David Suter

Matriculation Number: 20-065-405

MSc: Sport, Exercise & Health; Sports Science

Effectiveness of QPE-Aligned School-Based Physical Activity Interventions in Low- and Middle-Income Countries:

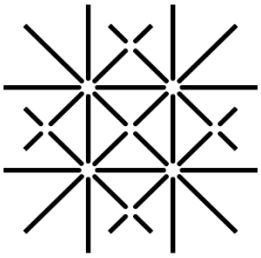
A Systematic Comparative Review of Randomised Controlled Trials

Master's Thesis

Submitted to the Department of Sports, Exercise and Health

of the University of Basel

Supervisor: Dr. Ivan Müller



Universität Basel

Basel, June 2026

Abstract

Background: UNESCO's Quality Physical Education (QPE) framework promotes inclusive, pedagogically robust physical education worldwide, yet it is unclear how effective QPE-aligned school-based interventions are across national contexts. This systematic comparative review examined the effectiveness of QPE-aligned school-based randomised controlled trials (RCTs) in low- and middle-income countries, the role of implementation fidelity, and which intervention components showed the most consistent effects.

Methods: Following PRISMA 2020 guidance, PubMed, SPORTDiscus, and Web of Science were searched for school-based RCTs published from 2015 onwards in low- and middle-income economies. Twelve RCTs comprising 4’999 randomised pupils met all inclusion criteria. Data was extracted at study and outcome level and synthesised narratively; heterogeneity in designs, outcomes, and reporting precluded meta-analysis.

Results: All included studies were conducted in upper-middle-income economies (China n = 8, Brazil n = 3, South Africa n = 1). Six studies were coded as effective and six as partly effective; none was ineffective. Structured fitness-oriented components were most consistently associated with improved physical fitness and movement-related outcomes. Teacher-focused components were linked to behavioural and psychosocial benefits, whereas additional PE time alone showed limited effects. Fidelity reporting was uneven, and no trial assessed post-intervention sustainability.

Conclusions: QPE-aligned school-based RCTs can improve fitness, movement, and selected psychosocial outcomes, with pedagogical quality appearing more decisive than PE dose. Cross-country conclusions remain descriptive. Trials from low- and lower-middle-income countries, systematic fidelity reporting, and follow-up assessments are needed.

Abstract (Deutsch)

Hintergrund: Das Quality-Physical-Education-Rahmenkonzept (QPE) der UNESCO fördert weltweit einen inklusiven, pädagogisch fundierten Sportunterricht. Unklar ist jedoch, wie wirksam QPE-orientierte schulbasierte Interventionen in unterschiedlichen nationalen Kontexten sind. Dieses systematische vergleichende Review untersuchte die Wirksamkeit QPE-orientierter schulbasierter randomisierter kontrollierter Studien (RCTs) in Ländern mit niedrigem und mittlerem Einkommen, die Rolle der Implementierungstreue sowie die Frage, welche Interventionskomponenten die konsistentesten Effekte zeigten.

Methoden: In Anlehnung an PRISMA 2020 wurden PubMed, SPORTDiscus und Web of Science nach schulbasierten RCTs durchsucht, die ab 2015 in Ländern mit niedrigem und mittlerem Einkommen publiziert wurden. Zwölf RCTs mit insgesamt 4'999 randomisierten Schülerinnen und Schülern erfüllten alle Einschlusskriterien. Die Daten wurden auf Studien- und Outcome-Ebene extrahiert und narrativ synthetisiert; die Heterogenität der Designs, Outcomes und Berichterstattung schloss eine Meta-Analyse aus.

Ergebnisse: Alle eingeschlossenen Studien stammten aus Ländern mit oberem mittlerem Einkommen (China n = 8, Brasilien n = 3, Südafrika n = 1). Sechs Studien wurden als wirksam und sechs als teilweise wirksam eingestuft; keine war unwirksam. Strukturierte, fitnessorientierte Komponenten waren am konsistentesten mit Verbesserungen der körperlichen Fitness und motorischer Outcomes verbunden. Lehrpersonenbezogene Komponenten zeigten verhaltensbezogene und psychosoziale Effekte, während zusätzliche Sportunterrichtszeit allein nur begrenzte Wirkungen aufwies. Die Berichterstattung zur Implementierungstreue war uneinheitlich; keine Studie erfasste die Nachhaltigkeit der Effekte.

Schlussfolgerungen: QPE-orientierte schulbasierte RCTs können Fitness, motorische und ausgewählte psychosoziale Outcomes verbessern, wobei die pädagogische Qualität bedeutsamer erscheint als der Umfang des Sportunterrichts. Länderübergreifende Schlussfolgerungen bleiben deskriptiv. Studien aus Ländern mit niedrigem und unterem mittlerem Einkommen, eine systematische Erfassung der Implementierungstreue sowie Follow-up-Erhebungen sind erforderlich.

Contents

Abbreviations

AI	Artificial intelligence
ANCOVA	Analysis of covariance
ASP	After-school programme
BF%	Body fat percentage
BIA	Bioelectrical impedance analysis
BM	Body mass
BMI	Body mass index
bpm	Beats per minute
CG	Control group
CHAMPS	Childhood Health, Activity and Motor Performance Study; used in the thesis for the Chinese CHAMPS school physical activity programme
CI	Confidence interval
CODT	Change-of-direction test
CON	Control or control condition
CRF	Cardiorespiratory fitness
CST	Core strength training
CT	Combined training
d2	d2 Test of Attention
DASH	Disease, Activity and Schoolchildren's Health
DXA	Dual-energy X-ray absorptiometry
F	F statistic
FFM	Fat-free mass
FM	Fat mass
FMS	Fundamental movement skills; Functional Movement Screen when referring to Liao et al.'s movement-quality battery
FST	Functional strength training
FT	Functional training
GMQ	Gross motor quotient
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HE	Health education
HIFT	High-intensity functional training
HIIT	High-intensity interval training
HR	Heart rate
HRmax	Maximum heart rate
HS	High school
IG	Intervention group
JRE	Jump-rope exercise
kg/m²	Kilograms per square metre
LMIC	Low- and middle-income country
MCID	Minimal clinically important difference
MD	Mean difference
mHealth	Mobile health
MHR	Mean heart rate
MICT	Moderate-intensity continuous training
mL/kg/min	Millilitres per kilogram per minute
MPA	Moderate physical activity
MS	Middle school
MSc	Master of Science
MVPA	Moderate-to-vigorous physical activity
n	Number of participants or observations
N/A	Not applicable
NCD	Non-communicable disease

NR	Not reported
ns	Not significant
NU	Nutrition component
p	Probability value
PA	Physical activity
PACER	Progressive Aerobic Cardiovascular Endurance Run
PAQ-A	Physical Activity Questionnaire for Adolescents
PE	Physical education
PE-PA	Physical education-based physical activity
POMS	Profile of Mood States
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PT	Plyometric training
QPE	Quality Physical Education
RCT	Randomised controlled trial
RMB	Renminbi (Chinese currency)
RPE	Rating of perceived exertion
RQ	Research question
RT	Resistance training; reaction time in n-back outcome reporting
RT4T	Resistance Training for Teens
SB	Sedentary behaviour
SD	Standard deviation
SE	Standard error
SES	Socioeconomic status
Sig.	Significance or significant
SLJ	Standing long jump
SPE	School physical education intervention
SRT	Shuttle run test
std β	Standardised beta coefficient
STH	Soil-transmitted helminth
TEO	Theory of Expanded, Extended, and Enhanced Opportunities
TGMD	Test of Gross Motor Development
TGMD-2	Test of Gross Motor Development, Second Edition
TMD	Total mood disturbance
TST	Traditional strength training
UNESCO	United Nations Educational, Scientific and Cultural Organization
VO_2max	Maximal oxygen uptake
VPA	Vigorous physical activity
vs.	Versus
WC	Waist circumference
WCST	Wisconsin Card Sorting Test
WHO	World Health Organization
wk	Week
Z-score	Standardised score
β	Beta coefficient
ηp^2	Partial eta squared

List of Figures

List of Tables

1. Introduction

Physical inactivity has emerged as one of the most significant global public health challenges of the 21st century, particularly among children and adolescents. Worldwide, 81% of adolescents fail to achieve the World Health Organization's (WHO) recommended daily levels of physical activity, contributing to increased risks of obesity, cardiovascular disease, type 2 diabetes, mental health problems, and lower overall well-being (World Health Organization, 2020). Projections indicate that, if current trends persist, physical inactivity will lead to substantial growth in non-communicable diseases (NCDs) and impose increasing economic burdens on health systems by 2030 (World Health Organization, 2022a). Schools have unparalleled reach to children and adolescents and are therefore the most effective means of providing all children with the skills, attitudes and values necessary to engage in lifelong physical activity, says the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2015b).

Recognising the strategic role of schools, UNESCO introduced the Quality Physical Education (QPE) framework in 2015 to promote inclusive, developmentally appropriate and pedagogically robust physical education worldwide. QPE emphasises competence development, meaningful participation, supportive motivational climates, and inclusive teaching practices that enable education regardless of background or ability (UNESCO, 2015b). To assist national systems in strengthening PE provision, UNESCO produced a comprehensive framework of resources, including the QPE Guidelines for Policy-Makers (UNESCO, 2015b), the Methodology for Policy Review (UNESCO, 2015c), the Advocacy Toolkit (UNESCO, 2021a), the QPE Infographic (UNESCO, 2015a), and a multi-country Analysis of Process, Content and Impact (UNESCO, 2021b). Collectively, these documents underscore the importance of high-quality pedagogy, teacher training, and structured PE curricula in promoting physical, social, emotional and cognitive development.

Parallel to these policy efforts, an expanding body of school-based randomised controlled trials (RCTs) has examined interventions that align with QPE principles, including student-centred instructional approaches, autonomy-supportive teaching, motor skill development programmes, and structured curriculum models. These interventions have demonstrated a wide range of effects on outcomes such as physical activity, fitness, motor competence, motivation and psychological well-being. However, evidence across trials remains highly heterogeneous. Studies conducted in different countries frequently report divergent outcomes, suggesting that contextual factors such as curriculum structures, teacher training systems, school resources, cultural norms and socioeconomic environments may substantially shape intervention effectiveness (Love et al., 2019; Suter, 2024).

Despite this growing evidence base, no systematic review has yet synthesised how the effectiveness of QPE-aligned interventions varies across countries, or which contextual factors may be associated with these differences. Because the available evidence is heterogeneous and unevenly distributed across settings, such a synthesis is necessarily descriptive rather than causal. Given UNESCO's global vision for scalable, adaptable and inclusive QPE, synthesising cross-country RCT findings is essential for understanding which intervention components demonstrate robustness across settings, and which require adaptation for local educational systems. A systematic review with a cross-country comparative focus is therefore timely and necessary to advance international knowledge on the transferability and contextual sensitivity of QPE-aligned school-based interventions.

1.1. Research Aim and Research Questions

The objective of this master's thesis is to systematically examine how the effectiveness of QPE aligned interventions delivered through RCTs varies across different national contexts. Although UNESCO's QPE framework provides a global vision for inclusive, equitable and pedagogically sound physical education (UNESCO, 2015), empirical evaluations of school-

based interventions show substantial variation in outcomes across countries (Love et al., 2019). These differences may reflect contextual factors such as curriculum structures, teacher training systems, socioeconomic environments, resource availability, and broader cultural norms surrounding physical activity and education (Ha et al., 2020; Smedegaard et al., 2017).

School-based RCTs carried out since the publication of the UNESCO QPE framework in 2015 provide an important evidence base for understanding how QPE principles such as inclusive pedagogies, competence development, motivational climates, and structured PE curricula translate into measurable outcomes for children and adolescents. However, this evidence remains widely dispersed, and little is known about how national or regional contexts moderate intervention effectiveness, or which QPE elements appear most transferable across countries. Addressing this gap is essential for informing the global scalability of QPE and supporting countries in adapting evidence-based approaches to their QPE adaptation.

Accordingly, this review will synthesise RCTs conducted in diverse countries and analyse their outcomes through a comparative lens, with particular attention to contextual moderators and implementation features. This approach enables the thesis to extend beyond effectiveness alone and contributes to a deeper understanding of the conditions under which interventions succeed, partially succeed, or show limited effects in specific educational environments.

Research Question 1

Do the outcomes of QPE-aligned randomised controlled trials differ across countries with varying educational, socioeconomic, and cultural contexts?

Research Question 2

How does implementation fidelity relate to the effectiveness of QPE-aligned RCTs across different national settings?

Research Question 3

Which intervention components of QPE-aligned RCTs show the most consistent effectiveness across countries?

Together, these research questions (RQ) aim to advance understanding of how QPE-aligned interventions operate within different national contexts and provide evidence to support more context-sensitive, scalable and globally relevant QPE policy-making.

2. Method and Data

2.1. Method

This section describes how the review was designed and conducted. It outlines the analytical framework, search strategy, eligibility criteria, data extraction procedures, and synthesis approach.

2.1.1. Study Design and Analytical Framework

This thesis is a systematic comparative review of school-based RCTs evaluating interventions aligned with QPE principles in low- and middle-income economy contexts. The review investigates how the effects of these interventions vary across national, educational, socioeconomic, and implementation contexts. It focuses on three analytical dimensions: cross-country differences in intervention outcomes, the relationship between implementation fidelity and effectiveness, and the consistency of intervention components across settings.

The review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) procedure (Page et al., 2021) based on predefined eligibility criteria, a structured search strategy, transparent study selection, and standardised data extraction. Eligible studies were primary intervention studies using a randomised controlled design, including individually randomised trials, cluster randomised trials, factorial randomised trials, and school-based interventions with clearly reported random allocation. The restriction to randomised controlled designs strengthens the review's focus on intervention effectiveness.

Studies were classified as QPE-aligned when their interventions reflected core principles of quality physical education, even if the term "QPE" was not used by the authors. Relevant principles included structured physical education, inclusive or context-sensitive delivery, teacher training, motor competence development, physical fitness, physical activity promotion, motivation, and broader physical, psychological, social, cognitive, or developmental outcomes. This operational definition allowed the review to capture interventions that corresponded to the substance of QPE rather than only studies that used the label explicitly.

The research questions were operationalised through three analytical dimensions: context, implementation fidelity, and intervention content. Context captured country, income group, school setting, participant characteristics, comparator condition, and outcome direction. Implementation fidelity captured adherence, attendance, dose delivery, intensity monitoring, attrition, and reported barriers. Intervention content captured the main components of each intervention, including teacher training, additional PE, modified curricula, resistance training, high-intensity interval training (HIIT), motor competence, motivational strategies, and multicomponent approaches.

The synthesis was conducted narratively and comparatively; the rationale for not pooling effects statistically is given in Section 2.1.5. Cross-country comparisons were planned to be interpreted cautiously in the event of geographic imbalance in the evidence base, and country-level differences were not treated as direct evidence of national effects. Findings were compared across studies by outcome domain, intervention component, implementation characteristics, and available contextual information.

2.1.2. Search Strategy and Study Selection

A systematic database search was conducted in PubMed, SPORTDiscus, and Web of Science. The search strategy was built around three core concepts: physical education/QPE, school-based settings, and randomised controlled trial design. This structure followed the consultation with the University of Basel medical research support team, where these three concepts were identified as the most relevant search blocks. The consultation also noted that "QPE" is rarely used explicitly in the scientific literature, meaning that the search could not rely only on the term "Quality Physical Education." Instead, broader physical education terms were used, and QPE alignment was assessed during screening.

The primary search string was developed in PubMed using a combination of MeSH terms, title/abstract terms, publication types, and an RCT filter. The PubMed search string was then translated into database-specific syntax for SPORTDiscus and Web of Science using the Polyglot Search Translator. The use of this tool is supported by Clark et al., (2020) who evaluated the Polyglot Search Translator for translating search strategies across databases. Additionally, a search filter was added to ensure only studies published in 2015 and onwards would be retrieved. The final database-specific search strings were as follows:

PubMed

```
(Physical Education and Training[mh] OR quality physical education[tiab] OR
"QPE"[tiab])
AND
(schools[mh] OR school*[tiab] OR "teacher*"[tiab] OR "student*"[tiab] OR
"pupil*"[tiab])
AND
((("randomized controlled trial"[pt] OR "controlled clinical trial"[pt] OR
"randomized"[tiab] OR "placebo"[tiab]) OR ("clinical trials as topic"
[mesh:noexp]) OR (randomly [tiab] OR trial [ti])) NOT (animals [mh] NOT
humans [mh]))
```

Web of Science

```
("Physical Education" AND Training OR "quality physical education" OR QPE)
AND
(schools OR school* OR teacher* OR student* OR pupil*)
AND
((("randomized controlled trial" OR "controlled clinical trial" OR
randomized OR placebo) OR ("clinical trials as topic") OR (randomly OR
trial)) NOT (animals NOT humans))
```

SPORTDiscus

```
("Physical Education" AND DE "Training" OR (TI "quality physical education"
OR AB "quality physical education") OR (TI "QPE" OR AB "QPE"))
AND
(DE "schools" OR (TI "school*" OR AB "school*") OR (TI "teacher*" OR AB
"teacher*") OR (TI "student*" OR AB "student*") OR (TI "pupil*" OR AB
"pupil*"))
AND
(((TX "randomized controlled trial" OR TX "controlled clinical trial" OR
(TI "randomized" OR AB "randomized") OR (TI "placebo" OR AB "placebo")) OR
(DE "clinical trials as topic") OR ((TI "randomly" OR AB "randomly") OR TI
"trial")) NOT (DE "animals" NOT DE "humans"))
```

All search results were exported into the literature review management software Covidence (Veritas Health Innovation, 2026) for deduplication and screening. The selection process followed the PRISMA flow structure. First, all records identified through the three databases were imported and duplicates or ineligible automated records were removed. Second, titles and abstracts were screened against the predefined eligibility criteria. Third, full texts were retrieved and assessed for eligibility. Reasons for exclusion at the full-text stage were documented. The complete study selection process is presented in Figure 1. In total, 697 records were identified through database searching. After the removal of 115 records, 582 records were screened by title and abstract. Of these, 158 full texts were assessed for eligibility. Following full-text screening, twelve studies met all inclusion criteria and were included in the final review.

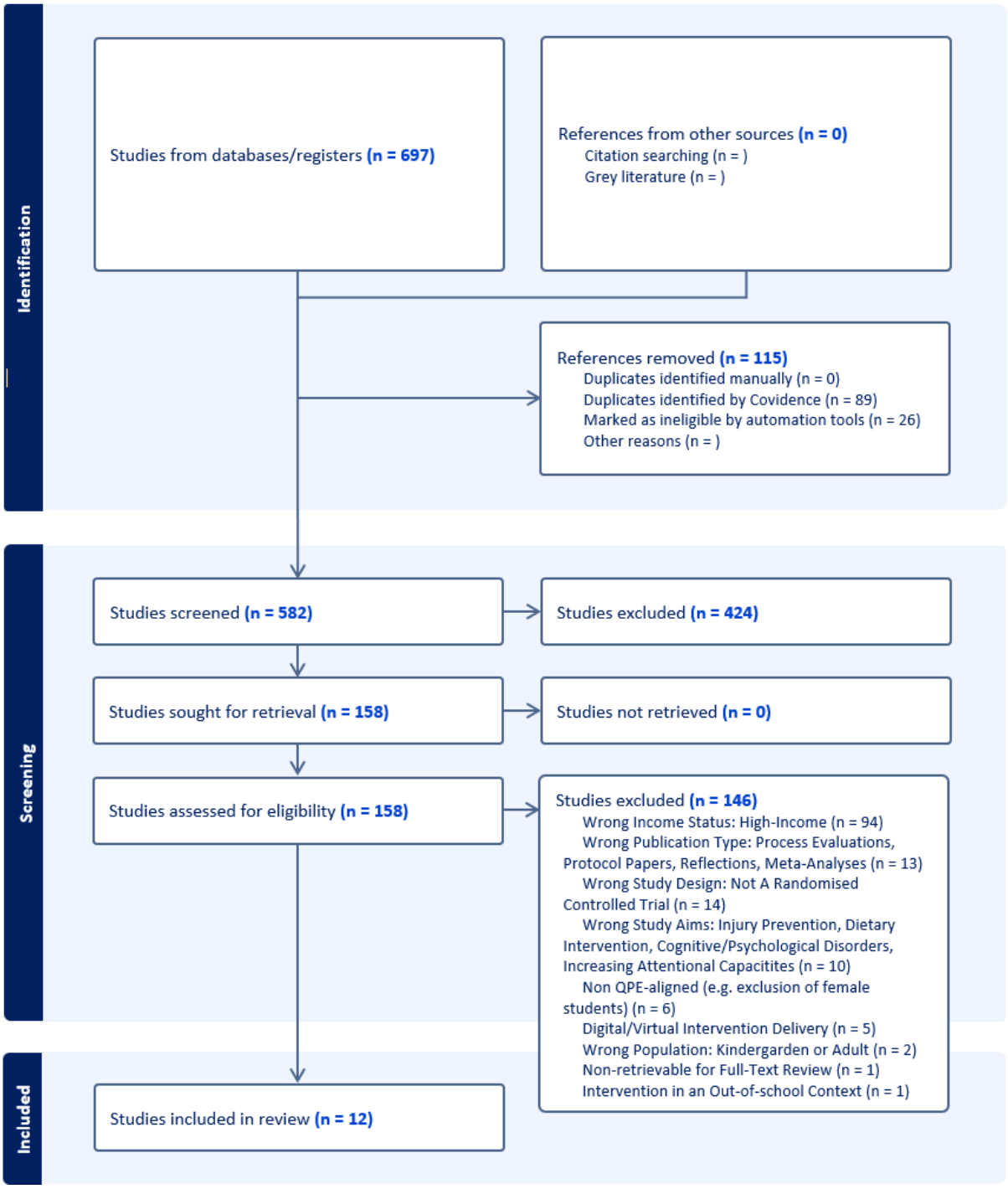


Figure 1
PRISMA Flow Diagram of Study Selection

2.1.3. Eligibility Criteria

Studies were selected according to predefined eligibility criteria. These criteria were applied during title and abstract screening and full-text assessment. They reflected the review's focus on school-based, QPE-aligned randomised controlled interventions conducted in low- and middle-income economy contexts. Studies published from 2015 onward were considered eligible, corresponding to the publication year of UNESCO's QPE guidelines. The publication year restriction was applied through database filters during the search process. The inclusion and exclusion criteria are listed in Table 1 and Table 2 respectively.

Table 1
Inclusion criteria for screening

Criterion	Inclusion rule
Study design	Randomised controlled trials, cluster randomised controlled trials, factorial randomised controlled trials, or clearly randomised controlled school-based interventions
Setting	School-based intervention
Population	Children or adolescents
School stage	Primary, secondary, or high school
Country income-group	Low-, lower-middle-, or upper-middle-income economy according to the current World Bank classification
Intervention type	QPE-aligned physical education, school physical activity, physical fitness, motor competence, or health-related school intervention
Delivery	Intervention delivered in person during PE, regular school time, or school-linked activity
Comparator	Control group, usual PE, standard PE, no intervention, or alternative intervention arm
Outcomes	At least one relevant physical, behavioural, psychological, cognitive, motor, or school-related outcome
Publication type	Full-text primary research article
Publication year	2015 onward
Language	No language restriction

Note. Criteria were applied during title and abstract screening and full-text assessment. Country income group follows the current World Bank classification. PE = physical education; QPE = Quality Physical Education; RCT = randomised controlled trial.

Table 2
Exclusion criteria for screening

Criterion	Exclusion rule
High-income economy	Studies conducted in high-income economies were excluded
Wrong design	Non-randomised studies, quasi-experimental studies, protocols, reviews, meta-analyses, process evaluations, reflections, editorials, and commentaries were excluded
Wrong population	Kindergarten and preschool populations were excluded
Wrong setting	Interventions delivered primarily outside the school context were excluded
Non-QPE alignment	Studies were excluded when the intervention did not plausibly align with QPE-relevant PE or school physical activity principles
Male-only eligibility	Studies explicitly restricted to boys only were excluded
Digital delivery	Studies were excluded when the intervention itself was delivered digitally or virtually
Dietary or clinical focus	Studies were excluded when the main intervention aim was dietary treatment, injury prevention, clinical treatment, or another non-PE-aligned aim
Full-text access	Studies were excluded when the full text could not be retrieved

Note. Criteria were applied during title and abstract screening and full-text assessment. Reasons for exclusion at the full-text stage are reported in the PRISMA flow diagram (Figure 1). PE = physical education; QPE = Quality Physical Education.

QPE alignment was assessed substantively rather than by explicit terminology. Studies did not need to use the phrase “Quality Physical Education” to be eligible. Instead, interventions

were considered QPE-aligned when they reflected principles such as structured PE, physical activity promotion, teacher-led delivery, motor competence development, fitness development, motivational support, inclusion, or broader developmental outcomes.

Female-only studies were retained when they were otherwise school-based and QPE-aligned. This decision was justified by the gendered pattern of youth physical activity: adolescent girls are less likely than boys to meet physical activity recommendations (World Health Organization, 2022b), participation declines during adolescence, and intervention effects for girls are typically small, making targeted female samples relevant rather than exclusionary (Pearson et al., 2015).

Digital components were not automatic grounds for exclusion. Studies were retained when a digital element formed only one component of an otherwise in-person, school-based intervention. Studies were excluded only when the intervention itself was delivered digitally or virtually, as this fell outside the review's focus on in-person QPE-aligned school practice.

Country income status was classified using the current World Bank country classification (World Bank Group, 2026). Although all included studies were conducted in upper-middle-income economies, this was not an inclusion requirement.

2.1.4. Data Extraction and Coding

Data was extracted using a self-developed extraction form in Covidence and exported into a structured spreadsheet (Figure 2).

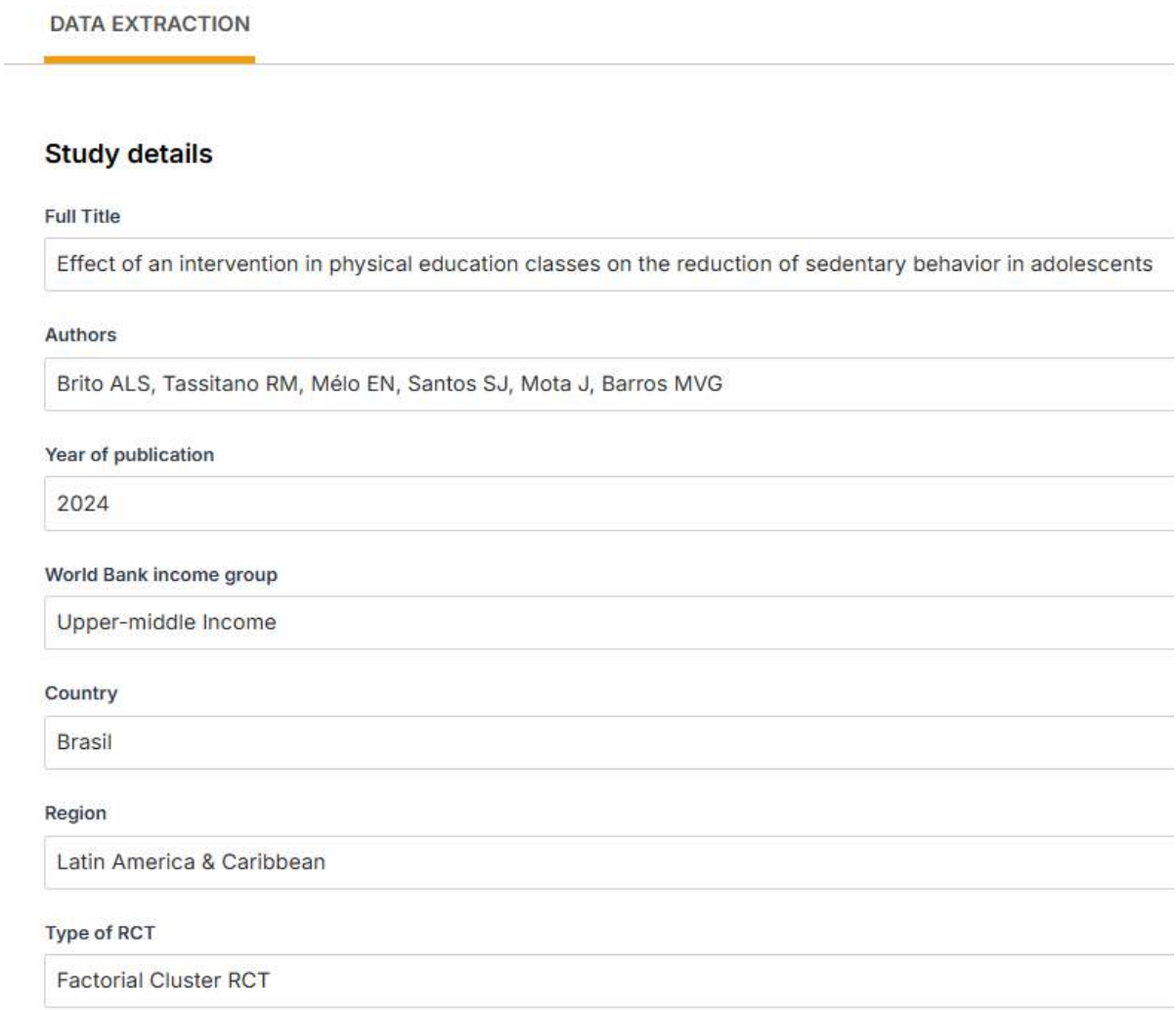


Figure 2
Extraction process sample using a prepared Covidence extraction template

The main extraction sheet contained one row per included study and covered the following categories displayed in Table 3:

Table 3
Main extraction categories

Extraction category	Corresponding spreadsheet columns
Study identification	Authors; Title
Country and income classification	Income Group; Country; Region
Study design	Type of RCT; Level of Randomisation
Educational context	Year Level; Educational Level; School Type; Setting; Socioeconomic Description; Low Resource Setting Specified
Population characteristics	Total Sample Randomised; Randomised n by Group; Mean Age (+/- SD); Sex Distribution; Baseline Weight Status; Ethnicity/Race; Baseline Comparability
Study aim	Aim of study
Intervention structure and components	Component Structure (single/multi); Component Composition; Intervention Component Description
Dosage and resources	Intervention Dosage; Intervention Dosage Specifications; Required Equipment
Comparator	Control Condition
Implementation fidelity and participant flow	Implementation Fidelity & Process Indicators; Dropout in %
Reviewer interpretation	Reviewer Inferred Feasibility/Scalability Notes
Author-stated barriers and limitations	Reported Barriers; Study limitations
Outcomes and subgroup findings	Relevant Outcomes; Sex/Subgroup Reporting & Findings
Effectiveness appraisal	Overall Effectiveness Judgement; Main Positive Outcome Domains; Key findings
Future research implications	Implications for future research

Outcome data was stored in a separate outcome-level table to allow multiple comparisons, subgroup findings, and outcome domains to be retained without overloading the study-level extraction sheet. Table 4 provides an overview of the outcome-level extraction categories.

Table 4
Outcome data columns

Extraction category	Corresponding outcome spreadsheet columns
Study and comparison	Study; Comparison; Subgroup
Outcome classification	Domain
Effect direction and significance	Direction; Significant

Extraction category	Corresponding outcome spreadsheet columns
Effect estimates	Effect; 95% CI / statistic; p-value
Arm-level component coding	Component composition
Notes	Notes

Note. CI = confidence interval. The "95% CI / statistic" column records a confidence interval where reported, otherwise the relevant test statistic. Where no between-group magnitude was reported, the effect estimate was coded as not reported.

Full-text articles were the primary source for extraction. Information was taken from the main text, tables, figures, and study flow diagrams. Where abstracts and full texts differed, full-text information was prioritised. Non-English studies were translated using the DeepL Translator (DeepL SE, n.d.). Extraction was conducted by a single reviewer; therefore, inter-rater reliability was not calculated. Unclear or internally inconsistent information was revisited in the full text and coded as unclear or not reported if it could not be verified. The study-level variables included the following:

- Author and Year
- Country
- World Bank Income Group
- School Level
- School Type
- Setting
- Study Design
- Randomisation Unit
- Number of Study Arms
- Participant Characteristics
- Intervention Content
- Comparator Condition
- Intervention Duration
- Frequency, Session Length
- Total Planned Exposure where Calculable

Country, region, and income group were coded separately. School type, setting, and socioeconomic context were coded only when they were explicitly reported or could be verified. Intervention components were coded using controlled categories while preserving free-text descriptions of the intervention. Component categories included:

- teacher training
- additional PE lessons
- modified PE curriculum
- motor competence or fundamental movement skill training
- strength or resistance training
- aerobic training
- high-intensity interval training

- motivational strategies
- school-environment changes
- and digital or mobile Health (mHealth) components

Multi-arm and factorial studies were coded by preserving the individual intervention arms rather than collapsing them into a single intervention category. Outcome data was extracted into a separate table at the level of study, comparison, subgroup, and outcome domain. Outcome domains included:

- cardiorespiratory fitness
- muscular strength or power
- speed or agility
- flexibility
- body composition
- fundamental movement skills
- physical activity
- sedentary behaviour
- cognition
- mental health
- wellbeing

Metabolic or clinical biomarkers were not included in the RQ-focused outcome synthesis unless they directly informed the physical, behavioural, psychological, cognitive, or motor outcome domains.

Both statistically significant and non-significant findings were extracted. Effect direction was standardised as favouring the intervention, showing no difference, or favouring the comparator, with the beneficial direction defined separately for each outcome. Implementation fidelity was extracted because of its relevance to the second research question. Fidelity indicators included:

- attendance
- adherence or compliance
- teacher delivery
- intensity monitoring
- dose delivered
- scalability
- sustainability
- and reported implementation barriers

Participant flow was coded separately, distinguishing overall attrition, arm-specific attrition, and outcome-specific missing data where reported. Device-validity loss, such as missing accelerometer data, was kept separate from participant dropout.

Author-stated information was separated from reviewer interpretation. Reported barriers and limitations were restricted to statements made by the original study authors. Reviewer observations on feasibility, scalability, methodological concerns, or interpretation were recorded separately. Practical problems affecting intervention delivery were coded as barriers,

whereas weaknesses in design, measurement, sample size, follow-up, blinding, or reporting were coded as limitations. Missing and ambiguous information was coded conservatively. Information that was not reported was coded as NR, information that could not be interpreted confidently was coded as Unclear, and variables that did not apply to a study were coded as N/A.

2.1.5. Data Synthesis

The extracted data was synthesised using a structured narrative comparative synthesis (Popay et al., 2006). A statistical meta-analysis was not conducted because the included studies differed substantially in intervention content, population characteristics, comparator conditions, randomisation unit, outcome domains, measurement instruments, intervention duration, and statistical reporting. Several studies also reported results in incompatible formats, including raw mean differences, standardised coefficients, standardised effect sizes, p-values without between-group magnitudes, and within-group changes only. Under these conditions, statistical pooling would not have produced a meaningful summary estimate. The synthesis was informed by principles of narrative synthesis and synthesis without meta-analysis (Campbell et al., 2020), with findings grouped, compared, and interpreted transparently rather than pooled statistically.

The synthesis framework is summarised in Table 5. It links the main synthesis dimensions (outcome domains, intervention components, implementation fidelity, and country/context) to the extracted data and the corresponding synthesis procedure. Outcome findings were grouped into predefined domains, including cardiorespiratory fitness, muscular strength or power, speed or agility, flexibility, body composition, fundamental movement skills, physical activity, sedentary behaviour, cognition, mental health, wellbeing, and metabolic or clinical outcomes where relevant. Domain-level grouping was used because studies often reported several tests within the same broader construct. Individual tests and reported estimates were retained in the outcome table, while the synthesis focused on broader patterns across domains.

For each outcome record, the direction of effect was standardised as favours intervention, no difference, favours comparator, mixed, or unclear. Statistical significance was coded primarily on the basis of between-group evidence where available, consistent with the outcome table structure. The beneficial direction was defined separately for each outcome. Increases in physical activity, fitness, motor competence, cognitive performance, or motivation were coded as beneficial, whereas reductions in sedentary behaviour, body fat, BMI, waist circumference, depressive symptoms, or metabolic risk markers were also coded as beneficial. Both statistically significant and non-significant findings were retained to avoid selective emphasis on positive results.

Table 5
Synthesis framework

Analytical dimension	Data used	Synthesis procedure	Output
Outcome domains	Outcome table: domain, direction, significance, effect estimate	Group findings by outcome domain and compare direction/significance	Domain-level pattern of effects
Intervention components	Component coding from extraction sheet	Compare component categories with domains showing positive/null/mixed effects	Descriptive component–outcome associations
Implementation fidelity	Implementation fidelity/process indicators, dropout percentage,	Compare fidelity reporting with study-level effectiveness judgement	Exploratory fidelity–

Analytical dimension	Data used	Synthesis procedure	Output
	feasibility/scalability notes, reported barriers		effectiveness patterns
Country/context	Country, school level, setting, income group	Compare patterns across countries and contextual descriptors	Descriptive cross-context interpretation

Note. The framework links each synthesis dimension to the extracted data and the corresponding synthesis procedure. The synthesis was narrative; no pooled effect estimates were calculated.

Effect estimates were reported descriptively where available but were not transformed into a common metric or pooled. Within-group changes were kept separate from between-group effects. When studies reported only direction and significance without a between-group magnitude, the effect magnitude was coded as not reported. Where the classification of an outcome depended on the statistical test used, such as a time-by-group interaction, group main effect, or post-intervention pairwise comparison, this dependency was flagged.

Multi-arm and factorial trials were analysed by preserving separate intervention-versus-comparator contrasts. Intervention arms were not collapsed when they represented different intervention components. This allowed component patterns to be examined across studies, especially for teacher training, additional PE lessons, modified PE curricula, resistance or strength training, HIIT, motor competence development, motivational strategies, and multicomponent interventions. These component–outcome associations were interpreted descriptively, because most interventions combined several elements and therefore did not allow the independent causal contribution of single components to be isolated.

Implementation fidelity was synthesised by comparing study-level effectiveness judgements with information recorded under implementation fidelity and process indicators, dropout percentage, reviewer-inferred feasibility/scalability notes, and reported barriers. Where studies reported specific process indicators such as attendance, adherence, teacher delivery, intensity monitoring, or intervention completion, these were retained in the notes and used descriptively. Missing fidelity information was treated as not reported, not as evidence of poor implementation. Because fidelity reporting was uneven across studies, the relationship between fidelity and effectiveness was interpreted exploratorily.

Country context was used to support interpretation of the findings. Because cross-country comparison depended on the geographic composition of the included studies (Section 2.2.1), apparent differences between countries were interpreted as descriptive patterns in the available evidence rather than as evidence of national or income-level effects. Apparent differences between countries were therefore interpreted as descriptive patterns in the available evidence rather than as evidence of national or income-level effects.

The overall effectiveness judgement assigned to each study was used as a descriptive synthesis aid. Studies were coded as effective when primary or main reported outcomes significantly favoured the intervention, partly effective when benefits were limited to some outcomes, domains, arms, or subgroups, and not effective when no significant main benefit was reported. This judgement was not a formal certainty-of-evidence rating. Because no formal risk-of-bias assessment was conducted, studies were not weighted according to methodological quality. Reporting limitations, such as missing between-group estimates, reliance on within-group analyses, incomplete attrition reporting, or unclear fidelity information, were noted descriptively where they affected interpretation.

2.2. Data

This section describes the final evidence base included in the review. It summarises the included studies, study designs, participant characteristics, intervention features, outcome domains, and available fidelity information.

The final dataset presented in Table 6 consisted of twelve included studies published between 2019 and 2026. The studies were conducted in three countries: China (n = 8), Brazil (n = 3), and South Africa (n = 1). All included studies were conducted in countries classified as upper-middle-income economies according to the current World Bank country classification. This was not an inclusion requirement; studies from low- and lower-middle-income economies were eligible but no such studies met all inclusion criteria. The total number of participants randomised across all included studies was 4’999, with individual study samples ranging from 79 to 1’279 participants. The included studies covered primary, lower secondary, upper secondary, and mixed or unspecified secondary school settings.

Table 6
Characteristics of included studies

Study	Country	Region	Income group	Type of RCT	Randomisation level	Educational level	Sample randomised
Brito et al., 2024	Brazil	Latin America & Caribbean	Upper-middle income	Four-arm factorial cluster RCT	School	Upper secondary	854
Cai et al., 2025	China	Asia	Upper-middle income	Two-arm parallel RCT	Individual	Secondary, unspecified	222
Castilho Dos Santos et al., 2024	Brazil	Latin America & Caribbean	Upper-middle income	Two-arm cluster RCT	School	Lower secondary	317
Liao et al., 2022	China	Asia	Upper-middle income	Two-arm parallel RCT	Individual	Mixed secondary	266
Lima et al., 2022	Brazil	Latin America & Caribbean	Upper-middle income	Four-arm cluster RCT	School	Upper secondary	1’279
Liu et al., 2026	China	Asia	Upper-middle income	Three-arm cluster RCT	Class	Lower secondary	135
Müller et al., 2019	South Africa	Africa	Upper-middle income	Cluster RCT	School	Primary	746
Shao & Cao, 2025	China	Asia	Upper-middle income	Three-arm parallel RCT	Individual	Unclear	120
Wan et al., 2025	China	Asia	Upper-middle income	Four-arm parallel RCT	Individual	Primary	79
Zhang et al., 2024	China	Asia	Upper-middle income	Two-arm parallel RCT	Individual	Primary	100

Study	Country	Region	Income group	Type of RCT	Randomisation level	Educational level	Sample randomised
Zhou et al., 2019	China	Asia	Upper-middle income	Four-arm factorial cluster RCT	School	Lower secondary	753
Zou et al., 2025	China	Asia	Upper-middle income	Two-arm parallel RCT	Individual	Primary	128

Note. n = number of participants randomised. RCT = randomised controlled trial. Level of randomisation refers to the unit of allocation (individual, class, or school). All included studies were conducted in upper-middle-income economies.

2.2.1. Study Designs and Participants

The included studies varied in design. Five studies randomised at the school level, one at the class level, and six at the individual level. The sample included two-arm, three-arm, four-arm, factorial, and cluster randomised trials. This variation was important for synthesis because several studies compared more than one intervention arm against a control condition, while others tested a single intervention against usual PE or no additional intervention. Four studies were conducted in primary school settings, three in lower secondary settings, two in upper secondary settings, one in mixed secondary settings, one in an unspecified secondary setting, and one had an unclear educational-level classification (Table 6). Participant characteristics also varied across studies, including general school populations, adolescents with overweight or obesity, female-only samples, and specific ethnic or regional populations. Female-only studies were retained when otherwise eligible, while male-only studies were excluded according to the eligibility criteria.

2.2.2. Interventions, Comparators, and Dosage

The included studies varied substantially in trial structure, intervention content, comparator condition, and dosage. Seven studies used multi-arm or factorial designs, whereas five tested a single intervention condition against a comparator. This distinction refers to the number of intervention conditions tested within a trial and should not be confused with whether an intervention was single- or multicomponent in content. For example, some single-arm studies delivered multicomponent interventions, while some multi-arm studies compared several relatively focused intervention conditions. Intervention components included teacher training, additional PE lessons, modified PE curricula, aerobic exercise, resistance or strength training, high-intensity training, motor competence or fundamental movement skill training, motivational strategies, school-environment changes, and digital or mHealth support. Most interventions were delivered during regular school time or within PE lessons. Some studies supplemented existing PE provision with additional lessons or structured activity, whereas others replaced part of the usual PE lesson with a specific training block. Comparator conditions included usual PE, regular PE, standard PE, no intervention, no-training control, usual school curriculum, or factorial control conditions. Intervention duration ranged from 8 weeks to 8 months. Planned dosage varied substantially by intervention frequency, session length, and whether the intervention replaced or supplemented existing PE. Several studies reported total planned exposure, while others reported dosage in terms of weekly lessons, intervention arms, teacher-training hours, or broader school-based activity schedules.

Table 7
Summary of intervention characteristics

Characteristic	Summary
Component structure	8 multicomponent interventions; 4 single-component interventions
Main intervention components	Teacher training, additional PE lessons, modified PE curricula, aerobic training, strength/resistance training, HIIT, motor competence/FMS, motivational strategies, school-environment changes, digital/mHealth support
Trial-arm structure	7 multi-arm or factorial designs; 5 single-intervention-condition designs
Intervention component composition	Interventions ranged from focused training programmes to multicomponent school-based interventions combining PE, teacher training, curriculum modification, activity support, strength/resistance training, aerobic/high-intensity training, motor competence/FMS, motivational strategies, school-environment changes, or digital/mHealth support
Comparator conditions	Usual PE, regular PE, standard PE, no intervention, no-training control, usual curriculum, factorial control conditions
Intervention duration	8 weeks to 8 months
Delivery context	Mostly delivered during PE lessons or regular school time; some interventions supplemented existing PE, others replaced part of PE
Resource requirements	Ranged from low-resource bodyweight or teacher-led activities to programmes requiring specialised equipment, monitoring tools, or external support

The interventions also differed in resource requirements. Several programmes were low-resource and relied mainly on bodyweight exercises, existing PE space, simple equipment, or teacher-led delivery. Others required more specialised resources, such as heart-rate monitoring, structured activity cards, external trainers, small-group supervision, Swiss balls, or additional training materials. These differences were retained in the extraction sheet because they are relevant to feasibility and scalability, particularly in low-resource school settings.

2.2.3. Outcome Domains and Fidelity Reporting

Outcome data was stored in a separate outcome-level table with the columns Study, Comparison, Subgroup, Domain, Direction, Significant, Effect (with unit), 95% CI/statistic, p-value, and Notes. The table contained 56 records, of which 50 represented effectiveness outcomes and six represented mediator records. Mediator records were retained to document proposed mechanisms but were not counted as effectiveness outcomes. The effectiveness outcome records covered the following domains:

- body composition
- cardiorespiratory fitness/physical fitness
- cognition/working memory
- executive function

- fundamental movement skills
- movement quality
- mental health/wellbeing
- muscular strength/power
- physical activity
- sedentary behaviour
- speed/agility
- flexibility
- static balance

Outcome direction and significance were coded separately so that both significant and non-significant findings could be retained in the synthesis. Implementation fidelity and participant flow were extracted in the main study-level table. Relevant fields included implementation fidelity and process indicators, dropout percentage, reviewer-inferred feasibility/scalability notes, and reported barriers. Fidelity reporting was uneven across studies. Some studies reported indicators such as attendance, adherence, teacher delivery, intensity monitoring, or implementation barriers, whereas others provided limited process information. Missing fidelity information was coded as not reported or unclear and was not interpreted as evidence of poor implementation.

3. Results

3.1. Overall Effectiveness and Outcome-Domain Patterns

This section summarises the overall effectiveness patterns across the twelve included studies, drawing on the descriptive effectiveness coding defined in the methods. Six studies were coded as effective and six as partly effective. No study was coded as entirely ineffective, although several showed effects only in specific arms, subgroups, or outcome domains. Table 8 summarises the study-level effectiveness pattern, and the full per-study extraction is provided in the Appendix.

Table 8
Overall effectiveness pattern by included studies

Study	Overall judgement	Main positive outcome domains	Key limitations or mixed findings
Brito et al. (2024)	Partly effective	Sedentary behaviour in the teacher-training arm, for boys and girls	Extra PE alone and the combined arm did not significantly reduce sedentary behaviour
Cai et al. (2025)	Effective	Cardiorespiratory fitness, working memory, body composition in girls	Boys' body-composition effects were not clearly supported by between-group evidence
Castilho dos Santos et al. (2024)	Partly effective	Lower-body muscular strength/power	Physical activity, cardiorespiratory fitness, and upper-body strength outcomes were not significant
Liao et al. (2022)	Effective	Movement quality, muscular strength/power, flexibility	Sprint and endurance measures were not significant
Lima et al. (2022)	Partly effective	Mental health/wellbeing in the teacher-workshop arm; executive function in the combined arm	Doubling PE alone did not significantly reduce depressive symptoms
Liu et al. (2026)	Effective	Cardiorespiratory fitness, muscular strength/power, mental health/wellbeing, body composition in HIFT	Effects are mainly based on group-by-time interactions; HIFT and MICT showed different psychological patterns
Müller et al. (2019)	Partly effective	Body composition	Cardiorespiratory fitness did not significantly improve
Shao & Cao (2025)	Effective	Cardiorespiratory fitness, body composition	Control VO_2max also improved, but less than in the jump-rope arms
Wan et al. (2025)	Partly effective	Muscular strength/power, especially push-ups across resistance-training arms	Speed/agility, cardiorespiratory fitness, and body composition were not clearly significant between groups
Zhang et al. (2024)	Partly effective	Fundamental movement skills, especially gross motor quotient and object control	Evidence was test-dependent
Zhou et al. (2019)	Effective	Physical/cardiorespiratory fitness and body composition (body fat %) across CHAMPS contrasts	Effects were reported through planned contrasts

Study	Overall judgement	Main positive outcome domains	Key limitations or mixed findings
Zou et al. (2025)	Effective	Cardiorespiratory/endurance, muscular strength/power, speed/agility, fundamental movement skills, static balance	Flexibility and dynamic balance were not significant

Note. PE = physical education; HIFT = high-intensity functional training; MICT = moderate-intensity continuous training; VO_2max = maximal oxygen uptake; CHAMPS = the school physical activity programme evaluated by Zhou et al. "Effective" and "partly effective" are descriptive synthesis judgements, not formal risk-of-bias or certainty-of-evidence ratings.

Across outcome domains, the clearest positive pattern was observed for physical fitness. Cardiorespiratory fitness or endurance improved in Cai et al., Liu et al., Shao and Cao, Zhou et al., and Zou et al., although not in Castilho dos Santos et al., Müller et al., or Wan et al. Muscular strength and power were also frequently improved, particularly in Castilho dos Santos et al., Liao et al., Liu et al., Wan et al., and Zou et al. Fitness-related outcomes showed the most consistent positive pattern.

Body composition showed a more mixed pattern. Positive effects were reported in Müller et al., Shao and Cao, Liu et al., Cai et al. among girls and Zhou et al. In contrast, Cai et al. did not clearly support body-composition effects among boys, and Wan et al. did not report meaningful between-group body-composition effects.

Movement-related outcomes were generally positive across the included studies, although the underlying constructs differed. Liao et al. reported improvements in movement quality measured with the Functional Movement Screen, whereas Zhang et al. and Zou et al. reported improvements in fundamental movement skills or motor-competence-related outcomes.

Behavioural outcomes were less consistent. Castilho dos Santos et al. did not report significant physical activity improvements. Brito et al. found sedentary-behaviour reductions only in the teacher-training arm, not in the extra-PE or combined arms. Mental health and cognition showed selective positive findings: Lima et al. reported improvements in depressive symptoms, social isolation, sleep quality, and executive function in specific arms; Liu et al. reported mood-related improvements; and Cai et al. reported working-memory improvements.

Table 9
Main outcome-domain pattern across included studies

Outcome domain	Studies reporting positive effects	Studies reporting null or mixed effects	Overall descriptive pattern
Cardiorespiratory fitness / endurance	Cai et al.; Liu et al.; Shao and Cao; Zhou et al.; Zou et al.	Castilho dos Santos et al.; Müller et al.; Wan et al.	Generally positive, especially in structured aerobic, high-intensity, jump-rope, and school-activity interventions
Muscular strength/power	Castilho dos Santos et al.; Liao et al.; Liu et al.; Wan et al.; Zou et al.	Some measures within Castilho dos Santos et al. and Wan et al. were null or only within-group positive	One of the clearest positive domains, but effects varied by test and intervention arm
Body composition	Cai et al. among girls; Liu et al.; Müller et al.; Shao and Cao; Zhou et al.	Cai et al. among boys; Wan et al.; Zhou et al. For BMI	Mixed; stronger in fitness-, adiposity-, or higher-intensity-focused interventions
Movement quality / motor competence	Liao et al.; Zhang et al.; Zou et al.	Some Zhang outcomes were test-dependent; Zou	Positive overall, but constructs differ and should be kept separate

Outcome domain	Studies reporting positive effects	Studies reporting null or mixed effects	Overall descriptive pattern
		dynamic balance and flexibility were not significant	
Physical activity	None clearly significant	Castilho dos Santos et al.; Müller et al. (self-reported)	Limited evidence of improvement
Sedentary behaviour	Brito et al., teacher-training arm only	Brito et al., extra-PE and combined arms	Selective positive effect; not consistently improved by additional PE time
Mental health/wellbeing	Lima et al., teacher-workshop arm; Liu et al.	Lima et al., extra-PE and combined arms for depressive symptoms	Selective positive effects
Cognition	Cai et al.; Lima et al.	Attention outcomes in Lima et al. were not significant	Promising but based on few studies

Note. The table summarises descriptive patterns, not pooled effects. "Positive effects" refers to outcomes coded as favouring the intervention on between-group evidence or specified interaction-based evidence. CRF = cardiorespiratory fitness; TGMD = Test of Gross Motor Development; Functional Movement Screen findings are kept separate from TGMD-style fundamental movement skills.

Overall, QPE-aligned school-based RCTs most consistently improved physical-fitness outcomes, especially cardiorespiratory fitness and muscular strength/power. Effects on body composition, motor competence, mental health, and cognition were more dependent on intervention type, subgroup, and statistical reporting. Behavioural outcomes such as physical activity and sedentary behaviour were least consistent. These patterns provide the basis for the following sections on country/context, implementation fidelity, and intervention components.

3.2. Cross-Country and Contextual Patterns

The included studies were conducted in China, Brazil, and South Africa. As described in Section 2.2.1, the evidence base was geographically imbalanced and limited to upper-middle-income economies; the synthesis therefore allowed descriptive comparison by country and school context, but not comparison across income groups.

Table 10
Cross-country and contextual pattern of included studies

Country	Nr. of Studies	Included studies	Main intervention types	Main positive outcome domains	Main interpretive limitation
China	8	Cai et al. (2025); Liao et al. (2022); Liu et al. (2026); Shao & Cao (2025); Wan et al. (2025); Zhang et al. (2024); Zhou et al. (2019); Zou et al. (2025)	Functional training; functional strength training; strength/resistance training; jump-rope exercise; HIFT; MICT; core strength training; school-based PA programme	Cardiorespiratory fitness; muscular strength/power; body composition; movement quality; fundamental movement skills; cognition; mental health/wellbeing	Findings were concentrated in one country but covered heterogeneous interventions and outcomes
Brazil	3	Brito et al. (2024); Castilho dos Santos et al. (2024); Lima et al. (2022)	Teacher training/workshops; additional PE lessons; multicomponent	Sedentary behaviour; mental health/wellbeing; executive function;	Effects were selective and often limited to specific arms,

Country	Nr. of Studies	Included studies	Main intervention types	Main positive outcome domains	Main interpretive limitation
			school-based intervention; mHealth support	lower-body muscular strength/power	subgroups, or outcomes
South Africa	1	Müller et al. (2019)	Multicomponent school-based PA and health intervention	Body composition	Single-study evidence did not allow broader country-level conclusions

Note. HIFT = high-intensity functional training; MICT = moderate-intensity continuous training; PA = physical activity; PE = physical education. Country-level patterns are descriptive and do not indicate causal national effects. All included countries were classified as upper-middle-income economies.

The China-based studies contributed most of the evidence and covered the widest range of intervention types. These included functional training, functional strength training, resistance training, jump-rope exercise, HIFT, MICT, core strength training, and a broader school physical activity programme. Positive findings in the Chinese studies were most visible for cardiorespiratory fitness, muscular strength/power, body composition, movement-related outcomes, cognition, and mood-related outcomes. However, these findings did not form a uniform country-level pattern because the Chinese studies differed substantially in school level, intervention content, target population, and outcome measurement.

The Brazil-based studies showed a more selective pattern. Brito et al. and Lima et al. reported their clearest effects in teacher-focused intervention arms (section 3.1.). Castilho dos Santos et al. did not report significant physical activity or cardiorespiratory fitness effects, but did show a positive lower-body muscular-fitness effect. Overall, the Brazilian studies showed the clearest positive findings where interventions included teacher-focused components or pedagogical support rather than simply increasing PE time.

The South African evidence was limited to Müller et al. This multicomponent intervention reported favourable effects on body composition but did not significantly improve cardiorespiratory fitness. Because only one South African study was included, the findings can only be treated as a single-study contribution to the overall synthesis.

Educational context also varied within countries. The Chinese studies included primary, lower-secondary, mixed-secondary, and unclear school-level samples, while the Brazilian studies were conducted mainly in lower- or upper-secondary settings and the South African study was conducted in primary schools. Socioeconomic information was unevenly reported, with some studies specifying public, disadvantaged, or lower-resource school contexts more clearly than others. Therefore, educational and socioeconomic context could be described but not isolated as an explanatory factor. Cultural context was not directly measured in the included studies and could therefore only be considered indirectly through country, school setting, and intervention context.

Overall, the cross-country pattern was shaped more by the available study designs and intervention types than by country alone. China contributed most of the physical-fitness and exercise-training evidence, Brazil contributed the clearest teacher-focused findings, and South Africa contributed one multicomponent body-composition finding. Given this composition of the evidence base, RQ1 could only be answered descriptively. Implementation Fidelity and Effectiveness

Implementation fidelity and process reporting varied across the included studies. As shown in Table 11, some studies provided detailed information on delivery agents, attendance, intensity monitoring, adherence, or process indicators, whereas others provided more limited implementation information. The relationship between fidelity and effectiveness was therefore interpreted descriptively. Clearer implementation reporting helped explain how interventions

were delivered, but it did not allow a firm conclusion that stronger fidelity reporting directly explained stronger effects.

Table 11
Implementation fidelity, process reporting, and effectiveness pattern by study

Study	Fidelity or process information reported	Delivery and feasibility pattern	Effectiveness pattern
Brito et al. (2024)	Teacher-training/engagement component reported; outcome missingness relevant for accelerometry	Teacher-focused intervention was feasible within the school context but effects were arm-specific	Partly effective; sedentary behaviour improved only in the teacher-training arm
Cai et al. (2025)	Limited process information; school-based delivery described	Intervention appeared feasible within regular school context	Effective; positive effects on cardiorespiratory fitness, working memory, and body composition in girls
Castilho dos Santos et al. (2024)	Some intervention and delivery information reported; fidelity evidence limited	Multicomponent intervention combined PE-based sessions and mHealth support	Partly effective; lower-body muscular fitness improved, but physical activity and CRF did not
Liao et al. (2022)	Structured supervision and instructor-to-student ratio reported	Delivery relied on research-linked or specialist instructors, which supports internal control but limits scalability	Effective; movement quality, muscular strength/power, and flexibility improved
Lima et al. (2022)	Teacher-workshop component reported; multi-arm delivery allowed comparison of added PE and workshop components	Teacher-focused professional development was more effective than increasing PE time alone	Partly effective; mental health improved in the workshop arm, and executive function improved in the combined arm
Liu et al. (2026)	Attendance, intensity monitoring, intervention completion, and safety indicators reported	HIFT and MICT were delivered within the school setting with high feasibility indicators	Effective; physical fitness, mood-related outcomes, and HIFT body-composition outcomes improved
Müller et al. (2019)	Multicomponent implementation described; delivery included external or university-linked support	Intervention was school-based but required external support and multiple components	Partly effective; body composition improved, but CRF did not
Shao & Cao (2025)	Exercise protocol and dose comparison reported; intervention delivery was structured	Jump-rope exercise was feasible as a school-based training format but relied on structured supervision	Effective; CRF and body composition improved in both jump-rope arms
Wan et al. (2025)	Intervention arms and resistance-training delivery described; intensity and training structure reported	School-based resistance-training formats were feasible but effects differed by outcome	Partly effective; clearest between-group effects were for muscular strength/power
Zhang et al. (2024)	Functional training structure and progression reported	Intervention was integrated into school-based training and showed motor-skill relevance	Partly effective; fundamental movement skill effects were positive but test-dependent

Study	Fidelity or process information reported	Delivery and feasibility pattern	Effectiveness pattern
Zhou et al. (2019)	Strong process information, including school-based implementation indicators	Multicomponent school programme combined PE, after-school PA, teacher training, and school support	Effective; broad physical/cardiorespiratory fitness effects were reported
Zou et al. (2025)	Core-training protocol reported; fidelity/process reporting was more limited	Intervention appeared feasible as a school-based core-strength programme	Effective; multiple fitness, motor-competence, and static-balance outcomes improved

Note. CRF = cardiorespiratory fitness; HIFT = high-intensity functional training; MICT = moderate-intensity continuous training; PA = physical activity; PE = physical education. Fidelity and process information refer to what was reported by the studies or coded in the extraction sheet. Limited fidelity reporting was not interpreted as evidence of poor implementation.

As shown in Table 11, the studies with the most detailed implementation reporting were generally easier to interpret. Zhou et al. reported broad school-based process information alongside consistent physical-fitness effects across the CHAMPS contrasts. Liu et al. also provided relatively detailed feasibility indicators, including attendance, intensity monitoring, completion, and safety information, which strengthened the interpretation that both HIFT and MICT could be implemented within a school setting. In these studies, implementation reporting supported the plausibility and feasibility of the reported effects, although it did not demonstrate that fidelity caused the outcomes.

Teacher-focused interventions formed a particularly important implementation pattern: in both Brito et al. and Lima et al., the clearest effects occurred in the teacher-training or teacher-workshop arms (Section 3.1)

Several studies showed positive effects but relied on more controlled or specialist-supported delivery. Liao et al. used structured functional strength training with close supervision, while Shao and Cao implemented a structured jump-rope training protocol. These designs support internal consistency and may explain the clear physical-fitness effects, but they also raise questions about scalability if similar supervision or specialist support is required in ordinary school settings. Müller et al. also involved external or university-linked support within a multicomponent programme, which strengthens delivery capacity but complicates interpretation of real-world implementation demands.

Participant flow and missing outcome data provided an additional implementation-related context for interpreting the findings. As shown in Table 12, dropout and missing-data patterns varied substantially across studies. Several trials reported no or low dropout, whereas others had moderate or uneven attrition across study arms. These patterns were not treated as evidence of poor implementation, but they affected how confidently study-level effects could be interpreted, especially where findings depended on specific outcome subsamples or where attrition differed between intervention and control conditions.

Table 12
Participant flow, dropout, and missing-data concerns by study

Study	Dropout or missing-data information	Relevance for interpretation
Brito et al. (2024)	Large accelerometry data loss. Per-arm loss from randomised accelerometry subsample to post-monitoring: IG1 = 43.7%, IG2 = 46.9%, IG3 = 28.9%, CG = 23.2%. Overall attrition: 703 to 391 valid cases = 44.4%.	Important missing-data concern because the significant sedentary-behaviour findings depended on accelerometry-valid subsamples.

Study	Dropout or missing-data information	Relevance for interpretation
Cai et al. (2025)	0% dropout reported.	Participant flow did not substantially limit interpretation, although generalisability remained limited by the single regional/ethnic sample.
Castilho dos Santos et al. (2024)	IG schools: 3.0% loss from allocation to 24-week follow-up; CG schools: 6.8%.	Attrition was relatively low; null PA and CRF findings are therefore less likely to be explained by dropout alone.
Liao et al. (2022)	IG: 0.8%; CG: 2.3%.	Low dropout supports interpretation of the intervention effects, although specialist delivery remains a scalability concern.
Lima et al. (2022)	IG1: approximately 19%; IG2: approximately 17%; IG3: approximately 9%; CG: 20%.	Moderate attrition should be considered when interpreting arm-specific mental-health and cognition effects.
Liu et al. (2026)	IG1: 13.3%; IG2: 11.1%; CG: 20%.	Attrition was higher in the control group; this should be considered when interpreting group-by-time effects.
Müller et al. (2019)	PA: 18.3%; PA + HE: 8.1%; PA + HE + NU: 12.9%; HE + NU: 23.7%; control schools: 48.1%.	High and uneven attrition, especially in control schools, limits confidence in interpreting body-composition and CRF patterns.
Shao & Cao (2025)	JRE-1: 15.0%; JRE-2: 12.5%; CG: 20.0%.	Moderate attrition should be considered, but positive CRF and body-composition effects were still reported for both jump-rope arms.
Wan et al. (2025)	0% dropout reported.	Participant flow did not substantially limit interpretation, though the small female-only sample limits generalisability.
Zhang et al. (2024)	IG: 8.0%; CG: 6.0%.	Low-to-moderate attrition; interpretation is affected more by test-dependent findings than by dropout.
Zhou et al. (2019)	SPE: 18.5%; ASP: 9.5%; SPE + ASP: 5.6%; CG: 3.4%.	Differential attrition should be considered, but broad planned-contrast effects were reported across the CHAMPS intervention conditions.
Zou et al. (2025)	0% dropout reported.	Participant flow did not substantially limit interpretation, although the single-school design and lack of long-term follow-up remain limitations.

Note. ASP = after-school programme; CG = control group; CRF = cardiorespiratory fitness; HE = health education; IG = intervention group; JRE = jump-rope exercise; NU = nutrition component; PA = physical activity; SPE = school physical education intervention. Dropout and missing-data information are reported descriptively and are not treated as evidence of poor implementation.

Table 12 shows that participant flow was not a uniform limitation across the evidence base. Cai et al., Wan et al., and Zou et al. reported no dropout, and several other studies reported relatively low attrition. In contrast, Brito et al. had a major missing-data concern because the sedentary-behaviour findings depended on accelerometry-valid subsamples, with substantial accelerometry-related data loss. Müller et al. also showed uneven attrition across intervention and control schools. Moderate attrition was reported in studies such as Lima et al., Liu et al., Shao and Cao, and Zhou et al., which should be considered when interpreting arm-specific or group-by-time effects. Overall, participant flow affected the interpretability of some findings but

did not provide a consistent explanation for why some interventions were more effective than others.

Across the included studies, no trial assessed sustainability in the sense of maintenance after the intervention period. Several studies provided evidence of feasibility during implementation, such as school-based delivery, attendance, protocol completion, or intensity monitoring, but these indicators did not show whether effects were maintained after the intervention ended. Therefore, RQ2 can only be answered in relation to implementation and feasibility during the trial, not long-term sustainability.

Overall, implementation fidelity appeared most useful as an explanatory context for interpreting effects rather than as an independent predictor of effectiveness. Studies with clearer process reporting allowed more confident interpretation of how interventions were delivered, while studies with limited reporting left more uncertainty. The evidence suggests that teacher training, structured delivery, and clear process monitoring may support effectiveness, but the relationship between fidelity and outcomes remained exploratory because fidelity reporting, participant flow, and study designs differed substantially.

3.3. Intervention Components and Consistency of Effects

The intervention components showed different levels of consistency across outcome domains. As shown in Table 13 the clearest pattern was observed for structured exercise components that directly targeted physical fitness, including functional strength training, resistance training, HIFT, jump-rope exercise, and core strength training. Teacher-focused components also showed a notable pattern, particularly in the Brazilian studies, where teacher training or workshops were associated with improvements in sedentary behaviour and mental health/wellbeing. In contrast, increasing PE time alone showed limited consistency.

Teacher-focused components showed one of the clearest non-fitness patterns: in both Brazilian multi-arm trials, the significant effects occurred in the teacher-focused arms (sedentary behaviour in Brito et al.; depressive symptoms, social isolation, and sleep quality in Lima et al.), whereas the additional-PE arms did not (Section 3.1).

Additional PE time alone showed limited evidence of effectiveness. In both Brito et al. and Lima et al., adding PE lessons without the teacher-focused component did not produce the clearest positive effects. Zhou et al. reported broad physical/cardiorespiratory fitness effects in a school PE intervention, but this was not a simple time-increase intervention. It combined PE provision with curriculum modification, teacher training, after-school activity, and broader school-level support. Across the included studies, additional PE time without further components did not show consistent positive effects.

Structured exercise components were the most consistent for physical-fitness outcomes. Functional strength training in Liao et al., HIFT in Liu et al., jump-rope exercise in Shao and Cao, resistance training in Wan et al., and core strength training in Zou et al. were all associated with improvements in one or more physical-fitness domains. These components were most clearly linked to cardiorespiratory fitness, muscular strength/power, body composition, speed/agility, or static balance. However, the strength of evidence differed between studies. For example, Wan et al. showed the clearest between-group effects for push-ups, while several other outcomes were less clearly significant between groups.

Movement-oriented components also showed positive findings. Liao et al. reported improvements in movement quality measured with the Functional Movement Screen. Zhang et al. reported improvements in fundamental movement skills, particularly gross motor quotient and object control. Zou et al. reported improvements in motor-competence-related outcomes and static balance. The outcome constructs differed across studies.

Evidence for mHealth and motivational-support components was limited as an independent component. Castilho dos Santos et al. included mHealth support as part of a multicomponent intervention, but physical activity did not significantly improve. Brito et al. reported motivational

mediation in the teacher-training arm, but these mediator findings were not coded as direct outcome effects. Therefore, motivational or mHealth elements could not be interpreted as independently effective components in the present synthesis.

High-intensity, aerobic, and jump-rope-based components were most consistently linked to cardiorespiratory fitness or endurance. Cai et al., Liu et al., Shao and Cao, Zhou et al., and Zou et al. all reported positive cardiorespiratory or endurance-related findings. Liu et al. additionally showed that HIFT and MICT were associated with different psychological patterns: HIFT was more clearly linked to positive mood states, whereas MICT was more clearly linked to reductions in negative mood states. Shao and Cao reported positive cardiorespiratory and body-composition effects in both jump-rope arms, while Zhou et al. showed broad fitness gains in a more comprehensive school-based programme.

Multicomponent interventions showed mixed consistency. Zhou et al. provided the clearest example of a combined intervention pattern, with the strongest effects observed where school PE and after-school activity components were combined. Müller et al. also showed positive body-composition effects in a multicomponent school-based intervention. However, Brito et al. and Lima et al. showed that combined arms were not necessarily more effective than single teacher-focused arms. Overall, the evidence did not show that multicomponent interventions were automatically superior. Instead, effects depended on the specific component combination, target outcome, and implementation context.

Overall, RQ3 was answered most clearly for physical-fitness components. Structured exercise components, including functional strength training, resistance training, high-intensity training, jump-rope exercise, and core strength training, showed the most consistent positive association with fitness-related outcomes. Teacher-focused components showed a smaller but important pattern for behavioural and psychosocial outcomes. Additional PE time alone was less consistently effective.

Table 13
Intervention components and consistency of effects

Component group	Studies or arms including the component	Main positive outcome domains	Null, mixed, or limited findings	Overall pattern
Teacher training / PE-teacher workshops	Brito et al. (2024); Lima et al. (2022); Zhou et al. (2019)	Sedentary behaviour; mental health/wellbeing; executive function; physical/cardiorespiratory fitness in Zhou et al. as part of a broader programme	Effects in Brito et al. and Lima et al. were arm-specific; teacher components were not always isolated from other school-level changes	Promising, especially for behavioural and psychosocial outcomes, but evidence is concentrated in Brazil and partly embedded in multicomponent designs
Additional PE lessons / increased PE time	Brito et al. (2024); Lima et al. (2022); Zhou et al. (2019)	Physical/cardiorespiratory fitness in Zhou et al. when embedded in a broader school PE intervention	Additional PE alone did not significantly improve sedentary behaviour in Brito et al. or depressive symptoms in Lima et al.	Limited consistency when implemented as increased time alone
Modified PE curricula / school-based PA programmes	Müller et al. (2019); Zhou et al. (2019); Castilho dos Santos et al. (2024)	Body composition; physical/cardiorespiratory fitness; lower-body muscular strength/power	Physical activity and CRF effects were not significant in Castilho dos Santos et al.; CRF was not significant in Müller et al.	Positive but context-dependent; strongest where curriculum change was embedded in a broader school programme
Strength, resistance, functional, or core training	Liao et al. (2022); Liu et al. (2026); Wan et al. (2025); Zhang et al. (2024); Zou et al. (2025)	Muscular strength/power; movement quality; fundamental movement skills; speed/agility; static balance	Some outcomes were test-dependent or only within-group positive, especially in Wan et al. and Zhang et al.	One of the most consistent component groups for fitness and movement-related outcomes
HIIT, HIFT, MICT, aerobic, or jump-rope training	Cai et al. (2025); Liu et al. (2026); Shao & Cao (2025); Zhou et al. (2019); Zou et al. (2025)	Cardiorespiratory fitness/endurance; body composition; muscular strength/power; mental health/wellbeing	Effects varied by intensity, protocol, and comparator; some control groups also improved	Consistently relevant for cardiorespiratory fitness, with additional effects depending on the intervention format
Motor competence / fundamental	Zhang et al. (2024); Zou et al. (2025); Liao et al. (2022)	Fundamental movement skills; motor-competence-related outcomes; movement quality	Liao et al. assessed Functional Movement Screen rather than TGMD-style FMS; some Zhang	Generally positive, but measurement constructs differed across studies

Component group	Studies or arms including the component	Main positive outcome domains	Null, mixed, or limited findings	Overall pattern
movement skill training			outcomes were test-dependent	
mHealth / motivational support	Castilho dos Santos et al. (2024); Brito et al. (2024); Zhou et al.(2019)	Lower-body muscular strength/power in Castilho dos Santos et al.; motivational mediation in Brito et al.	Physical activity was not significantly improved in Castilho dos Santos et al.; mediator findings in Brito et al. were not outcome effects	Limited direct effectiveness evidence as an independent component
Multicomponent interventions	Castilho dos Santos et al. (2024); Müller et al. (2019); Zhou et al. (2019); combined arms in Brito et al. and Lima et al.	Body composition; physical/cardiorespiratory fitness; lower-body muscular strength/power	Combined arms were not consistently stronger than single-component arms in Brito et al. or Lima et al.	Potentially effective, but more components did not automatically produce stronger effects

Note. CRF = cardiorespiratory fitness; FMS = fundamental movement skills; HIFT = high-intensity functional training; HIIT = high-intensity interval training; MICT = moderate-intensity continuous training; PA = physical activity; PE = physical education; TGMD = Test of Gross Motor Development. The table summarises descriptive component–outcome patterns. Because several interventions combined multiple elements, the patterns should not be interpreted as isolated causal effects of single components unless the study design separated intervention arms.

4. Discussion

4.1. Summary of Main Findings

This systematic comparative review examined QPE-aligned school-based randomised controlled trials in relation to country and context, implementation fidelity, and intervention components. The findings indicate heterogeneous effects across the included studies. Six studies were coded as effective and six as partly effective, but effects were usually limited to specific outcome domains, intervention arms, subgroups, or measurement instruments. The main interpretation is therefore not that QPE-aligned interventions were uniformly effective, but that observed effects varied according to the relationship between intervention content, target outcome, delivery conditions, and implementation support.

Physical fitness outcomes showed the most frequent positive findings. Cardiorespiratory fitness or endurance improved in several studies, especially where interventions included structured aerobic activity, high-intensity exercise, jump-rope exercise, core training, or broader school-based physical activity programmes. Muscular strength and power also improved in several interventions based on resistance training, functional strength training, high-intensity functional training, or core strength training. This pattern is consistent with wider school-based physical activity evidence indicating that structured interventions often show clearer effects on fitness-related outcomes than on broader daily behaviour outcomes such as total physical activity or sedentary time (Neil-Sztramko et al., 2021; Podnar et al., 2021; Wu et al., 2023).

Behavioural outcomes were less consistent. Physical activity did not clearly improve in either study that assessed it as an outcome (Castilho dos Santos et al., as a primary outcome; Müller et al., via self-report), and sedentary behaviour improved only in the teacher-training arm of Brito et al. (2024). These findings suggest that structured activity during school time does not necessarily change total daily activity or sedentary behaviour. Similar patterns have been reported in school-based physical activity trials using accelerometer-assessed outcomes, where effects on overall physical activity are often small or inconsistent (Love et al., 2019). Body composition also showed mixed findings. Positive effects were observed in some fitness- or adiposity-focused interventions, but not consistently across all studies or subgroups. This interpretation is plausible because body composition is influenced by multiple factors beyond school-based physical activity, including diet, maturation, baseline weight status, total activity outside school, intervention duration, and intervention intensity (Podnar et al., 2021; World Health Organization, 2020).

Movement-related outcomes require separate interpretation because the included studies measured different constructs. Liao et al. (2022) assessed movement quality through the Functional Movement Screen, whereas Zhang et al. (2024) and Zou et al. (2025) assessed fundamental movement skills or motor-competence-related outcomes. These findings indicate positive effects in movement-related domains, but they should not be collapsed into one general motor-competence conclusion. This distinction is relevant to the QPE framework, which defines PE quality through broader physical, cognitive, affective, and social learning aims rather than through physical performance alone (Dudley et al., 2022; UNESCO, 2015b). Mental health and cognitive outcomes were reported in fewer studies and were often arm-specific. Findings from Lima et al. (2022), Liu et al. (2026), and Cai et al. (2025) therefore indicate possible benefits in depressive symptoms, mood, executive function, or working memory, but not a consistent pattern across the full evidence base.

The cross-country comparison was limited by the composition of the included studies, which came from three upper-middle-income countries (Section 2.2.1). The review therefore cannot draw conclusions about income-group differences or causal country-level effects; this is taken up in Sections 4.3 and 4.7.

Implementation fidelity was relevant for interpretation but could not be analysed as an independent predictor. Some studies reported attendance, adherence, intensity monitoring, teacher delivery, or implementation barriers, whereas others provided limited process information. Studies with more detailed process reporting were easier to interpret because the reported information clarified how the planned intervention was delivered and monitored. This aligns with implementation-fidelity literature, which distinguishes dimensions such as adherence, dose, quality of delivery, and participant responsiveness (Cepni et al., 2025). Limited fidelity reporting was not treated as evidence of poor implementation, but it reduced the precision with which intervention effects could be linked to implementation processes. This supports the need for more systematic fidelity and process reporting in future QPE-aligned RCTs (Cepni et al., 2025; Errisuriz et al., 2018).

The component-level synthesis indicates that structured content and delivery quality were more informative than PE exposure alone. Structured fitness components, including resistance training, functional training, high-intensity training, jump-rope exercise, and core strength training, were most often associated with physical fitness outcomes. Teacher-focused components were relevant in the Brazilian studies, where teacher training or workshops were associated with behavioural or psychosocial effects more consistently than additional PE time alone. This interpretation is consistent with QPE policy literature, which emphasises teacher education, pedagogical quality, and structured curriculum provision as central features of quality PE (UNESCO, 2015c). Multicomponent interventions were not consistently superior to more focused intervention arms. This is consistent with concerns in PE-based intervention literature that multicomponent programmes are difficult to interpret when component effects cannot be isolated and process evaluation is limited (Errisuriz et al., 2018). Overall, the findings suggest that QPE-aligned interventions should be interpreted through the fit between component, target outcome, school context, and implementation conditions rather than through intervention presence alone.

4.2. Interpretation of Outcome-Domain Patterns

The outcome-domain pattern suggests that the included interventions were most likely to show effects when the measured outcome was directly targeted by the intervention content. This was most visible for physical fitness. Interventions that included structured aerobic exercise, high-intensity training, jump-rope exercise, resistance training, functional strength training, or core strength training more often reported positive findings in cardiorespiratory fitness, endurance, muscular strength, or muscular power. This pattern is consistent with previous school-based physical activity evidence, where structured exercise interventions have shown clearer effects on fitness-related outcomes than on broader daily behaviour outcomes (Neil-Sztramko et al., 2021; Podnar et al., 2021; Wu et al., 2023).

Cardiorespiratory fitness and endurance findings support this interpretation. Positive effects were reported in interventions that included higher-intensity or repeated aerobic stimuli, such as HIFT/MICT, jump-rope exercise, core strength training with repeated movement sequences, and broader school activity programmes. By contrast, studies that did not produce clear cardiorespiratory effects tended to use intervention designs in which the aerobic stimulus was either less central, shorter, harder to isolate, or embedded within multicomponent delivery. This does not imply that multicomponent PE interventions are ineffective for cardiorespiratory fitness. Rather, it indicates that cardiorespiratory outcomes are easier to interpret when the intervention provides a clearly specified aerobic dose and monitors whether that dose was delivered. This interpretation is consistent with reviews of school-based exercise and HIIT interventions, which emphasise the relevance of intensity, structure, and dose for fitness-related adaptations (Duncombe et al., 2022; Wu et al., 2023).

Muscular strength and power showed a similar pattern. Positive findings were most frequent in interventions that directly trained strength, power, or functional movement capacity. This included resistance-based, functional strength, high-intensity functional, and core strength interventions. The pattern is therefore consistent with a specificity interpretation: outcomes

improved most often when they were directly trained and assessed through related field tests. This is relevant for QPE-aligned interventions because improvements in strength, power, and functional movement capacity may support broader participation in PE, sport, and physically active play, but these outcomes remain only one part of the wider QPE framework. QPE is not limited to physical performance; it also includes cognitive, affective, social, and inclusive learning aims (Dudley et al., 2022; UNESCO, 2015a).

Behavioural outcomes were less consistent. The findings do not show a clear pattern of improvement in overall physical activity, and sedentary behaviour improved only in a specific teacher-training arm. This suggests that school-based PE or exercise interventions may increase activity within the intervention setting without necessarily changing total daily activity or sedentary behaviour. This distinction is important because intervention exposure during PE lessons represents only one part of the school day and an even smaller part of total daily behaviour. Previous school-based physical activity reviews have similarly reported small, inconsistent, or context-dependent effects on accelerometer-assessed physical activity, especially when interventions are expected to influence behaviour beyond the lesson or school setting (Love et al., 2019; Neil-Sztramko et al., 2021). Evidence from PE-based intervention reviews also suggests that increasing moderate-to-vigorous physical activity during PE does not necessarily lead to consistent changes in activity outside PE (Errisuriz et al., 2018; Lonsdale et al., 2013).

Body composition findings were mixed and should be interpreted cautiously. Several interventions reported improvements in BMI-related, fat-mass, body-fat, waist, or skinfold outcomes, but these effects were not consistent across all studies, arms, or subgroups. This pattern is plausible because body composition is influenced by factors beyond school-based physical activity, including dietary intake, maturation, baseline weight status, total activity outside school, intervention intensity, intervention duration, and growth-related changes during childhood and adolescence (Podnar et al., 2021; World Health Organization, 2020). Body composition outcomes may therefore require either a stronger energy-balance intervention, longer follow-up, or targeting of specific populations such as children with overweight or obesity. In this review, adiposity-focused effects were clearer in some interventions with higher baseline weight-related risk or explicit fitness/body-composition aims, whereas broader PE-based interventions did not produce uniform body-composition effects.

Movement-related outcomes were positive in several studies but should not be interpreted as one homogeneous domain. Liao et al. assessed movement quality using the Functional Movement Screen, whereas Zhang et al. and Zou et al. assessed fundamental movement skills or motor-competence-related outcomes. These measures are related but not equivalent. Functional Movement Screen performance reflects movement quality across selected functional patterns, whereas fundamental movement skill and motor competence measures more directly assess movement skills that support participation in physical activity and sport. The distinction matters because combining these outcomes too broadly would overstate the consistency of the evidence. Current literature on motor competence and school-based interventions also indicates that effects depend on the type of movement content, the specificity of instruction, and the assessment instrument used (Yin et al., 2025). The findings from this review therefore support the relevance of structured movement-focused PE, but not a single general conclusion about all motor outcomes.

Mental health and cognitive outcomes were less frequently assessed, but the available findings suggest that QPE-aligned interventions may affect outcomes beyond physical fitness. Effects were reported for depressive symptoms, social isolation, sleep quality, mood, executive function, and working memory in selected studies or intervention arms. These findings should be interpreted as preliminary within this review because they were based on fewer studies and less consistent measurement than the physical fitness outcomes. However, they are relevant to the broader QPE framing because QPE is intended to support physical, cognitive, affective, and social development rather than physical performance alone (Dudley et al., 2022; UNESCO, 2015b). Evidence from school-based physical activity and mental

health research also supports the need to assess psychological and cognitive outcomes more systematically, rather than treating them as secondary or incidental outcomes (Andermo et al., 2020; Vaquero-Solís et al., 2020).

Overall, the outcome-domain pattern indicates that the included interventions were most consistent for directly trained physical outcomes and less consistent for broader behavioural, psychosocial, cognitive, or body-composition outcomes. This does not reduce the relevance of QPE to fitness outcomes. Instead, it shows a measurement and evidence imbalance: the current RCT evidence is strongest for outcomes that are easier to target, quantify, and test over short intervention periods, whereas broader QPE outcomes are less consistently measured and less often linked to long-term follow-up. Future QPE-aligned trials should therefore retain fitness and movement outcomes while also measuring physical activity behaviour, sedentary behaviour, motivation, wellbeing, cognition, inclusion, and sustained participation more systematically.

4.3. Cross-Country and Contextual Interpretation

The first research question asked whether the outcomes of QPE-aligned RCTs differed across countries with varying educational, socioeconomic, and cultural contexts. The findings allow only a descriptive answer. As described in Section 2.2.1, all included studies were conducted in upper-middle-income economies in only three countries. Studies from low- and lower-middle-income economies were eligible but did not meet all inclusion criteria. Therefore, the review cannot compare effectiveness across income groups and cannot identify country-level causal effects. The country patterns reported in the Results should instead be interpreted as patterns in the available evidence base.

The concentration of studies in China shaped the overall synthesis. The China-based studies contributed most of the evidence on structured exercise and physical fitness, including functional training, functional strength training, resistance training, jump-rope exercise, high-intensity functional training, moderate-intensity continuous training, core strength training, and a broader school-based physical activity programme. Positive findings in these studies were reported across cardiorespiratory fitness, muscular strength/power, body composition, movement-related outcomes, cognition, and mood-related outcomes. However, these findings do not indicate a uniform China-specific intervention effect. The Chinese studies differed in school level, sample characteristics, target population, intervention content, intervention duration, delivery agent, and outcome measurement. Their contribution is therefore best interpreted as a concentration of structured exercise and fitness-focused trial evidence rather than as evidence that Chinese school contexts were more responsive to QPE-aligned interventions.

The Brazil-based studies contributed a different pattern. Their clearest findings were associated with teacher-focused components rather than with additional PE time alone (Sections 3.2 and 3.4). Castilho dos Santos et al. (2024) did not report clear improvements in physical activity or cardiorespiratory fitness, but reported a positive effect on lower-body muscular fitness. The Brazilian studies therefore suggest that teacher capacity and pedagogical intervention components may be relevant for behavioural and psychosocial outcomes. This interpretation is consistent with QPE policy literature, which treats teacher education, curriculum quality, and pedagogical support as central elements of quality PE provision (UNESCO, 2015c). It should not be read as evidence of a Brazil-specific effect, because the available Brazilian studies differed in intervention design, outcome focus, and study population.

The South African evidence was limited to one study. Müller et al. (2019) tested a multicomponent school-based physical activity and health intervention in a disadvantaged primary-school context and reported favourable body-composition effects but no significant cardiorespiratory fitness effect. Because only one South African study was included, this finding cannot support broader conclusions about South African school contexts or African LMIC settings. It should be treated as a single-study contribution showing that multicomponent

school-based interventions can produce selective effects under specific contextual and implementation conditions.

Across countries, contextual variables were not reported consistently enough to explain differences in effectiveness. School level, socioeconomic setting, teacher qualifications, PE timetable structure, facilities, delivery agent, and intervention resources varied across studies, but these factors could not be isolated as moderators. Cultural context was also not measured directly. It could only be considered indirectly through country, school setting, intervention design, and delivery conditions. This limits the interpretation of RQ1. Country and income group were useful for describing the evidence base, but they were not sufficient explanatory variables.

This limitation is relevant to the wider literature. Barbosa Filho et al. (2016) reported that physical activity intervention evidence for children and adolescents in LMICs was underrepresented in previous reviews and that findings from available studies could not easily be generalised across all LMIC contexts. The present review confirms a related problem for QPE-aligned RCT evidence. Although the review focused on non-high-income settings, the final evidence base did not include low- or lower-middle-income countries and was concentrated in a small number of upper-middle-income contexts. As a result, the findings cannot be used to infer how QPE-aligned interventions would function in lower-resource education systems.

The main conclusion for RQ1 is therefore cautious. The included studies differed by country, but the observed differences were more closely tied to available intervention types, study designs, outcome domains, and implementation conditions than to country or income group alone. The evidence suggests that QPE-aligned school-based RCTs should be interpreted through specific school and delivery contexts, including teacher capacity, PE structure, intervention resources, supervision, and fidelity reporting. This is consistent with QPE policy framing, which emphasises that PE quality depends on context-sensitive curriculum, trained teachers, inclusive provision, and sustainable school structures rather than on programme presence alone (UNESCO, 2015a, 2021b). Future studies should therefore report contextual information more systematically so that cross-country interpretation can move beyond descriptive comparison.

4.4. Implementation Fidelity as Explanatory Context

The second research question asked how implementation fidelity related to the effectiveness of QPE-aligned RCTs across national settings. The findings do not allow fidelity to be analysed as an independent predictor of effectiveness because implementation reporting differed substantially across studies. Instead, fidelity functioned as an explanatory context for interpreting how confidently intervention effects could be linked to the planned intervention.

Several studies reported implementation indicators such as attendance, adherence, intensity monitoring, teacher training, delivery agent, dropout, or barriers to delivery. These details were important because they clarified whether the intervention was likely to have been delivered as planned. In studies with more detailed process reporting, the interpretation of effects was more precise because planned dose, delivered dose, and participant exposure could be distinguished more clearly. This distinction is central to implementation-fidelity research, which separates dimensions such as adherence, dose, quality of delivery, and participant responsiveness (Cepni et al., 2025). It is also relevant in school-based PE interventions, where scheduled lesson time does not necessarily reflect active learning time, movement intensity, or student engagement (Hollis et al., 2017; Lonsdale et al., 2013).

The findings from this review indicate that stronger fidelity reporting was associated with clearer interpretation, not necessarily with stronger effectiveness. Studies such as Zhou et al. (2019) and Liu et al. (2026) provided more information on implementation structure, activity exposure, or intensity monitoring, which made it easier to assess whether observed effects were compatible with the intervention dose. By contrast, studies with limited fidelity information

could still report positive outcomes, but the mechanism linking intervention design, implementation, and outcome change was less clear. Limited fidelity reporting was therefore not treated as evidence of poor implementation. It was treated as a limitation in interpretability.

Teacher-focused interventions further illustrate the importance of fidelity as context. In the Brazilian studies, teacher training or teacher workshops were associated with more consistent behavioural or psychosocial findings than additional PE time alone. This pattern suggests that teacher capacity may influence how intervention content is translated into practice. This interpretation is consistent with QPE policy literature, which identifies trained teachers, pedagogical quality, and curriculum implementation as central elements of quality PE provision (UNESCO, 2015b). It also aligns with PE intervention literature showing that implementation quality and process evaluation are needed to understand why school-based interventions produce effects in some contexts but not others (Errisuriz et al., 2018; Smedegaard et al., 2017).

The distinction between researcher-supported delivery and scalable school delivery is also relevant. Some studies used specialist instructors, personal trainers, external monitoring, or highly structured protocols. These features may increase internal validity during the trial period, but they also limit conclusions about routine implementation. An intervention delivered by trained researchers or specialists under controlled conditions may not produce the same effects when delivered by regular PE teachers with larger classes, fewer resources, or competing school demands. This issue is especially relevant for QPE-aligned interventions because QPE is not only concerned with short-term efficacy, but also with sustainable, inclusive, and context-sensitive PE provision (UNESCO, 2015b, 2021b).

Attrition and participation also affected interpretation. Some studies reported low or no dropout, whereas others had substantial or uneven loss of data. This was particularly relevant where objective outcome measures required valid wear time or where follow-up participation differed between intervention arms. Attrition does not automatically invalidate findings, but it affects the certainty with which results can be interpreted. Participation and retention are therefore part of implementation quality, especially in school-based physical activity interventions where student exposure and data completeness may vary across classes, schools, or subgroups (Van Der Wurff et al., 2024).

A further limitation is that fidelity was usually assessed only during the intervention period. The included studies did not provide sufficient evidence on whether teachers, schools, or students maintained intervention practices after the trial ended. This restricts conclusions about sustainability. From a QPE perspective, this is important because policy-relevant school interventions should not only produce short-term effects under research conditions, but also be feasible within regular school structures. Sustainability requires attention to teacher workload, training needs, equipment, timetable constraints, administrative support, and compatibility with the existing PE curriculum (UNESCO, 2015b, 2021b).

Overall, implementation fidelity should be understood as a condition for interpreting effectiveness rather than as a separate outcome in this review. The available evidence suggests that QPE-aligned RCTs require more systematic reporting of adherence, delivered dose, delivery quality, participant responsiveness, attrition, and sustainability. Without this information, it remains difficult to determine whether mixed findings reflect weak intervention content, insufficient implementation, low exposure, contextual barriers, or measurement limitations.

4.5. Intervention Components and Consistency of Effects

The third research question asked which intervention components showed the most consistent effectiveness across countries. The findings indicate that structured fitness-oriented components were most consistently associated with positive physical outcomes. These included resistance training, functional strength training, high-intensity functional training, moderate-intensity continuous training, jump-rope exercise, core strength training, and

broader structured school-activity programmes. The consistency of these findings appears to reflect the direct relationship between intervention content and measured outcomes. Fitness outcomes were most often positive when the intervention provided a specific training stimulus for the same domain. This pattern is consistent with wider evidence on school-based exercise interventions. Reviews of school-based physical activity and fitness interventions indicate that structured exercise content can improve physical fitness in children and adolescents, particularly when intensity, frequency, and progression are specified (Duncombe et al., 2022; Wu et al., 2023). In the present review, interventions targeting cardiorespiratory fitness or endurance through repeated aerobic or high-intensity activity showed clearer effects than interventions in which aerobic exposure was less central or harder to isolate. Similarly, interventions targeting strength, power, or functional movement capacity more often reported improvements in related field-based fitness outcomes. These findings support a component-specific interpretation rather than a general conclusion that all PE-based interventions improve all physical domains.

Additional PE time alone was not consistently associated with better outcomes. As reported in Section 3.4, additional PE time without a teacher-focused component produced limited or inconsistent effects in the Brazilian trials. This suggests that the quality and organisation of PE may be more relevant than the amount of scheduled PE time alone. This interpretation is consistent with QPE policy literature, which identifies teacher education, curriculum quality, pedagogical support, and inclusive delivery as central features of quality PE provision (UNESCO, 2015b, 2021b). It is also consistent with PE-based intervention literature showing that scheduled PE time does not necessarily represent active, high-quality, or developmentally meaningful movement time (Hollis et al., 2017; Lonsdale et al., 2013).

Teacher-focused components were particularly relevant for interpreting behavioural and psychosocial outcomes. Teacher training, workshops, and curriculum support may influence how PE content is delivered, how students engage with lessons, and how physical activity is framed within the school environment (Ha et al., 2020). In this review, the clearest teacher-focused findings came from Brazil, where teacher-oriented components were more informative than additional PE time alone. However, the evidence was limited to a small number of studies, and the components differed across interventions. Therefore, the findings should be interpreted as support for the relevance of teacher capacity, not as proof that teacher training is sufficient by itself.

Movement-focused components also showed positive findings, but the evidence was heterogeneous. Functional strength training, functional training, and core strength training were associated with improvements in movement quality, fundamental movement skills, balance, or motor-competence-related outcomes. However, the studies used different assessment tools and outcome definitions. The findings therefore support the inclusion of structured movement content in QPE-aligned interventions, but they do not support a single pooled conclusion across all motor outcomes. This distinction is consistent with motor-competence literature, which indicates that intervention effects depend on the specific movement content, instructional design, age group, and assessment instrument (Yin et al., 2025).

Multicomponent interventions require cautious interpretation. Several included studies combined curriculum change, teacher training, additional activity time, structured exercise, environmental support, motivational components, parental engagement, or digital support. Such designs may be useful in school settings because they address multiple determinants of behaviour and learning. However, they also make it difficult to identify which component produced the observed effect. In the present review, multicomponent interventions were not consistently superior to more focused interventions. This aligns with PE-based intervention literature showing that limited process evaluation and component overlap can make it difficult to determine why an intervention was effective or ineffective (Errisuriz et al., 2018).

Digital or mHealth components could not be interpreted as independent drivers of effectiveness. In some studies, digital components were used as motivational support, information delivery, or monitoring tools, but they were not always isolated from other intervention elements. Where digital support was provided to both intervention and control groups, its independent contribution could not be assessed. These findings suggest that mHealth should be interpreted as a support component rather than as a stand-alone QPE intervention unless future trials test it separately and report engagement, exposure, and implementation data.

Overall, the component-level synthesis indicates that the most consistent effects occurred when intervention components were clearly matched to outcome domains. Structured fitness components were most consistently linked to physical fitness outcomes, movement-focused components were linked to movement-related outcomes, and teacher-focused components appeared relevant for behavioural, psychosocial, or cognitive outcomes. The evidence does not support the conclusion that more components automatically produce stronger effects. Instead, it suggests that component specificity, pedagogical quality, implementation support, and outcome alignment are central to interpreting QPE-aligned RCT effectiveness.

4.6. Relation to Quality Physical Education

The findings should be interpreted in relation to QPE alignment rather than as evidence on fully developed QPE systems. The included studies were eligible because they contained school-based PE or physical-activity components that aligned with QPE-relevant aims, such as structured movement learning, fitness development, teacher involvement, curriculum adaptation, inclusive school delivery, or broader child and adolescent development. However, most studies did not define themselves as QPE interventions. This distinction is important because QPE is a broader policy and educational framework, whereas the included interventions were specific trial-based programmes with defined components, outcomes, and follow-up periods.

The strongest findings in this review were concentrated in physical fitness and movement-related outcomes. This reflects the design of many included studies, which used targeted exercise, strength, aerobic, core, jump-rope, or functional movement content. These outcomes are compatible with QPE because physical development, movement competence, and participation in physical activity are central to PE. However, they do not represent the full scope of QPE. UNESCO's QPE framework emphasises teacher education, inclusive provision, curriculum quality, child-centred learning, and lifelong participation, rather than short-term physical performance alone (UNESCO, 2015b, 2021b). Similarly, PE-based intervention evidence has conceptualised quality PE through multiple learning domains, including psychomotor, cognitive, affective, and social outcomes (Dudley et al., 2022). This creates a difference between the breadth of the QPE framework and the outcome profile of the available RCT evidence. The review identified consistent evidence for several measurable physical outcomes, but fewer studies assessed outcomes such as motivation, enjoyment, inclusion, social learning, student voice, autonomy, or long-term participation. Mental health and cognitive outcomes were reported in some studies, but they were not measured consistently across the evidence base. Therefore, the available RCT evidence captures only part of what QPE aims to achieve. The findings are most informative for the physical and movement-related dimensions of QPE, while broader educational, affective, social, and inclusion-related outcomes remain less developed in the included trial evidence.

Teacher capacity was one of the clearest links between the review findings and QPE. In the Brazilian studies, teacher training or teacher workshops were more informative than additional PE time alone for behavioural, psychosocial, or cognitive outcomes. This pattern is consistent with QPE policy, which treats trained teachers and pedagogical quality as central conditions for effective PE provision (UNESCO, 2015b). It also suggests that QPE alignment should not be judged only by whether an intervention adds activity time or increases exercise dose. The way PE is planned, taught, supported, and integrated into school routines is also relevant.

The review also shows that QPE-aligned interventions vary in their degree of scalability. Some interventions used low-cost or easily transferable components, such as jump rope, bodyweight strength exercises, core training, or structured PE lesson content. Others required specialist instructors, personal trainers, heart-rate monitoring, extensive equipment, or close researcher supervision. These features may support internal validity in RCTs, but they limit conclusions about routine school implementation. From a QPE perspective, effectiveness under trial conditions is only one consideration. Programmes also need to be feasible for regular teachers, compatible with school timetables, adaptable to local resources, and sustainable after the research period (UNESCO, 2015b, 2021b).

The findings therefore support a cautious interpretation of QPE alignment. The included studies show that school-based, QPE-aligned interventions can improve selected physical, movement-related, behavioural, psychosocial, or cognitive outcomes. They do not show that short-term RCT effects are sufficient evidence of comprehensive QPE provision. QPE requires attention to educational quality, teacher development, inclusion, curriculum, context, and sustainability. Future trials should therefore measure not only fitness, body composition, physical activity, and motor outcomes, but also outcomes that reflect the wider aims of QPE, including motivation, perceived competence, enjoyment, inclusion, social participation, cognitive engagement, and sustained physical activity beyond the intervention period.

4.7. Strengths and Limitations

This review has several strengths and limitations that should be considered when interpreting the findings.

Its strengths include a predefined, PRISMA-guided procedure with a structured three-database search (PubMed, SPORTDiscus, and Web of Science) translated through the Polyglot Search Translator, transparent eligibility criteria, and standardised extraction into a two-level dataset that separated study-level information from outcome-level evidence. Effectiveness was coded on between-group rather than within-group evidence, direct-contrast and interaction-based findings were distinguished, and intervention content was coded at the arm level so that study design (single- versus multi-arm) was kept separate from component composition (single- versus multicomponent). Source tables and figures were used as the primary extraction source.

Several limitations also apply. As a one-person Master's thesis, the review was conducted by a single reviewer, so screening, extraction, and coding were not independently duplicated and no inter-rater reliability could be calculated. The synthesis was narrative rather than quantitative; no pooled effect estimate was calculated because the studies differed in design, outcomes, measurement tools, comparators, dosage, and statistical reporting. The final evidence base was restricted to twelve studies from upper-middle-income economies and was geographically uneven (eight from China, three from Brazil, one from South Africa), which constrained the cross-country comparison and prevented any comparison across income groups. No study reported post-intervention follow-up, so the maintenance of effects could not be assessed.

A further limitation concerns risk of bias. A formal risk-of-bias assessment using a standardised instrument, such as the Cochrane Risk of Bias 2 tool (Sterne et al., 2019), was not conducted, consistent with the single-reviewer scope of the thesis and the descriptive aim of the synthesis; risk of bias was therefore not used to weight or grade the evidence, and the effectiveness judgements should be read as descriptive synthesis aids rather than certainty-of-evidence ratings. Several bias-relevant signals were nonetheless noted descriptively during extraction and are reflected in the participant-flow reporting (Table 12), including substantial or uneven attrition in some trials (for example, the accelerometry data loss in Brito et al. and the differential school-level attrition in Müller et al.), reliance on subsample data for some outcomes, and the limited feasibility of participant or assessor blinding in school-based physical activity trials. These signals were treated as context for interpretation rather than as

a formal appraisal, and the absence of a structured risk-of-bias and certainty-of-evidence assessment such as GRADE (Guyatt et al., 2008) remains a limitation of the review.

Taken together, these limitations mean that the review describes patterns in the available QPE-aligned RCT evidence rather than establishing the comparative efficacy or causal mechanisms of specific intervention components. The findings should therefore be interpreted as a structured and transparent synthesis of a small and heterogeneous evidence base rather than as definitive estimates of effect.

5. Conclusion

This systematic comparative review examined whether QPE-aligned school-based RCTs were effective across different national contexts, how implementation fidelity related to effectiveness, and which intervention components showed the most consistent effects. Twelve RCTs from three upper-middle-income countries, comprising 4’999 randomised pupils, met all inclusion criteria. Six studies were coded as effective and six as partly effective; no study was entirely ineffective. The clearest and most consistent benefits were found for physical fitness and movement-related outcomes following structured exercise components such as resistance, functional, high-intensity, and jump-rope training. Teacher-focused components were associated with the most informative behavioural and psychosocial findings, whereas additional PE time alone showed limited and inconsistent effects.

Taken together, these findings support a central premise of UNESCO's Quality Physical Education framework: the quality, structure, and pedagogical organisation of physical education appear more decisive for intervention effectiveness than the amount of scheduled PE time. Component specificity also mattered, as outcomes improved most reliably when the intervention content directly targeted the measured domain. At the same time, the review demonstrates the current limits of the evidence base. Because all included studies were conducted in upper-middle-income economies and were concentrated in three countries, the first research question could only be answered descriptively, and no conclusions about national or income-group effects are possible. Fidelity reporting was uneven across trials, so the relationship between implementation and effectiveness could be explored but not systematically tested, and no trial assessed whether effects were maintained after the intervention ended.

Three implications follow for research and practice. First, more RCTs of QPE-aligned interventions in low- and lower-middle-income countries are needed before the QPE framework's claim to global applicability can be evaluated empirically; the present evidence base does not reach the education systems with the fewest resources. Second, future trials should report implementation fidelity and process indicators systematically, including attendance, adherence, teacher delivery, and dose delivered and should include post-intervention follow-up, because without such data neither the mechanisms nor the sustainability of effects can be established. Third, for policy-makers and schools, the findings suggest that investing in teacher training, pedagogical support, and well-structured movement content is likely to yield more consistent benefits than increasing PE exposure alone. In this respect, the review provides cautious but consistent empirical support for prioritising the quality of physical education over its mere quantity, while underlining that the international evidence base for QPE-aligned interventions remains narrow, uneven, and incomplete.

6. References

Andermo, S., Hallgren, M., Nguyen, T.-T.-D., Jonsson, S., Petersen, S., Friberg, M., Romqvist, A., Stubbs, B., & Elinder, L. S. (2020). School-related physical activity interventions and mental health among children: A systematic review and meta-analysis. *Sports Medicine - Open*, *6*(1), 25. https://doi.org/10.1186/s40798-020-00254-x

Brito, A. L. D. S., Tassitano, R. M., Mélo, E. N., Santos, S. J. D., Mota, J., & Barros, M. V. G. D. (2024). Efeito de uma intervenção em aulas de educação física sobre a redução do comportamento sedentário em adolescentes. *Cadernos de Saúde Pública*, *40*(11), e00211023. https://doi.org/10.1590/0102-311xpt211023

Cai, X., Liu, Y., Li, Y., & Chen, Z. (2025). Physical education exercise improves physical fitness and working memory function: A randomized controlled trial study of Uyghur adolescents in Xinjiang, China. *BMC Pediatrics*, *25*(1), 122. https://doi.org/10.1186/s12887-025-05511-0

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ*, l6890. https://doi.org/10.1136/bmj.l6890

Castilho Dos Santos, G., De Souza Silva, T. M., Da Silva, J. M., De Oliveira Barbosa, R., Kennedy, S. G., Lubans, D. R., & Stabelini Neto, A. (2024). Impact of the ActTeens Program on physical activity and fitness in adolescents: A cluster randomized controlled trial. *BMC Pediatrics*, *24*(1), 447. https://doi.org/10.1186/s12887-024-04922-9

Cepni, A. B., Ledoux, T. A., Walsh, D. W., & Johnston, C. A. (2025). Evaluating implementation fidelity of a school-based physical education intervention: Effects on child physical activity. *Health Education Research*, *40*(6), cyaf047. https://doi.org/10.1093/her/cyaf047

Clark, J. M., Sanders, S., Carter, M., Honeyman, D., Cleo, G., Auld, Y., Booth, D., Condron, P., Dalais, C., Bateup, S., Linthwaite, B., May, N., Munn, J., Ramsay, L., Rickett, K., Rutter, C., Smith, A., Sondergeld, P., Wallin, M., … Beller, E. (2020). Improving the translation of search strategies using the Polyglot Search Translator: A randomized controlled trial. *Journal of the Medical Library Association*, *108*(2), 195–207. https://doi.org/10.5195/jmla.2020.834

DeepL SE. (n.d.). *DeepL Übersetzer | Der präziseste Übersetzer der Welt*. Retrieved June 8, 2026, from https://www.deepl.com

Dudley, D., Mackenzie, E., Van Bergen, P., Cairney, J., & Barnett, L. (2022). What Drives Quality Physical Education? A Systematic Review and Meta-Analysis of Learning and Development Effects From Physical Education-Based Interventions. *Frontiers in Psychology*, *13*, 799330. https://doi.org/10.3389/fpsyg.2022.799330

Duncombe, S. L., Barker, A. R., Bond, B., Earle, R., Varley-Campbell, J., Vlachopoulos, D., Walker, J. L., Weston, K. L., & Stylianou, M. (2022). School-based high-intensity interval training programs in children and adolescents: A systematic review and meta-analysis. *PLOS ONE*, *17*(5), e0266427. https://doi.org/10.1371/journal.pone.0266427

Errisuriz, V. L., Golaszewski, N. M., Born, K., & Bartholomew, J. B. (2018). Systematic Review of Physical Education-Based Physical Activity Interventions Among Elementary School Children. *The Journal of Primary Prevention*, *39*(3), 303–327. https://doi.org/10.1007/s10935-018-0507-x

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, *336*(7650), 924–926. https://doi.org/10.1136/bmj.39489.470347.AD

Ha, A. S., Lonsdale, C., Lubans, D. R., & Ng, J. Y. Y. (2020). Increasing Students' Activity in Physical Education: Results of the Self-determined Exercise and Learning For FITness Trial. *Medicine and Science in Sports and Exercise*, *52*(3), 696–704. https://doi.org/10.1249/MSS.0000000000002172

Hollis, J. L., Sutherland, R., Williams, A. J., Campbell, E., Nathan, N., Wolfenden, L., Morgan, P. J., Lubans, D. R., Gillham, K., & 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), 52. https://doi.org/10.1186/s12966-017-0504-0

Liao, T., Duhig, S. J., Du, G., Luo, B., & Wang, Y. T. (2022). The Effect of a Functional Strength Training Intervention on Movement Quality and Physical Fitness in Adolescents. *Perceptual and Motor Skills*, *129*(1), 176–194. https://doi.org/10.1177/00315125211056865

Lima, R. A., de Barros, M. V. G., Bezerra, J., Dos Santos, S. J., Monducci, E., Rodriguez-Ayllon, M., & Soares, F. C. (2022). Universal school-based intervention targeting depressive symptoms in adolescents: A cluster randomized trial. *Scandinavian Journal of Medicine & Science in Sports*, *32*(3), 622–631. https://doi.org/10.1111/sms.14115

Liu, J., Zhao, Y., Ren, X., Wang, B., Chen, Y., & Tang, C. (2026). Differential effects of school-based high-intensity functional training and moderate-intensity continuous training on physical fitness and mood States in early adolescents: A cluster randomised trial. *BMC Psychology*, *14*(1), 262. https://doi.org/10.1186/s40359-026-04156-7

Lonsdale, C., Rosenkranz, R. R., Peralta, L. R., Bennie, A., Fahey, P., & Lubans, D. R. (2013). A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. *Preventive Medicine*, *56*(2), 152–161. https://doi.org/10.1016/j.ypmed.2012.12.004

Love, R., Adams, J., & van Sluijs, E. M. F. (2019). Are school-based physical activity interventions effective and equitable? A meta-analysis of cluster randomized controlled trials with accelerometer-assessed activity. *Obesity Reviews : An Official Journal of the International Association for the Study of Obesity*, *20*(6), 859–870. https://doi.org/10.1111/obr.12823

Müller, I., Schindler, C., Adams, L., Endes, K., Gall, S., Gerber, M., Htun, N. S. N., Nqweniso, S., Joubert, N., Probst-Hensch, N., du Randt, R., Seelig, H., Smith, D., Steinmann, P., Utzinger, J., Yap, P., Walter, C., & Pühse, U. (2019). Effect of a Multidimensional Physical Activity Intervention on Body Mass Index, Skinfolds and Fitness in South African Children: Results from a Cluster-Randomised Controlled Trial. *International Journal of Environmental Research and Public Health*, *16*(2). https://doi.org/10.3390/ijerph16020232

Neil-Sztramko, S. E., Caldwell, H., & Dobbins, M. (2021). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database of Systematic Reviews*, *2021*(9). https://doi.org/10.1002/14651858.CD007651.pub3

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*, n71. https://doi.org/10.1136/bmj.n71

Pearson, N., Braithwaite, R., & Biddle, S. J. H. (2015). The Effectiveness of Interventions to Increase Physical Activity Among Adolescent Girls: A Meta-analysis. *Academic Pediatrics*, *15*(1), 9–18. https://doi.org/10.1016/j.acap.2014.08.009

Podnar, H., Jurić, P., Karuc, J., Saez, M., Barceló, M. A., Radman, I., Starc, G., Jurak, G., Đurić, S., Potočnik, Ž. L., & Sorić, M. (2021). Comparative effectiveness of school-based interventions targeting physical activity, physical fitness or sedentary behaviour on obesity prevention in 6- to 12-year-old children: A systematic review and meta-analysis. *Obesity Reviews*, *22*(2), e13160. https://doi.org/10.1111/obr.13160

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic*

reviews: A product from the ESRC Methods Programme. https://doi.org/10.13140/2.1.1018.4643

Shao, S., & Cao, M. (2025). Integrating interval jump rope exercise into a school setting improve body composition, cardiorespiratory fitness and glycolipid metabolism parameters in adolescents with overweight and obesity: A randomized controlled trial. *BMC Pediatrics*, *25*(1), 945. https://doi.org/10.1186/s12887-025-06320-1

Smedegaard, S., Brondeel, R., Christiansen, L. B., & Skovgaard, T. (2017). What happened in the “Move for Well-being in School”: A process evaluation of a cluster randomized physical activity intervention using the RE-AIM framework. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 159. https://doi.org/10.1186/s12966-017-0614-8

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*, l4898. https://doi.org/10.1136/bmj.l4898

Suter, D. (2024). *A Literature Review of “Quality Physical Education” Interventions*. University of Basel.

UNESCO. (2015a). *Promote quality physical education policy—UNESCO Digital Library*. https://unesdoc.unesco.org/ark:/48223/pf0000233111

UNESCO. (2015b). *Quality physical education: Guidelines for policy makers*.

UNESCO. (2015c). *Quality physical education: Policy guidelines: Methodology*. https://unesdoc.unesco.org/ark:/48223/pf0000233812

UNESCO. (2021a). *How to influence the development of quality physical education policy: A policy advocacy toolkit for youth*.

UNESCO. (2021b). *Quality Physical Education Policy Project: Analysis of process, content and impact*. https://unesdoc.unesco.org/ark:/48223/pf0000376151

Van Der Wurff, I., Kirschner, M., Golsteijn, R., De Jonge, M., Berendsen, B., Singh, A., Savelberg, H., & De Groot, R. (2024). School-based physical activity interventions: Which intervention characteristics are associated with participation and retention? A meta-analysis. *Preventive Medicine*, *182*, 107925. https://doi.org/10.1016/j.ypmed.2024.107925

Vaquero-Solís, M., Gallego, D. I., Tapia-Serrano, M. Á., Pulido, J. J., & Sánchez-Miguel, P. A. (2020). School-based Physical Activity Interventions in Children and Adolescents: A Systematic Review. *International Journal of Environmental Research and Public Health*, *17*(3), 999. https://doi.org/10.3390/ijerph17030999

Veritas Health Innovation. (2026). *Covidence* [Computer software]. www.covidence.org

Wan, J., Li, Z., Sun, R., Gu, J., Li, Y., & Zhang, Q. (2025). In-school resistance training improves physical fitness in pubertal girls: A randomized controlled trial. *BMC Sports Science, Medicine and Rehabilitation*, *17*(1), 302. https://doi.org/10.1186/s13102-025-01351-8

World Bank Group. (2026). *World Bank Country and Lending Groups*. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

World Health Organization. (2020). *WHO Guidelines on Physical Activity and Sedentary Behaviour* (1st ed). World Health Organization.

World Health Organization. (2022a). *Global status report on physical activity 2022*. s.n.

World Health Organization. (2022b, April 3). *85% of adolescent girls don't do enough physical activity: New WHO study calls for action*. https://www.who.int/europe/news/item/04-03-2022-85-of-adolescent-girls-don-t-do-enough-physical-activity-new-who-study-calls-for-action

Wu, J., Yang, Y., Yu, H., Li, L., Chen, Y., & Sun, Y. (2023). Comparative effectiveness of school-based exercise interventions on physical fitness in children and adolescents: A systematic review and network meta-analysis. *Frontiers in Public Health*, *11*, 1194779. https://doi.org/10.3389/fpubh.2023.1194779

Yin, X., Zhang, D., Shen, Y., Wang, Y., Wang, Z., & Liu, Y. (2025). Effectiveness of school-based interventions on fundamental movement skills in children: A systematic review and meta-analysis. *BMC Public Health*, *25*(1), 1522. https://doi.org/10.1186/s12889-025-22696-2

Zhang, D., Kim Geok, S., Mun Chan, Y., Zaremohzzabieh, Z., Kim Lam, S., & He, S. (2024). Exploring the effects of a 12-Week functional training program on fundamental motor skills for primary school children aged 6–7. *Children and Youth Services Review*, *167*, 108008. https://doi.org/10.1016/j.childyouth.2024.108008

Zhou, Z., Li, S., Yin, J., Fu, Q., Ren, H., Jin, T., Zhu, J., Howard, J., Lan, T., & Yin, Z. (2019). Impact on Physical Fitness of the Chinese CHAMPS: A Clustered Randomized Controlled Trial. *International Journal of Environmental Research and Public Health*, *16*(22). https://doi.org/10.3390/ijerph16224412

Zou, X., Eungpinichpong, W., & Konharn, K. (2025). Effects of a 12-week core strength training program on physical fitness and fundamental movement skills in children aged 6–8 years. *Journal of Human Sport and Exercise*, *21*(1), 93. https://doi.org/10.55860/d10kqs35

Appendix

Study 1 of 12: Brito et al. (2024)

Effect of an intervention in physical education classes on the reduction of sedentary behavior in adolescents — Anísio Luis da Silva Brito; Rafael Miranda Tassitano; Edilânea Nunes Mélo; Simone José dos Santos; Jorge Mota; Mauro Virgílio Gomes de Barros

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
Brazil (Latin America & Caribbean)	Upper-middle income	Four-arm factorial cluster RCT (IG1 = additional PE; IG2 = PE-teacher training/engagement; IG3 = additional PE + PE-teacher training/engagement; CG = control)	School	854 = accelerometry subsample randomly drawn for monitoring at baseline 703 = valid accelerometry data at baseline (Substantial missing-accelerometry attrition) IG1 = 231 IG2 = 258 IG3 = 180 CG = 185	15.0 ± 1.1; 57.5% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
First year of Ensino Médio; Upper secondary	Public; Unclear	No; NR	NR	NR	Partly Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess the effect of different PE-class intervention strategies (additional PE lessons; teacher training and engagement; combination of both) on the reduction of daily sedentary time in high-school adolescents.
Study design	Multi-arm
Baseline comparability	Partly; baseline differences in age for boys and girls, and in girls' intrinsic/extrinsic motivation; baseline sedentary behaviour not significantly different

INTERVENTION	
Component composition	Teacher Training; Additional PE Lessons; Motivation / Enjoyment
Component description	IG1/Group A: PE lessons increased from 2 to 4 sessions/week; schools independently reorganised timetables to accommodate extra classes within the regular school day. IG2/Group B: 120-hour continuing-education course for PE teachers, certified by the University of Pernambuco; included 6 face-to-face sessions (30 h total) and a community action project with students (90 h total). Topics included physical activity, sedentary behaviour, physical fitness, adolescent brain function, mental health, social relations/isolation, sleep quality, and pedagogical skills. IG3/Group C: combination of Group A and Group B strategies. CG/Group D: control
Dosage	A: Additional PE lessons · Length: 6 months · Sessions/wk: 4 PE/wk · Total: NR (no class length reported) B: Teacher training · Length: 6 months · Sessions/wk: N/A · Total: N/A (120-h teacher course) C: A + B combined · Length: 6 months · Sessions/wk: 4 PE/wk · Total: NR D: Control · Length: 6 months · Sessions/wk: 2 PE/wk · Total: -
Dosage specifications	Structural (not intensity-based) intervention: no exercise-intensity target prescribed or HR-monitored. IG1 doubled PE (2→4 classes/wk), with extra classes scheduled within the regular timetable → additive PE time, not a substitution. IG2 = 120-h certified teacher course (Univ. of Pernambuco), 30 h across six in-person sessions with support materials each session (remainder delivered as a community action project). 6-month intervention (April–Oct 2017; July holiday excluded).

Required equipment	low · described by the authors as a "simple and low-cost" intervention; teacher training materials provided in IG2/IG3; no specialist equipment for IG1.
Control condition	CG/Group D: usual PE, 2 PE lessons/week; no additional PE lessons and no PE-teacher training/engagement strategy.

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: No · Teacher Delivery: Yes (routine + trained) · Intensity Monitored: No · Scalability: Yes · Sustainability: Not assessed · Attendance: No
Dropout (%)	Per-arm loss from the randomised accelerometry sample to those re-monitored at post: · IG1: 43.7% · IG2: 46.9% · IG3: 28.9% · CG: 23.2% · Overall valid-data attrition: · 703→391 = 44.4%; not reported per arm at the valid-data stage
Feasibility / scalability (reviewer-inferred)	1. Requires timetable reorganisation to fit the extra PE lessons within regular school hours; · 2. Teacher arm requires delivering a 120-h certified course and depends on teacher engagement
Reported barriers	NR

OUTCOMES & FINDINGS	
Relevant outcomes	Sedentary behaviour (daily sedentary time, min/day, accelerometry) · PE motivation (external, internal, amotivation) measured as mediators only, not outcomes
Sex / subgroup reporting	Reported by Sex: Yes · Reported by Subgroup: No Sex/Subgroup Details and Findings: Results stratified by sex throughout. Boys in IG2 showed a larger reduction in sedentary time (−61.5 min/day, d=0.61) than girls (−34.8 min/day, d=0.32). For IG1 (girls), a non-significant trend toward increased sedentary time, associated with increased amotivation and decreased intrinsic motivation. The IG2 reduction was significantly mediated by decreased amotivation in both sexes.
Main positive outcome domains	Sedentary behaviour (daily sedentary time - IG2/teacher-training arm only, both sexes)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*								
Comparison	**Subgroup**	**Domain**	**Direction**	**Sig. (between-group)**	**Effect (with unit)**	**95% CI / statistic**	**p-value**	**Notes**
IG1 (extra PE) vs control [Single-component arm]	Boys	Sedentary behaviour	Favours intervention	No	-0.106 (std β)	-0.610 to 0.540	0.53	Between-group regression coefficient
IG1 (extra PE) vs control [Single-component arm]	Girls	Sedentary behaviour	Favours control	No	+0.215 (std β)	-0.071 to 0.507	0.14	Increase in SB
IG2 (teacher training) vs control [Single-component arm]	Boys	Sedentary behaviour	Favours intervention	Yes	-61.5 min/day	-102.5 to -20.6	<0.004	d=0.61; std β -0.402
IG2 (teacher training) vs control [Single-component arm]	Girls	Sedentary behaviour	Favours intervention	Yes	-34.8 min/day	-67.3 to -2.4	<0.035	d=0.32; std β -0.376
IG3 (combined) vs control [Multi-component arm]	Boys	Sedentary behaviour	Favours intervention	No	-0.127 (std β)	-0.489 to 0.235	0.48	—
IG3 (combined) vs control [Multi-component arm]	Girls	Sedentary behaviour	Favours intervention	No	-0.265 (std β)	-0.580 to -0.049	0.09	CI excludes 0 but p=0.09
IG2 (teacher training) vs control	Boys	External motivation (mediator)	—	No	+0.108 (std β)	—	—	Mediator path (Group B); exclude from outcome vote-count
IG2 (teacher training) vs control	Girls	External motivation (mediator)	—	No	-0.071 (std β)	—	—	Mediator; exclude from vote-count
IG2 (teacher training) vs control	Boys	Internal motivation (mediator)	—	No	+0.032 (std β)	—	—	Mediator; exclude from vote-count

IG2 (teacher training) vs control	Girls	Internal motivation (mediator)	—	No	+0.146 (std β)	—	—	Mediator; exclude from vote-count
IG2 (teacher training) vs control	Boys	Amotivation (mediator)	—	Yes	-0.468 (std β)	—	<0.05	Mediator; indirect effect on SB 0.022 (p<0.05); exclude from vote-count
IG2 (teacher training) vs control	Girls	Amotivation (mediator)	—	Yes	-0.302 (std β)	—	<0.05	Mediator; indirect effect on SB 0.056 (p<0.05); exclude from vote-count

SUMMARY & APPRAISAL	
Key findings (summary)	Only the teacher-training arm (IG2) significantly reduced daily sedentary time, in both sexes; the additional-PE (IG1) and combined (IG3) arms were non-significant (see outcome-level results above). The IG2 effect was significantly mediated by a reduction in amotivation in both sexes. Among IG1 girls, sedentary time showed a non-significant increase, alongside increased amotivation and reduced intrinsic motivation.
Implications for future research	Include long-term follow-up assessments; · develop strategies to minimise attrition; · involve teachers as active agents of change through horizontal planning; · replicate in other school contexts; · recognise that adding PE lessons without teacher preparation may be insufficient or even counterproductive
Study limitations	1. Accelerometers may underestimate SB (cannot distinguish standing from sitting); 2. Varied literature criteria for data reduction and cut-points; 3. ≈54% loss to follow-up (refusals, absences, two schools not assessed, invalid data) may have biased results; 4. No follow-up data for long-term effects; 5. Limited accelerometer availability required sub-sampling

Study 2 of 12: Cai et al. (2025)

Physical education exercise improves physical fitness and working memory function: a randomized controlled trial study of Uyghur adolescents in Xinjiang, China — Xiaodi Cai; Yuefeng Liu; Yanhong Li; Zhiqiang Chen

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Two-arm parallel RCT (IG1 = physical education exercise; CG = no intervention)	Individual	222 IG1 n=114 CG n = 108	14.03 ± 0.17; 47.7% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
NR; secondary school students; Secondary, unspecified	Unclear; Urban	Yes - regional economic disadvantage specified, not school-resource level; Xinjiang as western China, multi-ethnic, mainly Uyghur, and states that the region's economic development is relatively backward	NR; Authors note Uyghur population has higher obesity rates	Chinese Uyghur	Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess the effects of 12 weeks of classroom physical education exercise on physical fitness and working memory function in Uyghur adolescents
Study design	Single-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Aerobic Training; Modified PE Curriculum; Motivation/Enjoyment; Other: culturally adapted physical exercise/Uyghur ethnic dance
Component description	IG1: Classroom PE exercise intervention developed from the Chinese PE and Health Curriculum Standards and adapted to the Xinjiang context. Delivered under guidance of a full-time PE teacher plus assistant. Content included soccer, basketball, aerobics combining Uyghur ethnic dance and aerobic exercise, and sports games. Activities were alternated to maintain motivation. Conducted on the school athletic field, or in the gymnasium during poor weather. CG: no-intervention/usual condition as reported.
Dosage	Exercise intervention · Length: 12 weeks · Sessions/wk: 3/wk · Total: 1,440 min (36 × 40 min)
Dosage specifications	36 sessions over 12 weeks. Intensity target 140–160 bpm, monitored in real time via Polar Gofit Team Heart Rate Monitoring System (worn on wrist/arm, displayed on tablet); students outside the HR range were alerted. Primarily high-intensity interval exercise combining aerobic and anaerobic work. No follow-up beyond immediate post-test.
Required equipment	moderate · sports balls and court equipment for basketball and soccer; Polar Gofit
Control condition	CG: no intervention; regular school/PE routine not described in detail.

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Unclear (only overall completion reported, no adherence metric) · Teacher Delivery: Yes (routine) · Intensity Monitored: Yes · Scalability: No · Sustainability: Not assessed · Attendance: Unclear (all 222 completed; session-level attendance not reported)
Dropout (%)	0% overall
Feasibility / scalability (reviewer-inferred)	1. Real-time HR monitoring (Polar Gofit + tablet) implies equipment demand; · 2. Structured multi-sport programme implies trained-instructor needs
Reported barriers	NR

OUTCOMES & FINDINGS	
Relevant outcomes	Cardiorespiratory fitness (VO_2max) · Body composition (height, weight, BMI, waist circumference) · Cognition/working memory (1- and 2-back reaction time)
Sex / subgroup reporting	Reported by Sex: · Yes · Reported by Subgroup: · No Sex/Subgroup Details and Findings: · Results reported separately by sex; both sexes improved within-group on VO_2max and on weight/BMI/waist. The main sex difference was a larger waist-circumference reduction in girls (−2.02 cm) than boys (−0.4 cm); VO_2max gain was somewhat larger in boys (+4.08 vs +2.87 mL/kg/min). Analyses are within-group; ANCOVA used for the overall between-group test.
Main positive outcome domains	Cardiorespiratory fitness (VO_2max - both sexes); Body composition (weight, BMI, waist - girls; boys not clearly supported); Cognition (working memory - n-back reaction time)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
Exercise vs control [Single-component arm]	Overall	Cardiorespiratory fitness	Favours intervention	Yes (both sexes)	VO_2max higher post (ANCOVA)	—	<0.001	Within-group +4.08 (boys)/+2.87 (girls) mL/kg·min
Exercise vs control [Single-component arm]	Girls	Body composition	Favours intervention	Yes	Weight, BMI, waist lower (ANCOVA)	—	<0.05	Within-group: BMI -0.26 kg/m2, waist -2.02 cm, weight -0.40 kg
Exercise vs control [Single-component arm]	Boys	Body composition	Favours intervention	No	-	—	—	Within-group changes sig but between-group ns (no ANCOVA superscript in Table 1); waist only -0.4 cm

Exercise vs control [Single-component arm]	Overall	Cognition (working memory)	Favours intervention	Yes	n-back RT between-group diff: boys 1-/2-back -115.21/-198.88 ms; girls -135.05/-192.84 ms	—	boys <0.05; girls <0.01	Between-group ANCOVA values (not within-group). Within-group: boys -150.5/-180.0; girls -204.8/-228.5 ms

SUMMARY & APPRAISAL	
Key findings (summary)	Twelve weeks of high-intensity classroom PE improved cardiorespiratory fitness, body composition and working memory within the intervention group, while the control group did not change. Between-group effects were clearest and most consistent for VO_2max in both sexes; for body composition the effect was clear in girls but not clearly supported in boys, whose clearest result was VO_2max. Working-memory (n-back) reaction time improved between-group in both sexes (see outcome-level results). The authors attribute part of the effect to ethnic dance sustaining students' motivation.
Implications for future research	Replicate in other ethnic and population groups; · use longer interventions with follow-up to assess durability; · use objective body-composition measures rather than BMI imputation or VO_2max projection; · incorporate culturally preferred activities (e.g. ethnic dance) to boost motivation in minority adolescents; · strengthen school PE time and extracurricular activity
Study limitations	1. Small sample (n=222), Uyghur-only → limited generalisability; 2. 12-week duration, no long-term follow-up; 3. Single-school concentration → possible Hawthorne effect; 4. BMI and VO_2max calculated/projected, not objectively measured

Study 3 of 12: Castilho dos Santos et al. (2024)

Impact of the ActTeens Program on physical activity and fitness in adolescents: a cluster randomized controlled trial — Géssika Castilho dos Santos; Thais Maria de Souza Silva; Jadson Marcio da Silva; Rodrigo de Oliveira Barbosa; Sarah G. Kennedy; David R. Lubans; Antonio Stabelini Neto

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
Brazil (Latin America & Caribbean)	Upper-middle income	Two-arm cluster RCT (IG1 = ActTeens Program; CG = control)	School	317 IG n = 169 CG n = 148	13.61 ± 0.7; 52.7% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
Grades 8 and 9; Lower secondary	Public; Urban	No; NR	NR	NR	Partly Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess the impact of the ActTeens Program on physical activity and health-related physical fitness in adolescents in Brazil
Study design	Single-arm
Baseline comparability	Partly; groups broadly similar, but significant baseline differences for age and sports practice (IG 13.4% vs CG 22.9% practising sport)

INTERVENTION	
Component composition	Modified PE Curriculum; Strength / Resistance Training; Aerobic Training; Motivation / Enjoyment; Other: mHealth/Digital Guidance
Component description	IG1: ActTeens intervention delivered within PE lessons and adapted from the Australian Resistance Training for Teens (RT4T) programme. Sessions combined muscle-strengthening and aerobic exercises, movement-based games, and a dynamic stretching warm-up. Session format: 3-min warm-up, 15-min structured PA, 2-min cool-down with static stretching. The

	programme targeted basic psychological needs: autonomy, competence, and relatedness. Students selected their own small groups (4–5 students) and selected activity cards from aerobic and resistance categories. mHealth component: weekly WhatsApp messages with videos/infographics promoting active behaviour and healthy eating to adolescents and parents. Content included education/counselling, practice tips, goal-setting, challenges, and healthy lifestyle guidance.
Dosage	ActTeens (structured PA + mHealth) · Length: 24 weeks · Sessions/wk: 2/wk · Total: 960 min (48 × 20 min) + mHealth (24 messages)
Dosage specifications	Structured sessions delivered in 83.3% of PE lessons. Reach 88.3% (317/359). Retention 95.3% (302/317). Attendance ≥75%: 79 students at 12-wk and 57 students at 24-wk (in IG). Assessments at baseline, 12-wk (PA only), and 24-wk (PA + fitness). mHealth messages read by <half (42.1%) of adolescents.
Required equipment	low · activity cards (aerobic and resistance categories) and bodyweight-based resistance exercises; minimal equipment; WhatsApp (smartphones) for the mHealth messages.
Control condition	CG: continued PE classes as usual; received the same WhatsApp/mHealth healthy-behaviour messages as IG.

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Yes · Teacher Delivery: Unclear · Intensity Monitored: Yes · Scalability: Yes · Sustainability: Not assessed · Attendance: Yes
Dropout (%)	IG schools: 3% from allocation to 24-week-follow-up · CG schools: 6.8% from allocation to 24-week-follow-up
Feasibility / scalability (reviewer-inferred)	NR
Reported barriers	1. Authors note typical muscular-fitness/resistance programmes need substantial space and equipment and run outside school; ActTeens is designed to overcome these (within PE, inexpensive minimal equipment - circuit cards, fitness packs, bodyweight) and is argued scalable incl. to disadvantaged/low–middle-income settings 2. Low mHealth uptake (only 42.1% of adolescents read the messages)

OUTCOMES & FINDINGS	
Relevant outcomes	Physical activity (PE-based and total PA, PAQ-A) · Muscular strength/power (standing long jump, push-ups, handgrip) · Cardiorespiratory fitness (20 m PACER)
Sex / subgroup reporting	Reported by Sex: No · Reported by Subgroup: No Sex/Subgroup Details and Findings:
Main positive outcome domains	Muscular strength/power (lower-body - standing long jump).

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
Intervention vs control [Multi-component arm]	Overall	Muscular strength/power	Favours intervention (partial)	Yes	+12.9 cm (standing long jump, post 24 wk)	3.2 to 22.2	<0.05	Upper-body/handgrip/push-ups ns (between-group)
Intervention vs control [Multi-component arm]	Overall	Cardiorespiratory fitness	No difference	No	ns	—	—	—
Intervention vs control [Multi-component arm]	Overall	Physical activity	No difference	No	+0.3 / -0.02 (PAQ-A score: PE-PA / total PA)	PE-PA -0.1 to 0.6	—	Primary outcome; between-group ns (group-time p=0.12/0.94 and 0.81/0.91)

SUMMARY & APPRAISAL	
Key findings (summary)	The intervention did not improve the primary outcome (physical activity); only lower-body muscular fitness (standing long jump) favoured the intervention, with no effect on upper-body muscular fitness or cardiorespiratory fitness (see outcome-level results). The authors attribute the null physical-activity result to activity being confined to PE lessons and to the use of a subjective measure.
Implications for future research	Use device-based physical-activity measures (accelerometry) rather than questionnaires; · add physical-activity opportunities beyond PE (recess, after-school, active transport) in line with the TEO framework; · conduct a scale-up/dissemination evaluation; · better isolate the mHealth component; · monitor session intensity objectively (e.g. heart rate)
Study limitations	1. No HR/intensity control of sessions; 2. PAQ-A self-report (recall bias, no intensity differentiation); 3. mHealth read by <half (42.1%) of adolescents; 4. mHealth delivered to both arms → effectiveness unverifiable; 5. Out-of-school sport (notably CG) not captured → may inflate CG fitness gains

Study 4 of 12: Liao et al. (2022)

The Effect of a Functional Strength Training Intervention on Movement Quality and Physical Fitness in Adolescents — Ting Liao; Steven J. Duhig; Geng Du; Bingting Luo; Yong Tai Wang

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Two-arm parallel RCT (IG1 = functional strength training; CG = regular PE)	Individual	266 IG1 n = 133 CG n = 133	14.35 ± 0.57; 50% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
Second year of middle or high school; Mixed secondary	Unclear; Urban	No; NR	NR	NR	Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess (1) whether the FST program produces greater gains in movement ability and physical performance than regular PE, and (2) the effect of FST on younger vs older adolescents
Study design	Single-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Modified PE Curriculum; Motor Competence / FMS; Strength / Resistance Training

Component description	IG1: 12-week functional strength training (FST) intervention delivered in three progressive phases: (1) flexibility and stability, (2) functional movement patterns, and (3) health-related functional strength. Programme included 55 exercises in total. CG: regular PE.
Dosage	FST · Length: 12 weeks · Sessions/wk: 3/wk · Total: 1,620 min (36 × 45 min)
Dosage specifications	2 sets × 8–12 reps or 10–20 s holds; moderate intensity (RPE 6–7, Borg); weekly progressive loading +5% upper body / +10% lower body
Required equipment	high · Phase 1 largely bodyweight; Phase 2 adds elastic bands, dumbbells, kettlebells, Swiss ball, Bosu ball, medicine ball, cable/pulley station, chair and wedge
Control condition	CG: regular PE classes, same frequency as FST group

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: No · Teacher Delivery: No · Intensity Monitored: Unclear (RPE 6–7 collected only at end of first weekly session) · Scalability: No · Sustainability: Not assessed · Attendance: No
Dropout (%)	IGs: 0.8% · CG: 2.3%
Feasibility / scalability (reviewer-inferred)	1. High staffing/expertise: Two FMS-certified specialist instructors at 1:10, not the regular PE teacher; unlikely to transfer to typical LMIC PE staffing; · 2. Equipment-intensive: Kettlebells, Swiss/Bosu balls, cable station, motion-capture for assessment → cost/space barrier in low-resource settings
Reported barriers	NR

OUTCOMES & FINDINGS	
Relevant outcomes	Movement quality (Functional Movement Screen total score)
Sex / subgroup reporting	Reported by Sex: · No · Reported by Subgroup: · No Sex/Subgroup Details and Findings:

Main positive outcome domains	Movement quality (Functional Movement Screen); Muscular strength/power (curl-ups, pull-ups); Flexibility (sit-and-reach)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
FST vs control [Single-component arm]	Overall	Movement quality (Functional Movement Screen)	Favours intervention	Yes	17.25 vs 13.21 (Functional Movement Screen total score, post)	F(1,260)=102.273	<0.01	Between-group (** horizontal comparison, Table 2). FMS here = Functional Movement Screen (movement-quality battery), not directly comparable to TGMD-style fundamental movement skills (Zhang) or the motor-competence/power items in Zou.
FST vs control [Single-component arm]	Overall	Muscular strength/power	Favours intervention	Yes	Curl-ups & pull-ups higher in FST	curl-ups F(1,127)=50.843; pull-ups F(1,131)=8.068	curl-ups <0.01; pull-ups <0.05	Curl-ups (females), pull-ups (males). FST>control at T3 (curl-ups 45.21 vs 36.17 reps; pull-ups 8.58 vs 6.44 reps).
FST vs control [Single-component arm]	Overall	Flexibility	Favours intervention	Yes	Sit-and-reach higher in FST	F(1,260)=21.854	<0.01	All participants; FST>control T1–T3 (T3 17.98 vs 11.55

								cm). Non-significant between-group: standing broad jump (power, F=0.511), 50 m sprint (speed, F=0.437), endurance run 800 m/1000 m (all p>0.05).

SUMMARY & APPRAISAL	
Key findings (summary)	Significant time-by-group effects emerged by weeks 4–8 (T2) and strengthened by week 12 (T3) for movement quality and for the strength and flexibility tests; endurance, power and speed did not differentiate the groups (see outcome-level results).
Implications for future research	Embed functional skills training within middle- and high-school PE; · examine sex and age differences directly; · equate exercise intensity between conditions
Study limitations	1. Control PE lacked FST's individualised instruction/attention → attention/instruction confound; 2. Limited recruitment across grade bands; 3. Sex and MS-vs-HS effects not directly tested; 4. Exercise intensity not equated between FST and PE

Study 5 of 12: Lima et al. (2022)

Universal school-based intervention targeting depressive symptoms in adolescents: A cluster randomized trial — Rodrigo Antunes Lima; Mauro Virgílio Gomes de Barros; Jorge Bezerra; Simone José dos Santos; Elena Monducci; Maria Rodriguez-Ayllon; Fernanda Cunha Soares

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
Brazil (Latin America & Caribbean)	Upper-middle income	Four-arm cluster RCT (IG1 = doubled PE; IG2 = PE-teacher workshop; IG3 = doubled PE + PE-teacher workshop; CG = control)	School	1279 IG1, IG2, IG3 in total n = ~1'000 CG n = ~250	NR; 90% 14-16 years old; 56.4% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
NR; high-school students aged 14–16; Upper secondary	Unclear; Unclear	No; NR	NR	NR	Partly Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess (1) the effects of three interventions (doubling PE; PE-teacher workshop; workshop + doubling PE) on depressive symptoms in first-year high-school adolescents in Brazil, and (2) whether social isolation, social anxiety, sleep quality, executive function, and attention mediate the intervention effect on depressive symptoms
Study design	Multi-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Teacher Training; Additional PE Lessons
Component description	IG1: doubled PE from 2 to 4 classes/week, increasing weekly PE time from 1 h 40 min to 3 h 20 min for 6 months. IG2: PE-teacher workshop consisting of five lectures of approximately 4 h each, covering pedagogical and health-related topics; teachers planned and delivered PE lessons with feedback sessions. IG3: combined doubled PE and PE-teacher workshop. CG: control.
Dosage	A: Doubled PE · Length: 6 months · Sessions/wk: 4 PE/wk · Total: ~4,800 min (200 min/wk × ~24 wk) B: PE-teacher workshop · Length: 6 months · Sessions/wk: N/A · Total: 1,200 min (5 × 4 h, one-off) C: A + B combined · Length: 6 months · Sessions/wk: 4 PE/wk · Total: ~4,800 min + 1,200 min D: Control · Length: 6 months · Sessions/wk: 2 PE/wk · Total: -
Dosage specifications	Structural (not intensity-based) intervention; no exercise-intensity target prescribed or monitored. Doubled-PE arm additive: 1:40 h→3:20 h PE/wk (2→4 classes) within the regular timetable, not a substitution. Teacher workshop = five lectures, ~4 h each. Implementation note: one school in the workshop+doubling-PE arm did not implement the intervention after baseline; its 96 students were retained under intention-to-treat analysis. 6-month intervention (April–Oct 2017; July holiday excluded).
Required equipment	none
Control condition	CG: usual PE; no doubled PE and no PE-teacher workshop

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: No · Teacher Delivery: Yes (routine + trained) · Intensity Monitored: No · Scalability: Yes · Sustainability: Not assessed · Attendance: No
Dropout (%)	IG1: ~19% · IG2: ~17% · IG3: ~9% · CG: 20%
Feasibility / scalability (reviewer-inferred)	1. Requires structured PE-teacher workshops (teacher time/training); · 2. Workshop appeared ineffective for adolescents already high at baseline → may need a bolt-on programme

Reported barriers	1. One school in the workshop + doubling-PE arm declined to implement the intervention after baseline (handled by intention-to-treat)

OUTCOMES & FINDINGS	
Relevant outcomes	Mental health/wellbeing (depressive symptoms, social isolation, social anxiety, sleep quality) Cognition (executive function - WCST; attention - D2)
Sex / subgroup reporting	Reported by Sex: · Yes · Reported by Subgroup: Yes Sex/Subgroup Details and Findings: · The workshop affected depressive symptoms and social isolation in girls only. Significant benefit (prevention of high symptomatology) only for adolescents low at baseline, no effect for those already high at baseline.
Main positive outcome domains	Mental health/wellbeing (depressive symptoms, social isolation, sleep quality - workshop arm IG2 only); Cognition (executive function - WCST perseverative errors, combined arm IG3 only),

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
IG1 (doubling PE) vs control [Single-component arm]	Overall	Mental health/wellbeing	No difference	No	Depressive symptoms -0.947 points	-3.176 to 1.281	—	—
IG2 (workshop) vs control [Single-component arm]	Overall (esp. girls)	Mental health/wellbeing	Favours intervention	Yes	Depressive symptoms -2.495 points	-4.668 to -0.323	<0.05	Also social isolation -4.759 [-9.025,-0.493]; sleep -0.560 [-1.108,-0.012]
IG3 (combined) vs control [Multi-component arm]	Overall	Mental health/wellbeing	No difference	No	Depressive symptoms +0.726 points	-1.558 to 3.009	—	Combined arm not effective on depression

IG3 (combined) vs control [Multi-component arm]	Overall	Cognition (executive function)	Favours intervention	Yes	WCST perseverative errors -1.294	-2.216 to -0.371	<0.05	Higher WCST = worse, so negative = improvement; attention (D2) ns all arms; executive function ns for IG1/IG2

SUMMARY & APPRAISAL	
Key findings (summary)	The teacher workshop (IG2) significantly reduced depressive symptoms, social isolation and poor sleep quality versus control, with social isolation mediating about 32% of the effect; the workshop and workshop-plus-doubling arms reduced the risk of high depressive symptomatology among adolescents low at baseline. The combined arm (IG3) also improved executive function (WCST perseverative errors), though attention (D2) was unaffected in all arms. Increasing PE quantity alone (IG1) was ineffective (see outcome-level results).
Implications for future research	Extend the intervention and follow-up period; · combine the workshop with an additional component to reach adolescents already high at baseline; · test in other settings; · adequately power and design analyses for sex differences
Study limitations	1. Testers not blinded to allocation; 2. No long-term follow-up (funding) → maintenance unknown; 3. Mediators and primary outcome measured at the same time point → weak causal inference; 4. Sex-difference findings post hoc and underpowered

Study 6 of 12: Liu et al. (2026)

Differential effects of school-based high-intensity functional training and moderate-intensity continuous training on physical fitness and mood States in early adolescents: a cluster randomised trial — Junwei Liu; Yanlin Zhao; Xianqi Ren; Bihan Wang; Yishan Chen; Changfa Tang

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Three-arm cluster RCT (IG1 = HIFT; IG2 = MICT; CG = usual PE)	Class	135 IG1 n = 45 IG2 n = 45 CG n = 45	11.98 ± 0.13; 50% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
Grade 7; Lower secondary	Public; Unclear	No; NR	NR	NR	Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess and compare the psychological and physical effects of HIFT vs MICT as 8-week school-based interventions for early adolescents
Study design	Multi-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Modified PE Curriculum; Strength / Resistance Training; Aerobic Training; Motivation / Enjoyment
Component description	IG1/HIFT: three-phase progressive high-intensity functional training. Weeks 1–3: 30 s work / 30 s rest at 75–80% HRmax. Weeks 4–6: 30 s work / 20 s rest at 80–85% HRmax. Weeks 7–8: 40 s work / 20 s rest at 85–88% HRmax. Each set included 2 rounds

	with 3-min rest. Exercises included dynamic movements such as high knees, push-ups, burpees, bodyweight resistance exercises, planks, Russian twists, hexagonal station rotation, and partner/relay games. Friday sessions included teamwork relays in two approximately 11-min segments. IG2/MICT: continuous rhythm-based aerobic circuit set to music at 65–75% HRmax. Three-phase constant-intensity format with five 6-min cycles, each including 4 min aerobic work and 2 min strength/toning. Friday session was a 30-min group fitness run with task/teamwork challenges. CG: usual PE.
Dosage	HIFT · Length: 8 weeks · Sessions/wk: 3/wk · Total: 960 min (24 × 40 min) · MICT · Length: 8 weeks · Sessions/wk: 3/wk · Total: 960 min (24 × 40 min)
Dosage specifications	Two sessions are 40-min regular morning PE; one is a 40-min Friday after-school extracurricular session. Standard school PE curriculum continued for all arms (intervention arms received the structured sessions additionally), i.e. the intervention is additive to, not a substitution of, usual PE
Required equipment	low · outdoor track/open space, mats, cones, and a music/sound system
Control condition	CG/CON: usual-curriculum control; continued standard PE curriculum without HIFT/MICT structured intervention.

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Yes · Teacher Delivery: Yes (routine) · Intensity Monitored: Yes · Scalability: Yes · Sustainability: Not assessed · Attendance: Yes
Dropout (%)	IG1: 13.3% · IG2: 11.1% · CG: 20%
Feasibility / scalability (reviewer-inferred)	1. High staffing ratio (2 PE teachers + 2 supervisors, small 4–6 groups); · 2. HR telemetry for all participants + tablet = tech/equipment demand; · 3. Third session is after-school extracurricular → requires non-curricular time, and the extra structured time vs control is an equity/transferability issue
Reported barriers	1. High-intensity formats impose greater physiological stress and require greater supervision → may affect acceptability for students with lower baseline fitness

OUTCOMES & FINDINGS	
Relevant outcomes	Cardiorespiratory fitness (VO_2max) · Mental health/wellbeing (mood - POMS: vigor, self-esteem, anger, depression, fatigue, total mood disturbance) · Body composition (BMI, waist, body-fat %) · Muscular strength/power
Sex / subgroup reporting	Reported by Sex: · No · Reported by Subgroup: No Sex/Subgroup Details and Findings:
Main positive outcome domains	Cardiorespiratory fitness (VO_2max); Muscular strength/power (lower-body power); Body composition (waist circumference); Mental health/wellbeing (mood - POMS)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
HIFT vs control [Single-component arm]	Overall	Cardiorespiratory fitness	Favours intervention	Yes (group×time interaction)	NR (vs-control MD not reported)	ηp^2=0.319	<0.001	HIFT within-group +5.26 [4.24,6.28]; MICT +1.10 [0.09,2.11]; control ns
HIFT vs control [Single-component arm]	Overall	Mental health/wellbeing	Favours intervention	Yes (interaction)	NR	TMD ηp^2=0.303	<0.001	POMS, group×time interaction. Both HIFT and MICT reduced total mood disturbance (TMD) vs control (HIFT p<0.001); HIFT mainly improved positive mood - vigor and self-esteem (p<0.01). Magnitude NR
MICT vs control [Single-component arm]	Overall	Cardiorespiratory fitness	Favours intervention	Yes (interaction)	NR	ηp^2=0.319	<0.001	MICT within-group +1.10 (modest) < HIFT; control ns

MICT vs control [Single-component arm]	Overall	Mental health/wellbeing	Favours intervention	Yes (interaction)	NR	TMD ηp^2=0.303	<0.01	POMS, group×time interaction. Both HIFT and MICT reduced total mood disturbance (TMD) vs control (MICT p<0.01); MICT mainly reduced negative mood - anger, depression, fatigue. Magnitude NR.
HIFT vs control [Single-component arm]	Overall	Muscular strength/power	Favours intervention	Yes (interaction)	NR	SLJ ηp^2=0.086	<0.001	Within-group: SLJ +11.13 cm [7.32,14.94]; push-ups +7.90; handgrip +1.78 kg
MICT vs control [Single-component arm]	Overall	Muscular strength/power	Favours intervention	Yes (interaction)	NR	—	<0.001	Within-group: push-ups +6.33; handgrip +1.34 kg
HIFT vs control [Single-component arm]	Overall	Body composition	Favours intervention	Yes (group×time interaction)	NR	WC ηp^2=0.087	<0.01	Sig interaction for BF%, WC, BMI (all p<0.01). HIFT within-group WC -1.18 cm [-1.83,-0.53], p<0.001 (exceeds MCID ~1.0 cm). Control BF% increased +1.43% [0.70,2.15], p<0.001; both exercise groups maintained stable adiposity. MICT WC change ns (p=0.484).

SUMMARY & APPRAISAL	
Key findings (summary)	Both HIFT and MICT improved cardiorespiratory fitness (VO_2max) and reduced total mood disturbance versus control (HIFT −11.69, MICT −12.78 vs control −6.86). HIFT was superior for cardiorespiratory fitness, lower-body power, waist circumference and positive affect, whereas MICT was superior for negative affect (anger, depression, fatigue); magnitudes in the outcome-level results. Neither improved flexibility; the control group worsened on body-fat % and BMI.

Implications for future research	Conduct larger multi-school trials with longer follow-up; · examine mediators and confounders; · test hybrid-modality programmes and AI-driven personalisation; · emphasise effect magnitude and precision over statistical significance
Study limitations	1. Unmeasured confounders (medication, pain, out-of-school activity, sleep, diet); 2. Short 8-week duration; 3. Pandemic-related disruptions; 4. Small number of clusters; 5. Intervention arms received additional structured time vs control; 6. Single-school recruitment; 7. Student/teacher blinding not feasible (assessors blinded)

Study 7 of 12: Müller et al. (2019)

Effect of a Multidimensional Physical Activity Intervention on Body Mass Index, Skinfolds and Fitness in South African Children: Results from a Cluster-Randomised Controlled Trial — Ivan Müller; Christian Schindler; Larissa Adams; Katharina Endes; Stefanie Gall; Markus Gerber; Nan S. N. Htun; Siphesihle Nqweniso; Nandi Joubert; Nicole Probst-Hensch; Rosa du Randt; Harald Seelig; Danielle Smith; Peter Steinmann; Jürg Utzinger; Peiling Yap; Cheryl Walter; Uwe Pühse

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
South Africa (Africa)	Upper-middle income	Cluster RCT with school-level intervention combinations; PA effect analysed as schools with PA vs schools without PA	School	746 IG1 n = 82 IG2 n = 148 IG3 n = 70 (IG4 n = 97 (non-PA)) CG n = 349	11.2 ± 0.9; 49.9% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
Grade 4; Primary	Public; Urban	Yes; Disadvantaged, low-SES communities	overweight/obesity prevalence reported; PA schools overweight 12.7%, obesity 5.0%; no-PA schools overweight 14.4%, obesity 5.4%	Coloured children, usually Afrikaans-speaking; Black African children, mainly Xhosa-speaking	Partly Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess the effect of a twice-10-week multidimensional, school-based physical activity intervention (DASH) on children's cardiorespiratory fitness and on excess increases in BMI and skinfold thickness in Port Elizabeth, with baseline adjustment for sociodemographic confounders and soil-transmitted helminth status
Study design	Multi-arm

Baseline comparability	Partly; Authors state no significant baseline differences in primary outcomes (obesity, skinfolds, CRF; all p>0.05). Baseline imbalances present, higher in PA schools: stunting 15.3% vs 9.0%; anaemia 22.3% vs 15.9%; STH infection 44.0% vs 23.1%; intestinal protozoa 20.3% vs 13.2%

INTERVENTION	
Component composition	Teacher Training; Additional PE Lessons; Modified PE Curriculum; Motivation / Enjoyment; School Environment / Infrastructure Change; Other: daily in-class activity breaks
Component description	IG1/PA intervention: two 40-min PE lessons/week, one weekly 40-min moving-to-music lesson, daily in-class activity breaks within the main curriculum, and PA-friendly school-environment enhancement through activity stations and painted games. Delivered during official school hours with life-orientation teachers. CG/no-PA schools: usual curriculum without PA intervention.
Dosage	PA intervention · Length: 2 × 10 weeks · Sessions/wk: 3/wk + daily breaks · Total: ~3,300 min (55 h, reported)
Dosage specifications	taught outdoors on grass/cemented areas; coach assisted one of two weekly lessons, teacher delivered the other. Active intervention only in 3 of 8 schools
Required equipment	low · donated sports equipment (bean bags, colour bands, skipping ropes, cones, diverse balls); plus painted games and activity stations for the environment enhancement
Control condition	CG/no-PA schools: usual school curriculum; no multidimensional PA intervention; total PA exposure ≈15 h

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Unclear (no quantitative metric; teacher motivation described as variable) · Teacher Delivery: No / Mixed · Intensity Monitored: No · Scalability: Yes · Sustainability: Not assessed · Attendance: No
Dropout (%)	PA (S1)18.3% · PA+HE (S2)8.1% · PA+HE+NU (S3)12.9% · HE+NU (S4)23.7% · Control (S5–8)48.1%
Feasibility / scalability (reviewer-inferred)	NR

Reported barriers	1. Variable teacher motivation 2. Lack of formal PE qualifications 3. PE not taught regularly in disadvantaged schools 4. Unequal class sizes 5. Paucity/low quality of sport and recreation facilities 6. Logistic constraints (PA in only 3 of 8 schools, two 10-week blocks) 7. Careful local adaptation required before scale-up 8. May not generalise beyond this African context

OUTCOMES & FINDINGS	
Relevant outcomes	Body composition (BMI Z-score, skinfold adiposity) · Cardiorespiratory fitness (20 m shuttle run / VO_2max) · Physical activity (self-reported)
Sex / subgroup reporting	Reported by Sex: · Yes · Reported by Subgroup: Yes Sex/Subgroup Details and Findings: Differential effects on skinfold thickness: Greater benefit in girls, in younger children, and at higher SES; opposite effects in anaemic vs non-anaemic children. For BMI, girls benefited less than boys. Authors suggest gender-specific interventions may be needed and that anaemia should be managed before PA interventions.
Main positive outcome domains	Body composition (BMI-for-age Z-score, skinfold thickness)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
PA intervention vs control [Multi-component arm]	Overall	Body composition	Favours intervention	Yes	BMI Z-score -0.17	-0.24 to -0.09	<0.001	Skinfold -1.06 mm (-1.83 to -0.29), p=0.007

PA intervention vs control [Multi-component arm]	Overall	Cardiorespiratory fitness	No difference	No	20 m SRT: ns	—	>0.05	—

SUMMARY & APPRAISAL	
Key findings (summary)	The intervention produced a smaller increase in BMI-for-age Z-score and skinfold thickness than control (see outcome-level results), with no between-group difference in cardiorespiratory fitness and no effect on self-reported physical activity (which rose more in the control group). Skinfold effects were modified by sex (greater benefit in girls), age (greater in younger children), socioeconomic status (greater at higher SES) and anaemia status.
Implications for future research	Extend the intervention across the full school term/year and to all grades; · deliver a more intensive intervention with greater community, parent and volunteer involvement (potentially requiring policy change); · add extracurricular physical-activity and healthy-nutrition components; · tailor by gender and manage anaemia before physical-activity interventions; · use longitudinal, age-, gender- and grade-specific monitoring; · adapt carefully before scaling up and disseminating across South Africa and sub-Saharan Africa
Study limitations	1. Teacher motivation/qualifications, irregular PE and unequal class sizes affected facilitation; 2. Larger effects may need longer duration plus community/policy involvement; 3. No objective (accelerometer) PA measure; 4. Small sample; PA in only 3 of 8 schools for two 10-week periods; 5. No concealment/blinding; cluster-design bias

Study 8 of 12: Shao & Cao (2025)

Integrating interval jump rope exercise into a school setting improve body composition, cardiorespiratory fitness and glycolipid metabolism parameters in adolescents with overweight and obesity: a randomized controlled trial — Shitong Shao; Meng Cao

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Three-arm single-blind parallel RCT (IG1 = JRE 30:30; IG2 = JRE 60:60; CG = no-training control)	Individual	120 IG1 n = 40 IG2 n = 40 CG n = 40	12.4 ± 0.6; 50% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
NR; Tanner stage 2-3 aged 11-13; Unclear	Unclear; Unclear	No; NR	Obesity-selected sample; Participants had BMI ≥95th percentile for age/sex	NR	Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess (1) the effects of 12 weeks of HIIT-based jump-rope exercise integrated into school PE lessons on body composition and cardiorespiratory fitness in adolescents with excessive body weight, and (2) its effects on glucose homeostasis and blood-lipid profiles
Study design	Multi-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	HIIT; Aerobic Training
Component description	IG1/JRE 30:30: HIIT-based interval jump-rope exercise integrated into the final 15 min of a 40-min PE class, 3 sessions/week for 12 weeks. Each session included 10 min of exercise: 10 × 30 s work / 30 s rest. Intensity progressed from 100 jumps/min in weeks 1–4, to 110 jumps/min in weeks 5–8, to 120 jumps/min in weeks 9–12. IG2/JRE 60:60: same structure, but 5 × 60 s work / 60 s rest. Cadence controlled by rhythmic MP3. PE warm-up was omitted; sessions ended with 3–5 min of stretching. Sessions were monitored by personal trainers. CG: no-training control.
Dosage	JRE-1 (30:30) · Length: 12 weeks · Sessions/wk: 3/wk · Total: 540 min (36 × 15-min block) · JRE-2 (60:60) · Length: 12 weeks · Sessions/wk: 3/wk · Total: 540 min (36 × 15-min block)
Dosage specifications	Total jump-rope work time identical across arms (10 min/session); arms differ only in interval structure (JRE-1 30:30 × 10 bouts; JRE-2 60:60 × 5 bouts). Average HR during exercise (MHR, Table 1): JRE-1 152–157 b.p.m, JRE-2 160–164 b.p.m. Protocol adapted from prior obese-population research
Required equipment	low · jump ropes; MP3 player for cadence pacing
Control condition	CG/CON: no-training control; regular PE and school activity continued; no additional JRE training

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Yes · Teacher Delivery: No · Intensity Monitored: Yes · Scalability: Yes · Sustainability: Not assessed · Attendance: Yes
Dropout (%)	IG1: 15.0% · IG2: 12.5% · CG: 20.0%
Feasibility / scalability (reviewer-inferred)	1. Delivered by external personal trainers/exercise instructors rather than regular PE teachers → uncertain whether typical PE staff could replicate fidelity/intensity
Reported barriers	NR

OUTCOMES & FINDINGS	
Relevant outcomes	Cardiorespiratory fitness (VO_2max) · Body composition (body mass, fat mass, body-fat %, fat-free mass)
Sex / subgroup reporting	Reported by Sex: · No · Reported by Subgroup: No Sex/Subgroup Details and Findings:
Main positive outcome domains	Cardiorespiratory fitness (VO_2max); Body composition (body mass, BMI, fat mass, body-fat %)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
JRE-1 (30:30) vs control [Single-component arm]	Overall	Cardiorespiratory fitness	Favours intervention	Yes	VO_2max +1.85 mL/kg/min	[1.14, 2.57]	<0.001	Between-group MD vs control (Table 4). Within-group +3.4; control also +1.6; JRE-1 vs JRE-2 ns (p=0.234); interaction η^2=0.423.
JRE-1 (30:30) vs control [Single-component arm]	Overall	Body composition	Favours intervention	Yes	BMI -1.67 kg/m²; FM -2.73 kg	BMI [-2.21, -1.12]; FM [-3.90, -1.57]	<0.001	Between-group MD vs control (Table 4); also BM (p<0.001), BF% -1.90% (p=0.024); FFM ns (p=0.053). Control increased BM/FM.
JRE-2 (60:60) vs control [Single-component arm]	Overall	Cardiorespiratory fitness	Favours intervention	Yes	VO_2max +2.37 mL/kg/min	[1.65, 3.08]	<0.001	Between-group MD vs control (Table 4). Within-group +4.0;

								control +1.6; JRE-2 vs JRE-1 ns (p=0.234).
JRE-2 (60:60) vs control [Single-component arm]	Overall	Body composition	Favours intervention	Yes	BMI -2.39 kg/m²; FM -3.39 kg	BMI [-2.93, -1.85]; FM [-4.55, -2.23]	<0.001	Between-group MD vs control (Table 4); also BM -6.19 kg, FFM -2.80 kg (p<0.001), BF% -1.89% (p=0.024). JRE-2 reduced BM/BMI more than JRE-1 (p=0.003/0.004).

SUMMARY & APPRAISAL	
Key findings (summary)	Both jump-rope arms were superior to control for body composition and cardiorespiratory fitness (VO_2max), as well as for blood pressure and resting heart rate; fat-free mass also decreased (an undesirable change). The higher-dose arm (JRE-2, 60:60) reduced body mass and BMI more than JRE-1, but the two doses did not differ on cardiorespiratory outcomes (see outcome-level results).
Implications for future research	Examine factors influencing effectiveness (optimal intensity, duration and frequency); · evaluate long-term effects; · conduct larger multi-school trials; · combine jump-rope exercise with resistance/whole-body strength training to offset fat-free-mass loss; · control and monitor diet, energy intake and sleep; · use accelerometers and diaries; · inform health policy on scalable interventions for childhood obesity
Study limitations	1. Body composition by BIA (InBody770), not DXA; 2. Diet/energy intake recorded but not controlled; 3. Sleep not monitored; 4. Single-school recruitment, possibly insufficient sample

Study 9 of 12: Wan et al. (2025)

In-school resistance training improves physical fitness in pubertal girls: a randomized controlled trial — Jiaqian Wan; Zhe Li; Ruiqing Sun; Jiamin Gu; Yuzhang Li; Qingwen Zhang

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Four-arm parallel RCT (IG1 = traditional strength training; IG2 = plyometric training; IG3 = combined training; CG = control)	Individual	79 IG1 n = 20 IG2 n = 19 IG3 n = 19 CG n = 21	NR; 10-12 years old; 100% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
Grade 6; Primary	Unclear; Urban	No; NR	Healthy BMI range only; overweight/obesity excluded	NR	Partly effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess (1) the effects of school-based resistance training on physical fitness in pubertal girls, and (2) training outcomes across three RT modalities (TST, PT, CT) to inform developmentally appropriate exercise programs for pubertal girls
Study design	Multi-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Modified PE Curriculum; Strength / Resistance Training; Other: Plyometric training
Component description	IG1/TST: traditional strength training delivered during the 10-min physical-fitness segment of 40-min PE classes, 3 sessions/week, supervised by a trained PE teacher with a researcher monitoring intensity using Polar Verity Sense HR monitors. Target intensity was 57–76% HRmax. Exercises targeted lower limb, upper limb, and core. Each programme used 3 exercises × 10 repetitions × 3 sets, 30-s inter-set rest, with progressive overload. TST included half squats, knee-bended push-ups, planks with bodyweight, and progression via elastic resistance bands (10–30 lb) and added load (+5–10% bodyweight every 2 weeks). IG2/PT: plyometric training including box jumps on a 30-cm platform, 20-cm hurdles, standing/declined push-ups, and overhead ball throws. IG3/CT: combined TST + PT elements at half volume, 5 repetitions each. CG: continued regular comprehensive fitness work focused on speed, agility, balance, and standard PE activities.
Dosage	TST / PT / CT (identical dose) · Length: 8 weeks · Sessions/wk: 3/wk · Total: 240 min (24 × 10-min RT segment) Control · Length: 8 weeks · Sessions/wk: 3/wk · Total: - (comprehensive fitness, no RT)
Dosage specifications	8-week intervention; 3 sessions/week during PE classes; RT = the 10-min physical-fitness segment of each 40-min PE lesson; ≥85% session attendance required (participants below the threshold excluded). Progressive overload across weeks (band load, plyometric height/difficulty).
Required equipment	low · largely bodyweight with minimal venue requirements; low-cost elastic bands, plyometric boxes and hurdles
Control condition	CG/CON: regular comprehensive fitness work during the same PE segment; no TST/PT/CT resistance-training intervention

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Yes · Teacher Delivery: Yes (trained) · Intensity Monitored: Yes · Scalability: Yes · Sustainability: Not assessed · Attendance: Yes
Dropout (%)	0% overall

Feasibility / scalability (reviewer-inferred)	1. Requires a trained PE teacher plus a researcher and a HR monitor for real-time intensity (feasibility constraint outside a study); · 2. Modality selection/progressive-overload prescription needs expertise; · 3. Tested only in healthy-weight girls → uncertain transfer to overweight/obese pupils
Reported barriers	NR

OUTCOMES & FINDINGS	
Relevant outcomes	Muscular strength/power (push-ups, sit-ups, standing long jump, handgrip) · Speed/agility (50 m sprint, Illinois change-of-direction) · Cardiorespiratory fitness (20 m shuttle run) · Body composition
Sex / subgroup reporting	Reported by Sex: · No · Reported by Subgroup: No Sex/Subgroup Details and Findings:
Main positive outcome domains	Muscular strength/power - push-ups (significant vs control across TST/PT/CT); Other musculoskeletal gains (1-min sit-ups, standing long jump, hand grip) and speed/agility (50 m sprint, Illinois CODT) improved within-group, but between-group (ANCOVA) significance was established only for push-ups; No between-group effect on cardiorespiratory fitness or body composition

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
TST vs control [Single-component arm]	Girls	Muscular strength/power	Favours intervention	Yes	Push-ups greater than control	—	—	Between-group: push-ups 'remarkable differences vs CON' (narrative). Within-group d: push-ups 1.11, sit-ups 1.65, SLJ 0.66 (grip ns)

TST vs control [Single-component arm]	Girls	Speed/agility	Favours intervention	No	50 m sprint	—	—	Within-group d=0.31 (CI crosses 0); between-group ns
PT vs control [Single-component arm]	Girls	Muscular strength/power	Favours intervention	Yes	Push-ups greater than control	—	—	Between-group: push-ups vs CON (narrative). Within-group d: push-ups 0.43, sit-ups 0.60, SLJ 0.67, grip 0.36
PT vs control [Single-component arm]	Girls	Speed/agility	Favours intervention	No	50 m sprint, Illinois CODT	—	—	Within-group d: sprint 0.40, Illinois 0.33; between-group ns
CT vs control [Multi-component arm]	Girls	Muscular strength/power	Favours intervention	Yes	Push-ups greater than control	—	—	Between-group: push-ups vs CON (narrative). Within-group d: push-ups 0.32, sit-ups 0.89, SLJ 0.69, grip 0.51
CT vs control [Multi-component arm]	Girls	Speed/agility	Favours intervention	No	50 m sprint, Illinois CODT	—	—	Within-group d: sprint 0.61, Illinois 0.63; between-group ns
CT vs control [Multi-component arm]	Girls	Cardiorespiratory fitness	Favours intervention	No	20 m shuttle run	—	—	Within-group gain $p<0.001$, $d=0.70$; between-group explicitly ns
All arms vs control [Mixed-component arm]	Girls	Body composition	No difference	No	ns	—	—	Muscle-mass↑/fat-mass↓ trends, not significant

SUMMARY & APPRAISAL	
Key findings (summary)	All three resistance-training arms showed within-group gains in muscular strength/power and speed/agility, and the control group declined on push-ups; the clearest between-group difference favouring the intervention was in push-ups (see outcome-level results). Body composition was unchanged. The combined arm (CT) produced the broadest improvements, the total-body arm (TST) the strongest abdominal/upper-limb gains, and the plyometric arm (PT) the strongest jump/power gains.
Implications for future research	Use the 10-minute resistance-training programme as an economical, time-efficient addition to PE; · match the resistance-training modality to participants' characteristics and fitness weaknesses; · consider a combined (TST+PT) programme for more comprehensive fitness gains; · allow longer interventions or integrate aerobic exercise to change body composition; · encourage girls to continue strength and conditioning into adulthood
Study limitations	1. School-setting variability; lifestyle factors (nutrition, growth, sleep) uncontrolled; 2. Only healthy-weight girls, small per-group sample → limited generalisability to boys and overweight/obese; 3. Mechanisms (neuromuscular/endocrine) and psychological dimensions omitted; 4. 10-min sessions over 8 weeks, no long-term follow-up

Study 10 of 12: Zhang et al. (2024)

Exploring the effects of a 12-Week functional training program on fundamental motor skills for primary school children aged 6–7 — Dong Zhang; Soh Kim Geok; Yoke Mun Chan; Zeinab Zaremohzzabieh; Soh Kim Lam; Shanshan He

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Two-arm parallel RCT (IG1 = functional training; CG = regular PE)	Individual	100 IG1 n = 50 CG n = 50	NR; IG1 7.12 ± 0.29 CG 7.10 ± 0.27; 50% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
First-grade students aged 6–7; Primary	Unclear; Unclear	No; NR	NR	NR	Partly Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess the effect of a 12-week functional training program on fundamental movement skills in primary-school children aged 6–7
Study design	Single-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Modified PE Curriculum; Motor Competence / FMS; Strength / Resistance Training; Other: Balance & agility
Component description	IG1/FT: functional training group. Sessions included a 5-min warm-up with jogging, stretching, and games, followed by 30 min of functional training. Content included bodyweight exercises without equipment, fitness-step training, medicine-ball training,

	balance-mat training, and bodyweight exercises using agility circles, rope ladders, exercise balls, and balance boards. Designed to align with school PE curriculum guidelines and the developmental characteristics of children aged 6–7. CG: regular PE.
Dosage	FT · Length: 12 weeks · Sessions/wk: 2/wk · Total: 960 min (24 × 40 min)
Dosage specifications	Session = 5-min warm-up + 30-min functional-training core + 5-min relaxation. Intensity low-to-moderate, 50–70% HRmax (FT core 60–70% HRmax per Table 2), regulated via wrist-worn heart-rate monitors and the Borg RPE Scale. IG1 and CG ran concurrently in the same slots (first class Wednesday afternoon, third class Friday morning). FT replaced the regular PE slot rather than adding time → no intervention-vs-control dose-time difference
Required equipment	moderate · agility circles/ladders, rope ladders, exercise balls, balance boards, fitness steps, medicine balls and balance mats
Control condition	CG/PE: regular physical education curriculum, 2 × 40 min/week for 12 weeks

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Yes · Teacher Delivery: Yes (trained) · Intensity Monitored: Yes · Scalability: Yes · Sustainability: Not assessed · Attendance: Yes
Dropout (%)	IG: 8% · CG: 6%
Feasibility / scalability (reviewer-inferred)	1. Delivery requires a trained teacher + administrator per group and per-child wrist HR monitors → feasibility constraint outside a study context
Reported barriers	NR

OUTCOMES & FINDINGS	
Relevant outcomes	Fundamental movement skills (gross motor quotient, locomotor and object-control subtests TGMD-2, dynamic balance (Pediatric Balance Scale))
Sex / subgroup reporting	Reported by Sex: · No · Reported by Subgroup: No Sex/Subgroup Details and Findings:

Main positive outcome domains	Fundamental movement skills (gross motor quotient, object-control subtest)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
FT vs PE [Single-component arm]	Overall	Fundamental movement skills	Favours intervention (test-dependent)	Yes (group main effect: GMQ & object control)	GMQ & object control improved (TGMD standard scores)	—	GMQ 0.003; object control 0.006	Show-both: time×group interaction sig all four (p<0.001); group main effect sig only GMQ & object control (locomotor p=0.110, balance p=0.276 ns); pairwise FT>PE at 12 wk on all four. GMQ d≈1.09. Between-group MD NR

SUMMARY & APPRAISAL

Key findings (summary)	Time-by-group interactions favoured functional training (FT) on all four fundamental-movement-skill outcomes, but the group main effect was significant only for the gross motor quotient and object control; pairwise, FT exceeded PE at both 6 and 12 weeks for these two outcomes, and at 12 weeks only for locomotor and balance. The largest between-group effect was the gross motor quotient (see outcome-level results).
Implications for future research	Adopt more rigorous measures to monitor and control external influences (e.g. children's extracurricular activity); · design long-term curriculum plans adjusted to students' characteristics across different schools when integrating functional training into school PE

Study limitations	1. External/extracurricular training/sports could not be fully controlled over 12 weeks → may affect FMS and accuracy (despite parental consent/instruction)

Study 11 of 12: Zhou et al. (2019)

Impact on Physical Fitness of the Chinese CHAMPS: A Clustered Randomized Controlled Trial — Zhixiong Zhou; Shiyu Li; Jun Yin; Quan Fu; Hong Ren; Tao Jin; Jiahua Zhu; Jeffrey Howard; Tianwen Lan; Zenong Yin

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Four-arm 2×2 factorial cluster RCT (IG1 = SPE; IG2 = ASP; IG3 = SPE + ASP; CG = control)	School	753 IG1 n= 200 IG2 n = 199 IG3 n = 178 CG n = 176	12.66 ± 0.56; 47% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
Grade 7; Lower secondary	Unclear; Unclear	No; Parent education and monthly family income reported; broad range across junior-high to postgraduate parental education and RMB ≤5000 to ≥15,000 family income. Sample not characterised as low-resource or disadvantaged.	NR	NR	Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess (1) whether incremental modification of school PE policy and curriculum, and the addition of an after-school PA program, influences change in cardiorespiratory fitness (primary) and other physical-fitness outcomes in middle-school students - testing an incremental dose-response (SPE, ASP and SPE+ASP each > control; SPE and SPE+ASP > ASP; SPE+ASP > SPE), and (2) more favourable change in other fitness measures in intervention vs control

Study design	Multi-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Teacher Training; Additional PE Lessons; Modified PE Curriculum; Aerobic Training; Motivation / Enjoyment; School Environment / Infrastructure Change; Other: parent mHealth engagement
Component description	Socio-ecological multi-level programme over 8 months. IG1/SPE: PE policy change to 3 PE classes/week plus daily 15-min PA-based recess; redesigned game-based, high-intensity physical function training curriculum; 45-min PE structure of 5-min warm-up, 12-min instruction, 25-min fitness activities, and 3-min cool-down; 2-day teacher training; portable equipment; fitness and nutrition handbook; bi-weekly WeChat health tips to students; parent WeChat campaign. IG2/ASP: unchanged PE policy plus mandatory extracurricular after-school PA programme, 45 min/session with 5-min warm-up, 35-min fitness activities, and 5-min cool-down using physical function training exercises; 1-day teacher training; portable equipment; handbook; WeChat tips; parent campaign. IG3/SPE+ASP: combined SPE and ASP with 3-day teacher training. CG: usual PE programme, 2 × 45 min/week, no teacher training and no curriculum change.
Dosage	SPE · Length: 8 months · Sessions/wk: 3 PE + 5 recess/wk · Total: 210 min/wk MVPA (reported) · ASP · Length: 8 months · Sessions/wk: 2 PE + 2 after-school/wk · Total: 180 min/wk MVPA (reported) · SPE+ASP · Length: 8 months · Sessions/wk: 3 PE + 5 recess + 2 ASP/wk · Total: 300 min/wk MVPA (reported) Control · Length: 8 months · Sessions/wk: 2 PE/wk · Total: 90 min/wk MVPA (reported)
Dosage specifications	Targets are weekly MVPA/VPA, not a single cumulative total. PE class internal structure 5/12/25/3 min; ASP 5/35/5 min; daily recess 15 min × 5 d/wk (SPE arms). Whole design confined to on-campus hours (8:00 a.m.–5:00 p.m.). Process data (Table 6, posttest): ≥140 bpm of PE-class time reached ~61% in both SPE (60.7%) and SPE+ASP (61.8%); ≥150 bpm reached ~26% only in SPE+ASP (25.6%), vs ~6% in SPE (5.7%). Mean PE-class HR rose to ~141 bpm (SPE) and ~142 bpm (SPE+ASP) vs ~106 bpm in control.
Required equipment	low · "small portable exercise equipment" (not further specified)
Control condition	CG: usual PE programme, 2 × 45 min/week; no teacher training, no after-school PA programme, and no curriculum change

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: Yes · Teacher Delivery: Yes (trained) · Intensity Monitored: Yes · Scalability: Yes · Sustainability: Not assessed · Attendance: No
Dropout (%)	IG1: 18.5% · IG2: 9.5% · IG3: 5.6% · CG: 3.4%
Feasibility / scalability (reviewer-inferred)	1. Differential dose vs control - IG2/IG3 add structured after-school PA time the control never receives (equity/transferability issue); · 2. Teacher training required (1–3 days); · 3. After-school programme requires staff time at day's end
Reported barriers	1. Authors note changing daily school health routines involves "many challenges and barriers" 2. PE-policy change (3 classes/wk + daily recess) requires institutional/administrative buy-in - authors note the need to institutionalise PE/recess time

OUTCOMES & FINDINGS	
Relevant outcomes	Cardiorespiratory/physical fitness (CRF / 20 m shuttle run, lower-limb power, speed, agility, flexibility, abdominal/core endurance) · Body composition (BMI) · Physical activity behaviour
Sex / subgroup reporting	Reported by Sex: · No · Reported by Subgroup: No Sex/Subgroup Details and Findings:
Main positive outcome domains	Cardiorespiratory/physical fitness (CRF, lower-limb power, speed, agility, flexibility, core endurance); Body composition (reduced body fat - all arms vs control)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
All interventions pooled vs control [Mixed-component arm]	Overall	Physical/cardiorespiratory fitness	Favours intervention	Yes	20 m SRT +15.2 laps; broad jump +17.0 cm; plank +31.8 s; sit-ups	—	<0.001	Covariate-adjusted planned contrast (Hypothesis 1)

					+5.1; sit-reach +3.5 cm; 50 m -0.4 s; agility -1.0 s; body fat -1.6%			
SPE-containing (SPE+SPE+ASP) vs ASP-only [Mixed-component arm]	Overall	Physical/cardiorespiratory fitness	Favours intervention	Yes	20 m SRT +9.4 laps; broad jump +9.7 cm; plank +37.2 s	—	<0.001	Incremental SPE effect (Hypothesis 2)
SPE+ASP incremental vs ASP [Multi-component arm]	Overall	Physical/cardiorespiratory fitness	Favours intervention	Yes	20 m SRT +13.6 laps; broad jump +15.0 cm; plank +48.6 s; body fat -2.4%	—	<0.001	Largest contrast - dose-response (Hypothesis 3)

SUMMARY & APPRAISAL	
Key findings (summary)	Results followed an incremental dose-response: improvement was greatest in the combined arm (SPE+ASP), then SPE, then ASP, each exceeding control, with SPE exceeding ASP; all three planned orthogonal contrasts were significant for every fitness measure after covariate adjustment, including body fat (see outcome-level results). Process data indicated high fidelity, with accelerometer-measured MPA/VPA/MVPA roughly two- to three-fold higher in intervention arms and PE-class heart-rate targets met in the SPE arms (details under dosage specifications).
Implications for future research	Replicate while examining the influence of student motivation and parental engagement on fitness improvement; · assess the acceptability and feasibility of the parent-engagement (mHealth) component; · examine the physical-activity–fitness dose-effect relationship; · test sustainability beyond eight months; · evaluate effectiveness using an implementation-science framework
Study limitations	1. No follow-up on sustainability beyond 8 months; 2. Accelerometry from only part of the sample → PA–fitness dose-effect not examinable; 3. Attendance data not collected; 4. Mediators (motivation/enjoyment) not collected; 5. Weather-day health/nutrition education not evaluated; 6. WeChat parent-engagement feasibility/effectiveness not assessed

Study 12 of 12: Zou et al. (2025)

Effects of a 12-week core strength training program on physical fitness and fundamental movement skills in children aged 6–8 years: A randomized controlled trial — Xueliang Zou; Wichai Eungpinichpong; Kurusart Konharn

Country / region	Income group	Design	Randomisation level	Total sample (n by group)	Mean age (±SD) / sex
China (Asia)	Upper-middle income	Two-arm parallel RCT (IG1 = core strength training; CG = standard PE)	Individual	128 IG1 n = 64 CG n = 64	7.01 ± 0.82; 48.4% female
Year / educational level	**School type / setting**	**Low-resource setting**	**Baseline weight status**	**Ethnicity / race**	**Overall effectiveness**
NR; primary school students aged 6–8; Primary	Unclear; Urban	No; NR	NR	NR	Effective

STUDY AIM, DESIGN & BASELINE	
Aim of study	To assess (1) the effects of a 12-week core strength training program on physical fitness and fundamental movement skills in children aged 6–8, and (2) the correlations between core strength and physical-fitness/movement-ability indicators, to support PA-promoting training programs for children
Study design	Single-arm
Baseline comparability	Yes

INTERVENTION	
Component composition	Motor Competence / FMS; Strength / Resistance Training; Aerobic Training; Other: core strength training
Component description	IG1/CST: 12-week core strength training programme. Progressive programme: weeks 1–2 used jumping and bridge bodyweight patterns with 2 sets of 8–10 repetitions and 60-s intervals; weeks 5–8 included reverse curls, side plank, push-ups and abdominal jumps, and Swiss-ball bridge exercises; weeks 9–12 used predominantly Swiss-ball exercises such as incline plank, dynamic hip bridge, supine leg curl, and related trunk-control exercises. Exercises included abdominal bracing, quadruped abdominal contractions, lateral/supine bridges, and lateral trunk rotation on a Swiss ball. Each session included a 5-min warm-up with jogging and dynamic stretching, 30 min of core training, and a 5-min static-stretch cool-down. CG: standard/regular PE.
Dosage	CST · Length: 12 weeks · Sessions/wk: 3/wk · Total: 1,440 min (36 × 40 min)
Dosage specifications	Session structure 5 + 30 + 5 min; 30-min active core block; 2 sets throughout, reps/duration and rest intervals progressed across the six 2-week blocks (8–10 reps/60 s → 12 reps or 40–60 s holds/45 s). CG matched on session count and length (3×/wk × 40 min) per Table 2.
Required equipment	low · Swiss balls and mats
Control condition	CG: standard/regular PE, 3 × 40 min/week for 12 weeks

IMPLEMENTATION & FEASIBILITY	
Fidelity & sustainability	Adherence: No · Teacher Delivery: Unclear · Intensity Monitored: No · Scalability: Yes · Sustainability: Not assessed · Attendance: No
Dropout (%)	0% overall
Feasibility / scalability (reviewer-inferred)	1. Per-participant Swiss balls required; · 2. Delivery agent unspecified, and CST progression requires a competent instructor; · 3. Post-test-only/no-ANCOVA → baseline imbalance may confound (methodological critique, not feasibility)
Reported barriers	NR

OUTCOMES & FINDINGS	
Relevant outcomes	Physical fitness (speed - 50 m sprint; muscular/core endurance - plank; back strength; flexibility - sit-and-reach); · Fundamental movement skills (lower-limb power - single-leg jump; coordination - side-to-side jump; dynamic balance - backward balance walking)
Sex / subgroup reporting	Reported by Sex: · No · Reported by Subgroup: No Sex/Subgroup Details and Findings:
Main positive outcome domains	Cardiorespiratory/endurance (jump rope); Muscular strength/power (plank, back strength, sit-ups); Speed/agility (50 m sprint, lateral movement); Fundamental movement skills (single-leg jump - lower-limb power; side-to-side jump - coordination); Balance – static (one-legged stand)

OUTCOME-LEVEL RESULTS *(from outcome extraction; grey rows = mediator analyses, excluded from outcome vote-count)*

Comparison	Subgroup	Domain	Direction	Sig. (between-group)	Effect (with unit)	95% CI / statistic	p-value	Notes
Experimental vs control [Single-component arm]	Overall	Speed/agility	Favours intervention	Yes	50 m sprint (d=0.37); lateral movement (d=-0.40)	—	sprint .01; lateral .04	Between-group d (Table 5). Within-group sprint change -0.57 s
Experimental vs control [Single-component arm]	Overall	Cardiorespiratory/endurance	Favours intervention	Yes	1-min jump rope (d=-0.43)	—	.009	Between-group (Table 5);
Experimental vs control [Single-component arm]	Overall	Muscular strength/power	Favours intervention	Yes	plank (d=-0.73); back strength (d=-0.60); sit-ups (d=-0.37)	—	plank <.01; back .005; sit-ups .04	Between-group (Table 5). Within-group: plank +26.92 s; back strength +5.28 kg

Experimental vs control [Single-component arm]	Overall	Fundamental movement skills	Favours intervention	Yes	single-leg jump (d=-0.51); side-to-side jump (d=-0.44)	—	.006 / .008	Between-group (Table 5). Within-group: single-leg jump +5.86 s; side-to-side +10.75 reps
Experimental vs control [Single-component arm]	Overall	Balance (static)	Favours intervention	Yes	one-legged stand (d=-0.39)	—	.04	Between-group (Table 5). Dynamic balance (backward walking p=.07) and flexibility (sit-and-reach p=.15) remain ns

SUMMARY & APPRAISAL	
Key findings (summary)	The experimental group significantly outperformed control across almost the whole fitness battery — sprint, jump rope, lateral movement, single-leg and side-to-side jumps, sit-ups, plank, back strength and static balance — with only sit-and-reach flexibility and dynamic balance (backward walking) non-significant; all significant effects favoured the intervention (see outcome-level results). Core endurance correlated with back strength (r=0.412) and inversely with sprint time (r=−0.477).
Implications for future research	Use a longitudinal design with longer follow-up; · recruit a larger, more diverse sample; · investigate psychological and external factors (e.g. nutrition, sleep, other activities) affecting children's fitness
Study limitations	1. 12-week duration may be insufficient for long-term effects; 2. Single-school → limited generalisability; 3. Confounders (nutrition, sleep, other activities) not accounted for; 4. Psychological benefits may be overlooked

Verwendete Hilfsmittel

Abschnitt	Hilfsmittel	Einsatzform	Anmerkung
Kapitel 1	ChatGPT-5.4 (OpenAI)	Verbesserung von Textstrukturen und Formulierungen	
Kapitel 2	ChatGPT-5.4 (OpenAI)	Verbesserung von Formulierungen, Hilfestellung bei der Formatierung APA-konformer Tabellen	
Kapitel 3	ChatGPT-5.4 (OpenAI)	Hilfestellung bei der Formatierung APA-konformer Tabellen	
Kapitel 4	ChatGPT-5.4 (OpenAI)	Verbesserung von Textstrukturen und Formulierungen	
Kapitel 5	ChatGPT-5.4 (OpenAI)	Verbesserung von Textstrukturen und Formulierungen	
Kapitel 6	Zotero	Erstellung des Literaturverzeichnisses gemäss Richtlinien	
Anhang	Claude Opus 4.8 (Anthropic)	Hilfestellung bei der Überführung der Roh-Extraktionstabelle (ca. 40 Zeilen) in eine präsentable Form	
Gesamte Arbeit	ChatGPT-5.4 (OpenAI), Claude Opus 4.8 (Anthropic)	Korrektur der Rechtschreibung	

Selbstständigkeitserklärung

Ich versichere hiermit, dass ich die vorgelegte Arbeit selbstständig angefertigt und keine anderen als die in der Arbeit angegebenen Hilfsmittel benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Darüber hinaus bestätige ich, dass die vorgelegte Arbeit nicht an einer anderen Hochschule als Seminar-, Projekt- oder Abschlussarbeit oder als Teil solcher Arbeiten eingereicht wurde.

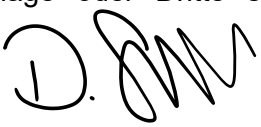
Des Weiteren versichere ich, sämtliche Textpassagen, die unter Zuhilfenahme KI-gestützter Programme verfasst wurden, entsprechend gekennzeichnet sowie mit einem Hinweis auf das verwendete KI-gestützte Programm versehen zu haben.

Eine Überprüfung der Arbeit auf Plagiate und KI-gestützte Programme – unter Einsatz entsprechender Software – darf vorgenommen werden. Bei begründetem Verdacht auf eine unerlaubte oder nicht gekennzeichnete Anwendung von KI bei schriftlichen Leistungsüberprüfungen ist die Kandidatin bzw. der Kandidat auf Einladung hin verpflichtet, an der Klärung des Verdachts mitzuwirken, z.B. durch Teilnahme an einem Gespräch. Ich habe zur Kenntnis genommen, dass unlauteres Verhalten zu einer Bewertung der betroffenen Arbeit mit einer Note 1 oder mit «nicht bestanden» bzw. «fail» oder zum Ausschluss vom Studium führen kann (gemäss §25 der Ordnung für das Masterstudium «Sport, Bewegung und Gesundheit» an der Medizinischen Fakultät der Universität Basel vom 19. Dezember 2016).

Datum: 15.06.2026 Unterschrift: 

Autorenrechte

Hiermit bestätige ich, dass die Publikation der vorliegenden Arbeit oder Teile des Inhalts – auch in Auszügen beziehungsweise als Zusammenfassungen oder in Rohdatenform – sowie die Abgabe der Autorenrechte (auch unentgeltlich) an Verlage oder Dritte stets der Einwilligung des Gutachters bedarf.

Datum: 15.06.2026 Unterschrift: 

Department of Sport, Exercise and Health

Effectiveness of QPE-Aligned School-Based Physical Activity Interventions in Low- and Middle-Income Countries:

A Systematic Comparative Review of Randomised Controlled Trials

David Suter
Supervisor: Dr. Ivan Müller

1 | Introduction

Physical inactivity is among the greatest global health challenges: worldwide, 81% of adolescents fail to meet WHO physical-activity recommendations, raising the risk of non-communicable diseases. Schools reach children and adolescents comprehensively and are therefore a key lever. With the *Quality Physical Education* (QPE) framework (2015), UNESCO promotes inclusive, pedagogically robust physical education. Numerous school-based RCTs have examined QPE-aligned interventions, yet findings are heterogeneous and inconsistent across countries. To date, no systematic synthesis has compared their effectiveness across national contexts. This review therefore examines QPE-aligned school-based RCTs in low- and middle-income countries.

2 | Method & Data

This systematic comparative review followed PRISMA 2020. PubMed, SPORTDiscus, and Web of Science were searched for school-based randomised controlled trials (RCTs) published from 2015 onward in low- and middle-income economies. Records were screened in Covidence; eligible studies were QPE-aligned, in-person PE or school physical-activity interventions with a control arm and at least one relevant outcome. Eligible populations were school-aged children and adolescents. Records were screened and data extracted in Covidence using a structured form, then synthesised narratively, as heterogeneity in designs, outcomes, and reporting precluded meta-analysis. Each study was coded as effective, partly effective, or not effective, based primarily on between-group evidence.

3 | Results

Twelve RCTs (4'999 randomised pupils in total) from three upper-middle-income countries met all criteria: China (n=8), Brazil (n=3), South Africa (n=1). Six studies were coded as effective and six as partly effective; none was ineffective, though effects were often limited to specific arms, subgroups, or outcomes.

Physical fitness showed the clearest and most consistent improvements, especially cardiorespiratory fitness and muscular strength/power. Effects on body composition, motor competence, cognition, and mental health were more dependent on intervention type and subgroup. Physical activity and sedentary behaviour were least consistent and assessed in few studies.

Because all studies came from three upper-middle-income countries, cross-country comparison (RQ1) is descriptive only; no income-group or causal national effects can be inferred.

Structured exercise components such as resistance, functional, high-intensity, and jump-rope training were most consistently linked to fitness gains, and effects were strongest where the intervention directly targeted the measured outcome. Teacher-focused components, rather than additional PE time alone, were associated with the clearest behavioural and psychosocial findings.

Table 1
Outcome domains and the components most linked to them

Outcome Domain	Pattern Actoss Studies	Most linked components
Physical / cardiorespiratory fitness	Most consistent improvements	Resistance, functional, HIIT, jump-rode, core
Movement / motor skills	Consisten where directly targeted	Functional training, FMS programmes
Body composition	Frequent positive effects	Fitness-/adiposity-focused intervention
Psychosocial / behavioural	Clearest in teacher-focused arms	Teacher training workshops
Physical Activity	Limited: assessed in only 2 studies	Added PE time alone (weak)

Note. Descriptive associations, not causal effects. HIIT = high-intensity interval training; FMS = fundamental movement skills.

Table 2
Included studies by country, focus, and main outcomes

Country (n)	Intervention Focus	Main positive outcomes
China (8)	Structure exercise / fitness training	Fitness, strength/power, body composition
Brazil (3)	Teacher-focused, additional PE, mHealth	Sedentary behaviour, mental health
South Africa (1)	Multicomponent school-basel	Body composition

Note. All three countries are upper-middle-income; patterns are descriptive only and do not support causal or national-level comparisons.

4 | Discussion and Conclusion

Effects were not uniform but varied with the fit between intervention content, target outcome, and delivery conditions. Structured exercise components (resistance, functional, high-intensity, jump-rope, and core training) were most consistently linked to fitness gains, and effects were strongest where the intervention directly targeted the measured outcome (Love et al., 2019). Teacher-focused components, rather than additional PE time, produced the clearest behavioural and psychosocial findings — consistent with QPE's emphasis on teacher education and pedagogical quality (UNESCO, 2015).

- QPE-aligned school-based RCTs can improve fitness, movement, and selected psychosocial outcomes.
- Pedagogical quality appears more decisive than the amount of PE time.
- The evidence base is narrow and uneven: all trials came from three upper-middle-income countries, so cross-country conclusions remain descriptive.
- Trials in low- and lower-middle-income countries, systematic fidelity reporting, and post-intervention follow-up are needed

Literature

Love, R., Adams, J., & van Sluijs, E. M. F. (2019). Are school-based physical activity interventions effective and equitable? A meta-analysis of cluster randomized controlled trials with accelerometer-assessed activity. *Obesity Reviews : An Official Journal of the International Association for the Study of Obesity*, *20*(6), 859–870. https://doi.org/10.1111/obr.12823
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*, n71. https://doi.org/10.1136/bmj.n71
UNESCO. (2015). *Quality physical education: Policy guidelines: Methodology*. https://unesdoc.unesco.org/ark:/48223/pf0000231101
Veritas Health Innovation. (2026). *Covidence* [Computer software]. https://www.covidence.org/
WHO Guidelines on Physical Activity and Sedentary Behaviour: At a Glance (1st ed). (2020). World Health Organization.

David Suter | DSBG | University of Basel | Delsbergerallee 45 | CH-4053 Basel | daviid.suter@stud.unibas.ch | www.dsbg.unibas.ch