Sciences, 10*(4), 110. https://doi.org/10.3390/educsci10040110

Bae, M.-H. (2023). The Effect of a Virtual Reality-Based Physical Education Program on Physical Fitness among Elementary School Students. *Iranian Journal of Public Health.* https://doi.org/10.18502/ijph.v52i2.11890

Rymarczyk, K., Makowska, I., & Hyniewska, S. (2024). The Impact of the Interactive Floor Device and Aerobic Training on Executive Functions in Children. *Children, 11*(12), 1489. https://doi.org/10.3390/children11121489

Schmid, R.F., Bernard, R.M., Borokhovski, E., Tamim, R.M., Abrami, P.C., Surkes, M.A., Wade, C.A., & Woods, J. (2014). The effects of technology use in postsecondary education: A meta-analysis of classroom applications. *Computers & Education,* 72, 271-291. https://doi.org/10.1016/j.compedu.2013.11.002

Kraft, M.A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241-253. https://doi.org/10.3102/0013189X20912798

Ben Rakaa, O., Lourenço, C., Bassiri, M., & Lotfi, S. (2025b). The Effect of Adapted Physical Activity and Inclusive Sport on the Motivation and Psychological Health of Children with Disabilities: A Randomized Control Trial. *Physical Education Theory and Methodology, 25*(3), 642-651. https://doi.org/10.17309/tmfv.2025.3.21

Ben Rakaa, O., Bassiri, M., & Lotfi, S. (2024). The Influence of School Pathologies on the Feeling of Pedagogical Incompetence in Teaching Inclusive Physical Education. *Physical Education Theory and Methodology, 24*(4), 626-634. https://doi.org/10.17309/tmfv.2024.4.15

Вплив цифрових технологій на результати фізичного виховання: систематичний огляд та метааналіз

Мудеттір Юнес[1ABCDE], Сіхам Ухрір[1ABCDE], Саїд Лотфі[1ABCDE]

[1]Університет Хасана II, Касабланка

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 21 с., 4 табл., 7 рис., 69 джерел.

Обгрунтування. Фізичне виховання стикається зі зростаючим тиском щодо інтеграції цифрових технологій на тлі зниження залученості та мотивації учнів. Попри зростання рівня впровадження, систематизовані докази ефективності цих інструментів у різних доменах результатів залишаються обмеженими, особливо щодо умов, за яких вони є найбільш корисними.

Мета. Це систематичне оглядове дослідження та метааналіз вивчали вплив цифрових технологій на результати фізичного виховання, досліджуючи загальну величину ефекту, доменно-специфічні ефекти (фізичні, мотиваційно-афективні, когнітивні) та модератори гетерогенності.

Матеріали і методи. Систематичний пошук у чотирьох базах даних (Scopus, Web of Science, ERIC, SPORTDiscus) дозволив ідентифікувати експериментальні та квазіекспериментальні дослідження, опубліковані в період 2015–2025 рр. Дванадцять досліджень, що відповідали критеріям включення, забезпечили 64 розміри ефекту на основі вибірки з 1 346 учасників. Для розрахунку зведених розмірів ефекту Hedges' g було проведено метааналізи з випадковими ефектами з використанням оцінювання REML, доповнені прогнозними інтервалами для врахування екстремальної гетерогенності.

Результати. Загальний зведений ефект становив $g = 0,745$ (95% CI [0,590; 0,900], $p < 0,001$), із доменно-специфічними оцінками $g = 0,854$ (фізичні), $g = 0,616$ (мотиваційно-афективні) та $g = 0,936$ (когнітивні). Проте екстремальна гетерогенність у всіх моделях ($I^2 = 99,98\%$, $\tau^2 = 0,373$) робить ці зведені оцінки нестабільними та такими, що не підлягають узагальненню. 95% прогнозний інтервал [−0,471; 1,961] свідчить про те, що істинний ефект у конкретному контексті впровадження може варіювати від помірно негативного до дуже великого позитивного, що відображає високу контекстну залежність спостережуваних ефектів.

Висновки. Цифрові технології демонструють потенціал для підвищення результатів фізичного виховання; однак широкі прогнозні інтервали та екстремальна гетерогенність не дозволяють робити універсальні твердження щодо ефективності. Практичний вплив конкретної реалізації критично залежить від типу технології, педагогічної інтеграції, компетентності вчителя та освітнього контексту. Майбутні дослідження повинні зосереджуватися на визначенні умов, за яких цифрові технології є найбільш та найменш ефективними, а не лише на оцінюванні середніх ефектів.

Ключові слова: фізичне виховання, цифрові технології, метааналіз, гетерогенність, педагогічна інтеграція, освітні технології.

Information about the Authors:

Youness Moudettir: youness.moudettir-etu@etu.univh2c.ma; https://orcid.org/0000-0002-6793-740X; Multidisciplinary Laboratory in Education Sciences and Training Engineering (LMSEIF), Assessment in Sport Sciences and in Physical Activity Didactic, Normal Superior School (ENS), Hassan II University of Casablanca, BP 50069, Ghandi, Morocco.

Ouhrir Siham: siham.ouhrir@gmail.com; https://orcid.org/0000-0001-9870-5655; Multidisciplinary Laboratory in Education Sciences and Training Engineering (LMSEIF), Assessment in Sport Sciences and in Physical Activity Didactic, Normal Superior School (ENS), Hassan II University of Casablanca, BP 50069, Ghandi, Morocco.

Lotfi Said: said.lotfi@etu.univh2c.ma; https://orcid.org/0000-0002-0008-6145; Multidisciplinary Laboratory in Education Sciences and Training Engineering (LMSEIF), Assessment in Sport Sciences and in Physical Activity Didactic, Normal Superior School (ENS), Hassan II University of Casablanca, BP 50069, Ghandi, Morocco.

Cite this article as: Youness, M., Ouhrir, S., & Lotfi, S. (2026). The Effects of Digital Technologies on Physical Education Outcomes: A Systematic Review and Meta-Analysis. *Physical Education Theory and Methodology, 26*(2), 230-250. https://doi.org/10.17309/tmfv.2026.2.02

Received: 09.01.2026. Accepted: 20.02.2026. Published: 30.03.2026