# Improving physical self-perception and health-related fitness through gamification in physical education: A cluster randomized controlled trial

Ferdinando Cereda[1*,A-F]

[1] The Catholic University of the Sacred Heart in Milan, Italy

***Corresponding author:** Ferdinando Cereda; The Catholic University of the Sacred Heart in Milan, Largo Fra Agostino Gemelli, 1, 20123 Milano MI, Italy; email: ferdinando.cereda@unicatt.it

**Received:** 2026-01-08
**Accepted:** 2026-03-02
**Published:** 2026-04-01

A – Research concept and design
B – Collection and/or assembly of data
C – Data analysis and interpretation
D – Writing the article
E – Critical revision of the article
F – Final approval of article



## ABSTRACT

**Background:** The global decline in adolescent physical activity necessitates pedagogical innovation to counter disengagement. Traditional norm-referenced physical education (PE) often alienates less skilled students, whereas gamification offers a promising criterion-referenced strategy to enhance motivation and learning.
**Objectives:** This study investigated the efficacy of a 12-week gamified didactic intervention on perceived physical self-concept, health-related fitness, and physical activity enjoyment among high school students.
**Methods:** A two-arm, parallel-group, cluster randomized controlled trial (Cluster-RCT) was conducted with 210 Italian adolescents (mean age 15.4 ± 0.9 years) nested within 10 intact classes (average cluster size = 21). Intact classes were randomised to an experimental group (gamified 'MoveQuest' module, k = 5) or a control group (traditional direct instruction, k = 5). Data were collected using the Revised Physical Self-Perception Profile (PSPP-R-IT), the ALPHA-FIT test battery, and the Physical Activity Enjoyment Scale (PACES). Following a per-protocol analytical approach, Linear mixed models (LMMs) were employed to analyse the data, accounting for the nested structure of the educational setting and handling cluster-level variance intraclass correlation coefficients (ICCs). Significance levels were interpreted considering the multiplicity of outcomes.
**Results:** Significant group × time interactions were observed, with the gamified group demonstrating superior improvements in sports competence ($d = 0.65$), physical self-worth ($d = 0.58$), and enjoyment ($d = 0.82$) compared to the control group. Physiologically, the experimental group showed significantly greater gains in cardiorespiratory fitness ($\dot{V}O_2$ max) and agility, while body attractiveness remained stable in both groups.
**Conclusions:** Gamified didactics significantly enhance psychological engagement and specific motor performance parameters by prioritising individual progress over social comparison. This approach represents a valuable pedagogical tool for fostering physical literacy in secondary PE.

**Keywords:** adolescents, cluster-RCT, gamification, health-related fitness, physical education, physical literacy, physical self-perception, self-determination theory.

**How to cite this article:** Cereda, F. (2026). Improving physical self-perception and health-related fitness through gamification in physical education: A cluster randomized controlled trial. *Physical Education and Sports: Studies and Research*, *5*(1), 7-27. https://doi.org/10.56003/pessr.v5i1.677

## INTRODUCTION

Regular engagement in physical activity is widely recognised as a fundamental determinant of physical, mental, and social health throughout the lifespan (World Health Organization [WHO], 2020). Despite the documented benefits, global surveillance data indicate a pervasive decline in physical activity levels, particularly during the transition from childhood to adolescence (Guthold et al., 2020). Recent evidence suggests that approximately 80% of adolescents worldwide do not meet the recommended guidelines of 60 minutes of moderate-to-vigorous physical activity daily, a phenomenon often described as a pandemic of physical inactivity (Bull et al., 2020; Piotrowski et al., 2025). This decline is not merely a behavioural issue but is deeply rooted in the developmental changes occurring during adolescence, a critical period characterised by rapid biological maturation and significant psychosocial shifts (Burton et al., 2023).

Central to understanding this disengagement is the construct of motor competence (MC), defined as an individual's proficiency in executing a wide range of motor skills (Robinson et al., 2015). According to the conceptual model proposed by Stodden et al. (2008), the relationship between MC and physical activity is reciprocal and dynamic. In early childhood, physical activity drives the development of motor skills; however, as children transition into adolescence, higher levels of motor competence become a prerequisite for participation in sports and active pursuits. Adolescents with low MC often experience a "proficiency barrier" (Seefeldt, 1980), leading to frustration, reduced motivation, and eventual withdrawal from physical activity. This negative spiral is exacerbated by the phenomenon of "adolescent awkwardness," where rapid growth spurts can temporarily disrupt motor coordination, further impacting performance and confidence (Quatman-Yates et al., 2012; Burton et al., 2023).

However, competence alone does not dictate behaviour; the individual's *perception* of that competence is equally decisive. The Physical Self-Perception Profile (PSPP) model posits that global self-esteem is hierarchically organised, with physical self-worth being influenced by specific subdomains such as sports competence, physical conditioning, body attractiveness, and physical strength (Fox & Corbin, 1989; Nicolosi et al., 2024). During adolescence, the gap between actual and perceived competence often widens. As cognitive capabilities mature, adolescents become more accurate in assessing their abilities relative to peers, but they also become more susceptible to social comparison and body image concerns (Harter, 1999; Bourke et al., 2025). Furthermore, recent validation studies in the Italian context have highlighted that the impact of perceived competence on self-esteem is moderated by the subjective importance assigned to specific domains (James's hypothesis), a variable frequently overlooked in intervention research (Nicolosi et al., 2024).

This interplay between competence, confidence, and motivation constitutes the core of physical literacy (PL), a holistic framework that is increasingly guiding physical education (PE) curricula globally (Whitehead, 2010; Iannaccone et al., 2025). Physical Literacy is defined not merely as the ability to move, but as the motivation, confidence, physical competence, knowledge, and understanding to value and take responsibility for engagement in physical activities for life (Edwards et al., 2017). Consequently, the primary goal of secondary school PE should be to foster this literacy. Moving beyond traditional frameworks, this aligns with the call for a modern pedagogical paradigm (Blain et al., 2022; Grauduszus et al., 2024; Cereda, 2023a). Viewed through a contemporary interdisciplinary lens (Cereda, 2026), this paradigm dictates that motor skills are perceived not as terminal achievements but as conduits for broader erudition and holistic student maturation. Yet, traditional pedagogical approaches, often characterised by direct instruction, repetitive drills, and norm-referenced assessments (comparing students against standardized norms), have been criticised for failing to engage the "digital native" generation (Kirk, 2013). Indeed, the structural weaknesses of this traditional technique-centered paradigm—specifically its tendency to exclude less-skilled students and offer insufficient learning time—have been highlighted as a primary barrier to effective physical education (Barrientos Hernán et al., 2023; Slingerland et al., 2024; Cereda, 2023b). Norm-referenced evaluations, in particular, can be detrimental to the motivation of less skilled students, fostering anxiety rather than mastery (Yang et al., 2024; Harte et al., 2024). Consequently, there is a strong pedagogical call to shift towards criterion-referenced standards that provide non-judgmental, health-focused feedback (Welk et al., 2022). Indeed, recent qualitative inquiry confirms that traditional fitness testing practices often evoke feelings of embarrassment and confusion about their purpose, potentially rendering the experience 'mis-educative' for students (Alfrey, 2024).

To address these pedagogical shortcomings, Gamification—the application of game-design elements (e.g., points, leaderboards, badges, narratives) in non-game contexts—has emerged as a promising strategy to enhance motivation and learning in educational settings (Hamari et al., 2014; Camacho-Sánchez et al., 2023; Wang et al., 2025). This approach is grounded in Self-Determination Theory (Ryan & Deci, 2020), which explicitly identifies gamified environments as promising contexts for satisfying the needs for *autonomy* (through choice of tasks), *competence* (through progressive levels and immediate feedback), and *relatedness* (through team-based challenges). By shifting the focus from normative comparison to criterion-referenced mastery (e.g., "beating your own score" to unlock a level), gamification may offer a didactic solution to the engagement crisis in PE (Yang et al., 2024; Sotos-Martínez et al., 2024).

Despite the growing popularity of gamification, the existing body of research presents several methodological limitations that this study aims to address. Firstly, while numerous interventions have reported positive effects on physical activity, evidence regarding the impact on the multidimensional physical self-concept in adolescents remains fragmented (Wang & Chen, 2022; Zamorano-García et al., 2023; see also recent school-based trials by Fernández-Bustos et al., 2024, and Marín-Rubio et al., 2024). Many studies fail to account for the "importance" variable, potentially obscuring why some students improve in self-worth while others do not (Nicolosi et al., 2024). Secondly, rigorous fidelity checks are often absent, making it difficult to ascertain whether the intervention was delivered as intended or if teacher variability influenced the outcomes (Li & Xiang, 2023). Thirdly, a pervasive issue in school-based research is the "unit of analysis error," where data collected from students nested within classes are analysed using standard ANOVA techniques that violate independence assumptions, thereby inflating Type I error rates (Li et al., 2017; Silverman & Solmon, 1998; see Rosenstiel et al., 2022, and Salmon et al., 2023, for modern cluster-RCT applications). Finally, few studies have combined validated psychometric tools with gold-standard field-based fitness assessments (such as the ALPHA-FIT battery) to provide a comprehensive picture of both the psychological and physiological adaptations to gamified teaching (Ruiz et al., 2011).

To date, literature lacks interventions that simultaneously evaluate physiological, multidimensional psychological, and emotional adaptations to gamification while accounting for individual differences in domain valuation. Systematic reviews have consistently identified this fragmentation, noting that PE intervention studies rarely integrate objective fitness assessment with validated psychological measures (Carl et al., 2022), and that gamification research has overwhelmingly focused on self-reported motivational outcomes to the exclusion of physical activity levels and motor skills (Arufe-Giráldez et al., 2022; Sal-de-Rellán et al., 2025). Therefore, this study fills this knowledge gap by combining objective health-related fitness testing (ALPHA-FIT) with multidimensional psychological assessments (PSPP-R-IT), uniquely testing the moderating role of perceived importance (PIP-IT). Furthermore, its key methodological contribution lies in the use of a robust Cluster Randomized Controlled Trial design combined with advanced hierarchical statistical modelling (LMMs) to effectively eliminate the pervasive unit-of-analysis error.

### *Aim*

The primary aim of this study was to investigate the effects of a 12-week gamified didactic intervention on (1) Perceived Physical Self-Concept, (2) Health-Related Fitness, and (3) Physical Activity Enjoyment in Italian high school students. A secondary aim was to explore the moderating role of Perceived Importance. Based on the theoretical framework, the following hypotheses were formulated:

a. **H1:** The Experimental Group (EG) will demonstrate significantly greater improvements in Sports Competence (SC) and Physical Self-Worth (PSW) compared to the Control Group (CG).
b. **H2:** The EG will show significantly greater enhancements in health-related fitness (specifically agility and cardiorespiratory fitness) and physical activity enjoyment compared to the CG.
c. **H3:** Perceived importance (PIP-IT scores) will significantly moderate the intervention's effects, with students attributing higher value to specific domains experiencing greater psychological benefits.

## METHODS

### Study Design and Participants

The study employed a two-arm, parallel-group, cluster randomized controlled trial (Cluster-RCT) design. Given the educational setting, randomisation was performed at the class level rather than the individual student level to maintain ecological validity and prevent contamination between experimental conditions within the same instructional environment (Li et al., 2017). Intact classes were randomly assigned to either the experimental group (EG), which received the gamified didactic module, or the control group (CG), which followed a traditional direct instruction curriculum. Assessments were conducted at two time points: baseline (T0) and post-intervention (T1), separated by a 12-week intervention period. The randomisation involved exactly 10 intact classes (clusters), assigned equally to the EG (k = 5) and the CG (k = 5), with an average cluster size (*m*) of 21 students.

The study adhered strictly to the CONSORT guidelines extended for cluster trials (Campbell et al., 2012), and a detailed flowchart of participant enrolment, allocation, and analysis is provided in Figure 1.

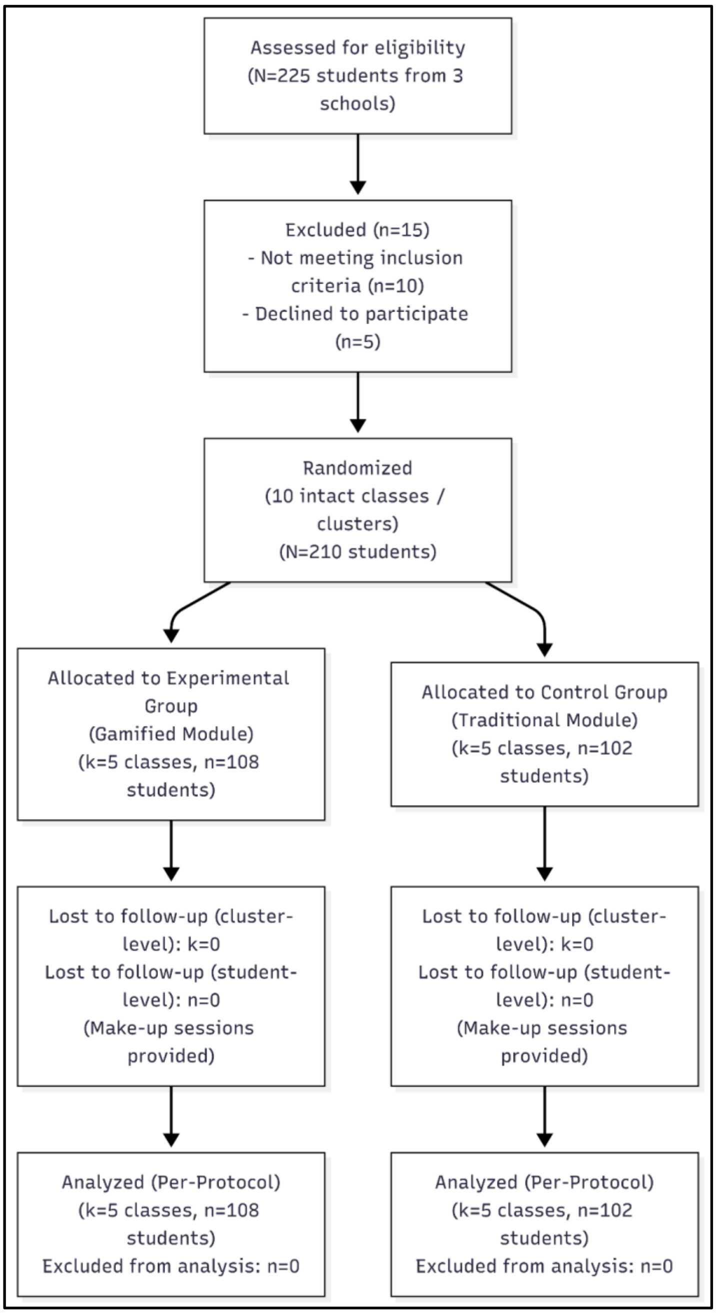


**Figure 1.** CONSORT flow diagram for the cluster randomised controlled trial, illustrating the flow of classes (clusters) and students through the phases of enrolment, allocation, follow-up, and per-protocol analysis.

Participants were recruited from three public high schools located in Northern Italy. The initial pool consisted of 225 students. Inclusion criteria were: (a) enrolment in the first or second year of secondary school (aged 14–17 years); (b) absence of physical or cognitive disabilities that would preclude safe participation in motor testing; and (c) willingness to adhere to the measurement protocol. Students who were professional athletes (training > 10 hours/week) were excluded to avoid confounding effects related to advanced motor competence (Burton et al., 2023).

Based on these criteria, 15 students were excluded. The final sample consisted of 210 adolescents (118 males, 92 females; mean age = 15.4 ± 0.9 years).

An a priori power analysis was initially conducted using G*Power 3.1 (Faul et al., 2007) for a repeated measures framework ($f$ = 0.25), power = .80, $\alpha$ = .05), indicating a baseline requirement of 128 individuals. However, considering the nested structure, the Design Effect (DE) was explicitly calculated to determine the required sample size for a cluster trial. Assuming an average cluster size ($m$) of 21 and a conservative a priori Intraclass Correlation Coefficient (ICC) of .05 based on similar school-based trials (Salway et al., 2024), the DE was estimated as $1 + (m - 1) \times ICC = 2.0$. Consequently, the cluster-adjusted required sample size was 128 × 2.0 = 256 students. While the final recruited sample ($N$ = 210) was slightly below this optimal target, the risk of underpowering was mitigated through the use of robust parameter estimation within the linear mixed models (LMM) and the inclusion of powerful individual-level covariates (e.g., maturity offset, baseline scores) known to substantially reduce residual variance and enhance statistical power in mixed-effects designs.

## Ethical approval statement

Ethical approval for this study was granted by the University Territorial Ethics Committee (Reference: UTEC 2024–0029). The research was conducted in accordance with the ethical principles of the Declaration of Helsinki. Written informed consent was obtained from the parents or legal guardians of all participants, and written assent was provided by the students before participation. Personal data were collected and processed in compliance with the General Data Protection Regulation (GDPR).

## Instruments

### Perceived Physical Self-Concept

To assess the multidimensional physical self, the Italian Revised Physical Self-Perception Profile (PSPP-R-IT) was utilized (Nicolosi et al., 2024). This instrument, validated specifically for the Italian adolescent population (15–20 years), consists of 30 items assessing four specific subdomains: sports competence (SC), physical conditioning (PC), body attractiveness (BA), and physical strength (PS). Additionally, a superordinate domain, physical self-worth (PSW), evaluates general feelings of happiness and satisfaction with the physical self. Items are rated on a 4-point structured alternative format to minimise social desirability bias. The Italian version has demonstrated excellent internal consistency (Cronbach's $\alpha$ = 0.89–0.95) and construct validity (Nicolosi et al., 2024).

### Perceived importance

The Italian Perceived Importance Profile (PIP-IT) was administered to measure the subjective value students attach to each of the PSPP-R subdomains (e.g., how important it is for the student to be good at sports). This scale allows for the examination of the discrepancy between competence and importance, a key predictor of self-esteem according to James's hypothesis (Fox & Corbin, 1989; Nicolosi et al., 2024).

### Health-related fitness

Objective physical fitness was assessed using the High-Priority ALPHA Health-Related Fitness Test Battery, an evidence-based gold standard for European adolescents (Ruiz et al., 2011). Importantly, the variables included in the present study (i.e., cardiorespiratory fitness and musculoskeletal strength) were employed as operational proxies of physical performance related to perceived competence, rather than as direct indicators of skill-based motor competence. The selected tests were chosen to align with the domains of the PSPP-R, as detailed below.

Cardiorespiratory fitness (corresponding to physical condition, PC) was assessed using the 20-m shuttle run test. Participants ran repeatedly between two lines set 20 m apart, synchronising their pace with audio signals that increased progressively in frequency at one-minute intervals. The final stage successfully completed was recorded and used to estimate maximal oxygen uptake ($\dot{V}O_2max$) according to the equation proposed by Léger et al. (1988).

Musculoskeletal fitness (corresponding to physical strength, PS) was evaluated using two complementary measures. Lower-body explosive strength was assessed through the standing broad jump, in which participants jumped horizontally from a stationary position; the longest distance achieved across two attempts was recorded in centimetres. Upper-body strength was assessed via handgrip strength using a Jamar Plus+ digital dynamometer, with grip span individually adjusted according to hand size, following standardised recommendations (Ruiz et al., 2011).

Agility and speed were assessed using a timed 4 × 10 m Shuttle Run Test. While recognizing that this test does not measure contextual or tactical sporting intelligence (as required by broader definitions of sport competence), it is a validated measure of speed of movement, agility, and coordination (Ortega et al., 2008), which serve as fundamental physical proxies for the sport competence domain assessed in the PSPP-R-IT.

### Physical Activity Enjoyment

The emotional response to the PE lessons was measured using the Physical Activity Enjoyment Scale (PACES), Italian version (Carraro et al., 2008). The scale comprises 16 items rated on a 5-point bipolar Likert scale (e.g., “I enjoy it” vs “I hate it”). Prior to data analysis, the seven items with negative valence were reverse-scored. A total enjoyment score was then computed by summing the responses to all 16 items, in accordance with the protocol established by Carraro et al. (2008). A higher total score indicates greater enjoyment.

### Biological maturation

To control for the confounding effects of growth and maturation on physical performance (Burton et al., 2023), standing height, sitting height, and body mass were measured. These values were used to calculate the maturity offset (years from peak height velocity) using the sex-specific equations proposed by Mirwald et al. (2002).

## Procedures and Intervention

### Fidelity of implementation

To ensure the internal validity of the study, a treatment fidelity checklist was developed based on the framework by Li & Xiang (2023). External researchers observed a random subsample (20%) of the lessons in both groups to verify adherence to the assigned protocol. The checklist assessed key indicators such as "use of gamified language," "provision of autonomy-supportive feedback," and "adherence to lesson content." Only data from classes meeting a fidelity threshold of >80% were included in the final analysis.

**Experimental group (gamified didactics)**

Students assigned to the EG participated in *MoveQuest*, a 12-week gamified instructional module designed to support basic psychological needs. The didactic approach was implemented through a dedicated mobile web application, in which physical activities were framed as progressively structured "missions".

Game mechanics were based on the completion of motor tasks (e.g., completing five laps of the shuttle run), which rewarded students with experience points (XP). The accumulation of XP enabled progression through hierarchical levels (e.g., from novice to elite), thereby providing structured feedback on engagement and progression.

Assessment followed a criterion-referenced logic (Yang et al., 2024), whereby students' performance was evaluated against predefined standards rather than through normative peer comparison. For instance, an individual improvement of 5 cm in the standing broad jump resulted in the award of a "personal best" badge, irrespective of the absolute distance achieved.

From a pedagogical perspective, the teacher assumed the role of a "game master", fostering autonomy by allowing students to select task difficulty levels ("challenge zones") that were congruent with their perceived competence and current ability.

**Control group (traditional didactics)**

Students in the CG followed a standard multi-activity physical education curriculum covering the same core physical content (i.e., endurance, strength, and agility) as the experimental group, but delivered through direct instruction. Lessons typically adhered to a conventional structure consisting of a warm-up phase, teacher-led demonstration, structured practice in the form of drills, and a closing phase.

Assessment and feedback were based on a norm-referenced approach (Yang et al., 2024). Students' performance was evaluated relative to their peers and commonly aligned with standard school grading norms. For example, specific performance thresholds (e.g., six completed laps corresponding to Grade 6, eight laps to Grade 8) were used to rank or grade students, thereby emphasising social comparison.

## Data Analysis

Data were analysed using IBM SPSS Statistics (version 29.0). Preliminary screening included the Shapiro–Wilk test to assess normality and Levene's test to evaluate homogeneity of variance.

Given the hierarchical structure of the data, with students nested within classes, the use of standard analysis of variance would have violated the assumption of independence of observations (Li et al., 2017). Consequently, linear mixed models (LMMs) were employed to examine intervention effects. Models were estimated using Restricted Maximum Likelihood (REML) to provide unbiased estimates of variance and covariance parameters, which is particularly appropriate for small-to-moderate numbers of clusters. Degrees of freedom for fixed effects were computed using Satterthwaite's approximation.

The final models specified fixed effects for group, time, and the group × time interaction. With respect to the random-effects structure, random intercepts were included at both the class level (Class ID), to account for clustering effects, and the individual level (Student ID), to model repeated measurements. Random slopes for Time at the class level were initially tested but subsequently removed, as they did not improve model fit according to Akaike's information criterion (AIC) and Bayesian information criterion (BIC), or resulted in convergence issues.

Baseline scores, mean-centered Perceived Importance (PIP-IT scores), and Maturity Offset were included as covariates to control for pre-existing differences and inter-individual variability in biological maturation.

Outcomes were hierarchically defined a priori, with physical self-perception, health-related fitness, and enjoyment specified as primary outcomes, and Perceived Importance as a secondary outcome. Given the multiple outcomes assessed, the risk of inflated Type I error was acknowledged. However, rather than applying a highly conservative global adjustment (e.g., Bonferroni) that could obscure theoretically grounded specific effects, exact p-values were reported alongside the corresponding effect sizes (d), with significance interpreted primarily on the basis of effect magnitude and a priori hierarchical hypotheses. Effect sizes were calculated as Cohen's *d* for between-group differences at T1, dividing the difference in estimated marginal means by the pooled baseline standard deviation ($SD_{pooled}$). Magnitude was interpreted as small (0.2), medium (0.5), or large (0.8), in accordance with established conventions (Wang & Chen, 2022). Statistical significance was set at $\alpha < .05$.

## RESULTS

### Preliminary analysis

No missing data were observed for the primary outcomes as make-up testing sessions were organized for absent students. The analysis of treatment fidelity, conducted on 24 randomly selected lessons (20% of the total), indicated high adherence to the protocol. The mean fidelity score was 92% for the EG (gamified) and 94% for the CG (traditional), exceeding the 80% threshold required for inclusion (Li & Xiang, 2023). Internal consistency reliability for the psychometric instruments in the current sample was satisfactory. Cronbach's alpha coefficients for the PSPP-R-IT subscales ranged from .83 to .88 (sports competence $\alpha$ = .86; physical conditioning $\alpha$ = .84; body attractiveness $\alpha$ = .88; physical strength $\alpha$ = .83; physical self-worth $\alpha$ = .87). The PIP-IT subscales also demonstrated adequate reliability ($\alpha$ range = .79–.85), and the PACES showed excellent internal consistency ($\alpha$ = .92).

Baseline characteristics of the participants are presented in Table 1. Independent samples t-tests revealed no significant differences between the experimental and control groups at baseline for age ($p$ = .64), BMI ($p$ = 0.58), maturity offset ($p$ = .72), or any of the motor competence and psychological variables ($p$ > .05). This confirms that randomisation was successful in creating equivalent groups.

The Intraclass Correlation Coefficients (ICC) for the primary outcomes indicated non-negligible clustering effects, confirming the necessity of using linear mixed models to account for the class-level variance. Specifically, the baseline variance components (between-class variance [$\tau_{00}$] and residual individual variance [$\sigma^2$]) were as follows: Sports competence ($\tau_{00}$ = 0.024, $\sigma^2$ = 0.324, ICC = .07), physical self-worth ($\tau_{00}$ = 0.031, $\sigma^2$ = 0.312, ICC = .09), VO$_2$max ($\tau_{00}$ = 1.05, $\sigma^2$ = 25.1, ICC = .04), and PACES ($\tau_{00}$ = 0.041, $\sigma^2$ = 0.472, ICC = .08).

**Table 1.** Baseline characteristics of participants by group

| Characteristic | Experimental Group (Gamified) (*n* = 108) | Control Group (Traditional) (*n* = 102) | Total (*N* = 210) | *p*-value |
|---|---|---|---|---|
| **Demographics** | | | | |
| Age (years) | 15.3 (0.8) | 15.5 (0.9) | 15.4 (0.9) | .64 |
| Gender (*n* / %) | | | | |
| Male | 60 (55.6%) | 58 (56.9%) | 118 (56.2%) | .84 |
| Female | 48 (44.4%) | 44 (43.1%) | 92 (43.8%) | |
| **Anthropometrics** | | | | |
| Height (cm) | 168.4 (8.2) | 169.1 (7.9) | 168.7 (8.0) | .52 |
| Weight (kg) | 61.2 (9.5) | 62.0 (8.9) | 61.6 (9.2) | .51 |
| BMI (kg/m²) | 21.5 (2.8) | 21.6 (2.6) | 21.5 (2.7) | .58 |
| Maturity Offset (years) | 1.8 (0.6) | 1.9 (0.7) | 1.8 (0.7) | .72 |
| **Psychological baseline** | | | | |
| Physical Self-Worth (1–4) | 2.65 (0.58) | 2.68 (0.61) | 2.66 (0.59) | .70 |
| Sports Competence (1–4) | 2.45 (0.60) | 2.42 (0.58) | 2.43 (0.59) | .71 |
| Enjoyment (PACES) (1–5) | 3.20 (0.75) | 3.25 (0.70) | 3.22 (0.72) | .61 |
| **Motor competence baseline** | | | | |
| 20-m Shuttle Run (laps) | 32.5 (12.1) | 33.1 (11.8) | 32.8 (11.9) | .71 |
| Standing Broad Jump (cm) | 158.4 (22.3) | 160.2 (21.5) | 159.3 (21.9) | .55 |
| 4 x 10m Shuttle Run (s) | 12.1 (1.4) | 11.9 (1.3) | 12.0 (1.4) | .28 |

*Note. Data are presented as Mean (Standard Deviation) unless otherwise indicated. BMI = Body Mass Index; PACES = Physical Activity Enjoyment Scale. p-values derived from independent samples t-tests (for continuous variables) and Chi-square tests (for categorical variables).*

## Perceived Physical Self-Concept

The linear mixed model revealed a significant group × time interaction for the sports competence subscale ($F_{(1,208)}$ = 14.52, $p$ < .001) ($\beta$ = 0.35, *SE* = 0.09, 95% CI[0.17, 0.53]). Follow-up pairwise comparisons indicated that while both groups improved, the EG showed a significantly larger increase from T0 (*M* = 2.45, *SD* = 0.60) to T1 (*M* = 2.98, *SD* = 0.55) compared to the CG (T0: *M* = 2.42, *SD* = 0.58; T1: *M* = 2.60, *SD* = 0.62). The effect size for the between-group difference at post-test was medium-to-large (*d* = 0.65). Similarly, a significant interaction was found for the superordinate domain of physical self-worth ($F_{(1,208)}$ = 11.34, $p$ = .001) ($\beta$ = 0.31, *SE* = 0.09, 95% CI[0.13, 0.49]). The Gamified group demonstrated a substantial improvement (*d* = 0.58), whereas the CG showed only marginal gains.

In contrast, no significant interaction effect was observed for the body attractiveness subscale (*p* = .42). Scores for this domain remained relatively stable in both groups over the 12-week period, suggesting that perception of body image was not significantly altered by the short-term didactic intervention.

## Health-related fitness

Results for the ALPHA-FIT battery tests are summarized in Table 2. For cardiorespiratory fitness (20-m Shuttle Run), a significant main effect of Time was observed ($F_{(1,208)}$ = 35.40, $p$ < .001), with both groups improving their VO$_2$max estimates. However, the group × time interaction was also significant (*p* = .023) ($\beta$ = 2.21, *SE* = 0.96, 95% CI[0.31, 4.11]), with the EG showing a greater magnitude of improvement (*d* = 0.42) compared to the CG (*d* = 0.21).

Regarding musculoskeletal fitness (standing broad jump), both groups showed significant improvements over time (*p* < .001), but the interaction effect was not significant (*p* = .18), indicating similar gains regardless of the teaching method.

For the 4 x 10m shuttle run (sports competence/agility), the analysis revealed a significant interaction ($p$ = .004) ($\beta$ = -0.50, $SE$ = 0.17, 95% CI[-0.84, -0.16]). Students in the gamified condition improved their agility times significantly more than those in the traditional condition ($d$ = 0.51) vs ($d$ = 0.25).

**Table 2**. Effects of gamified vs. traditional didactics on physical self-perception, enjoyment, and health-related fitness parameters (Linear Mixed Models results).

| Outcome | Group | T0 (Baseline) M (SD) | T1 (Post) M (SD) | Group × Time Interaction Estimate $\beta$ [SE] | 95% CI for Interaction | $p$-value | Effect Size ($d$) | $\tau_{00}$ (Class Var) | $\sigma^2$ (Resid. Var) | ICC |
|---|---|---|---|---|---|---|---|---|---|---|
| **Perceived Physical Self (PSPP-R)** | | | | | | | | | | |
| Physical Self-Worth | Exp | 2.65 (0.58) | 2.98 (0.55) | 0.31 [0.09] | [0.13, 0.49]*** | .001 | 0.58 | 0.031 | 0.312 | .09 |
| | Con | 2.68 (0.61) | 2.70 (0.63) | | | | | | | |
| Sports Competence | Exp | 2.45 (0.60) | 2.98 (0.55) | 0.35 [0.09] | [0.17, 0.53]*** | < .001 | 0.65 | 0.024 | 0.324 | .07 |
| | Con | 2.42 (0.58) | 2.60 (0.62) | | | | | | | |
| Body Attractiveness | Exp | 2.30 (0.65) | 2.35 (0.68) | 0.07 [0.09] | [-0.10, 0.24] | .420 | 0.08 | 0.018 | 0.405 | .04 |
| | Con | 2.32 (0.62) | 2.30 (0.64) | | | | | | | |
| Physical Conditioning | Exp | 2.50 (0.55) | 2.80 (0.52) | 0.27 [0.10] | [0.07, 0.47]** | .009 | 0.44 | 0.021 | 0.289 | .07 |
| | Con | 2.55 (0.58) | 2.58 (0.59) | | | | | | | |
| Physical Strength | Exp | 2.40 (0.59) | 2.65 (0.56) | 0.13 [0.07] | [-0.01, 0.27] | .065 | 0.25 | 0.015 | 0.335 | .04 |
| | Con | 2.38 (0.61) | 2.50 (0.60) | | | | | | | |
| **Psychological** | | | | | | | | | | |
| Enjoyment (PACES) | Exp | 3.20 (0.75) | 4.10 (0.65) | 0.95 [0.18] | [0.60, 1.30]*** | < .001 | 0.82 | 0.041 | 0.472 | .08 |
| | Con | 3.25 (0.70) | 3.20 (0.74) | | | | | | | |
| **Health-Related Fitness** | | | | | | | | | | |
| 20-m Shuttle Run ($\dot{V}O_2max$) | Exp | 42.5 (5.2) | 45.8 (4.8) | 2.21 [0.96] | [0.31, 4.11]* | .023 | 0.42 | 1.05 | 25.1 | .04 |
| | Con | 42.8 (5.0) | 43.9 (4.9) | | | | | | | |
| Standing Broad Jump (cm) | Exp | 158.4 (22.3) | 165.2 (20.5) | 2.20 [1.63] | [-1.01, 5.41] | .182 | 0.15 | 32.5 | 452.1 | .07 |
| | Con | 160.2 (21.5) | 164.8 (21.0) | | | | | | | |
| Handgrip Strength (kg) | Exp | 28.5 (6.2) | 30.1 (5.8) | 0.10 [0.80] | [-1.48, 1.68] | .331 | 0.10 | 1.85 | 35.2 | .05 |
| | Con | 29.0 (5.9) | 30.5 (5.7) | | | | | | | |
| 4 x 10m Agility/Speed (s)† | Exp | 12.1 (1.4) | 11.2 (1.2) | -0.50 [0.17] | [-0.84, -0.16]** | .004 | 0.51 | 0.12 | 1.84 | .06 |
| | Con | 11.9 (1.3) | 11.7 (1.3) | | | | | | | |

*Note. Exp = Experimental Group; Con = Control Group. $\beta$ represents the fixed effect estimate for the Group × Time interaction derived from Linear Mixed Models (REML estimation, Satterthwaite df). $\tau_{00}$ represents the between-class variance (random intercept variance at cluster level); $\sigma^2$ represents the residual variance (within-individual variance). ICC = Intraclass Correlation Coefficient ($\tau_{00}/(\tau_{00}+\sigma^2)$). †Lower values indicate better performance (negative $\beta$ indicates greater improvement in Exp); *p < .05, **p ≤ .01, ***p ≤ .001.*

### Physical activity enjoyment

A strong group × time interaction was found for PACES scores ($F_{(1,208)}$ = 28.6, $p$ < .001) ($\beta$ = 0.95, $SE$ = 0.18, 95% CI[0.60, 1.30]). The EG reported a marked increase in enjoyment ($d$ = 0.82, large effect), whereas enjoyment scores in the CG remained static or slightly declined ($d$ = 0.10).

### Role of perceived importance (covariate analysis)

The inclusion of mean-centered PIP-IT scores as covariates in the LMM revealed a significant moderating effect. For physical self-worth, the three-way interaction between Group × Time × Sports Competence Importance score was significant ($p$ = .015, $\beta$ = 0.18, $SE$ = 0.07, 95% CI[0.04, 0.32]). As illustrated in the interaction plot (Figure 2), students in the gamified group who attributed high importance to sports competence (+1 SD above the mean PIP-IT score) experienced the greatest gains in self-worth $d$ = 0.72) compared to those with low perceived importance (-1 SD; $d$ = 0.35). This suggests that the intervention was particularly effective for students who already value the domain being trained.

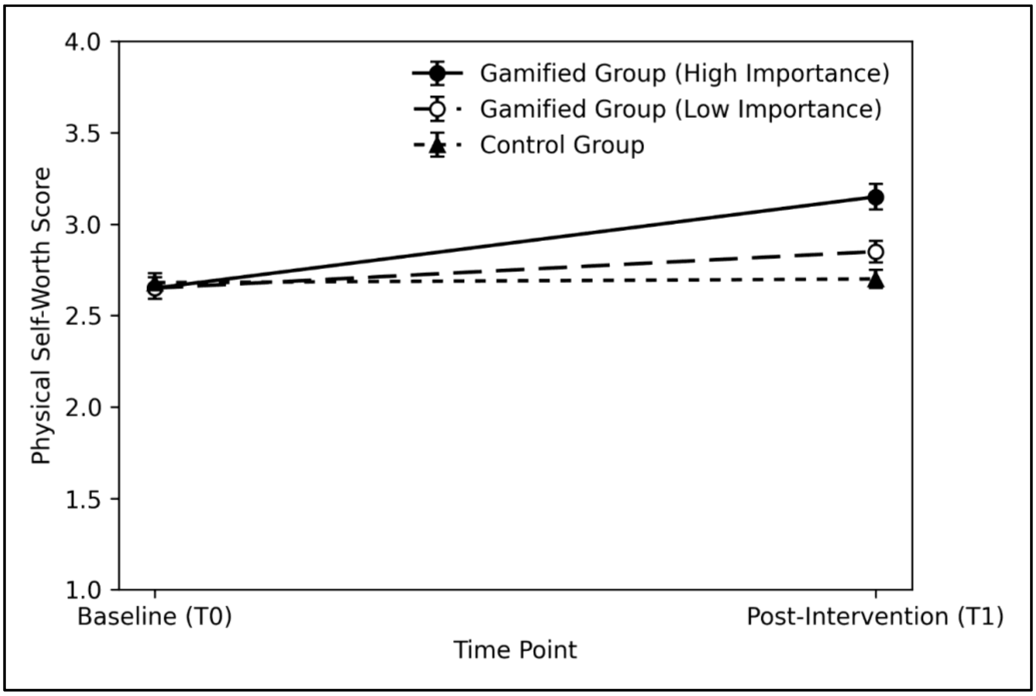


**Figure 2.** Interaction plot illustrating the moderating role of Perceived Importance (PIP-IT Sports Competence subscale) on the intervention effect for Physical Self-Worth. The graph depicts the estimated marginal means at baseline (T0) and post-intervention (T1). The solid line represents students in the Experimental (Gamified) group with high perceived importance (+1 *SD*), the dashed line represents students in the Experimental group with low perceived importance (-1 *SD*), and the dotted line represents the Control group. Error bars represent standard errors.

## DISCUSSION

The present study aimed to evaluate the efficacy of a gamified didactics intervention on physical self-perception, actual health-related fitness, and enjoyment among Italian high school students. The findings provide empirical support for the integration of gamification in Physical Education (PE), demonstrating its potential to positively influence specific domains of the physical self and motor performance, albeit with some nuances.

### Impact on physical self-perception and enjoyment

A key finding of this investigation was the superior improvement in sports competence and physical self-worth observed in the gamified group compared to the traditional instruction group. Furthermore, the medium-to-large effect size observed for physical self-worth ($d = 0.58$) is substantial, especially when benchmarked against recent meta-analyses indicating that physical activity-based interventions typically yield more modest pooled effects on physical self-concept in adolescents (Zamorano-García et al., 2023). This result aligns with the theoretical tenets of Self-Determination Theory embedded in the intervention design. As anticipated, replacing normative peer comparisons with criterion-referenced mastery goals likely fostered a more supportive environment for competence satisfaction (Yang et al., 2024). This shift is supported by recent findings indicating that reducing the sense of public judgment and surveillance in fitness activities is essential for fostering student agency and comfort (Alfrey & Young, 2025). According to Ryan & Deci (2020), comparative grading schemes typically provide little competence-relevant feedback, merely informing students of their rank relative to others. By removing this 'controlling' pressure—which inadvertently highlights incompetence in less able students (Li et al., 2017)—the gamified intervention allowed feedback to be perceived as informational rather than evaluative. This distinction is critical: as observed in a previous qualitative analysis of Italian students (Cereda, 2025), technology can create a detrimental "digital panopticon" when used for public comparison. By contrast, the current intervention utilized digital tools to foster private progression, effectively mitigating the anxiety and 'shame' associated with public performance tracking.

The significant increase in enjoyment (PACES) within the experimental group further corroborates this interpretation. The magnitude of this effect ($d$ = 0.82) is particularly notable when compared to recent meta-analytical pooled estimates for game-based PE interventions (e.g., 0.53 reported by Mo et al., 2024), confirming the high emotional engagement elicited by the didactic format. This observation is consistent with Ashley & Kawabata (2023), who identified the experience of enjoyment as a pivotal factor in shaping positive perceptions of fitness testing, specifically within a Self-Determination Theory framework. As suggested by Iannaccone et al. (2025), the affective domain is a critical component of physical literacy. This is consistent with recent interdisciplinary reviews highlighting that mastery-oriented learning environments, which prioritize personal improvement over normative comparison, are essential for fostering positive motivational patterns and comprehensive physical literacy (Cereda, 2026). The "fun" element inherent in gamification—achieved through badges, levels, and narrative—appears to act as a catalyst, transforming repetitive motor tasks into engaging challenges. This emotional engagement is crucial during adolescence, a period often marked by a decline in PE interest and participation (Burton et al., 2023).

However, it is noteworthy that Body Attractiveness perceptions did not change significantly in either group. This stability is consistent with the findings of Nicolosi et al. (2024), who suggest that body image is a deeply entrenched construct, heavily influenced by external societal factors and pubertal changes, and thus less permeable to short-term school-based interventions compared to competence-based domains.

**Health-related fitness gains**

Regarding health-related fitness, the results were mixed. While both groups improved in musculoskeletal fitness (standing broad jump), suggesting that the physiological stimulus of the lessons was effective regardless of the pedagogical wrapper, the gamified group showed significantly greater gains in cardiorespiratory fitness and agility. This differential effect might potentially reflect variations in behavioral engagement and "time on task". In the gamified condition, the desire to unlock the next level or earn a badge may have encouraged students to sustain a higher volume of active practice during the shuttle runs, potentially contributing to greater physiological adaptations. However, as internal load (e.g., heart rate or Ratings of Perceived Exertion) was not monitored, definitive conclusions regarding higher cardiovascular intensity in the gamified group remain unwarranted. Therefore, any causal inferences regarding physiological mechanisms require cautious interpretation and future objective verification. This supports the notion that didactic strategies influence how students move, not just that they move (Wang & Chen, 2022). Indeed, recent evidence indicates that participation in structured physical activity alone does not guarantee motor competence if the pedagogical focus remains on repetitive technical drills rather than diverse movement experiences. The use of the ALPHA-FIT battery (Ruiz et al., 2011) ensured that these findings are robust and comparable with European normative data, adding validity to the observation that gamification can support health-related fitness without compromising skill acquisition. Furthermore, a distinction must be drawn between the fitness components assessed here (fitness and agility) and motor competence (skill mastery). However, recent theoretical reconceptualization suggests that improved product-oriented outcomes (e.g., jump distance, running speed) may reflect a shift from stability to mobility movement strategies, signalling a breakthrough of the proficiency barrier (Brian et al., 2020). Thus, the improvements observed likely reflect both physiological adaptations and enhanced movement efficiency.

**The role of perceived importance**

A novel contribution of this study, enabled by the use of the PIP-IT (Nicolosi et al., 2024), was the identification of perceived importance as a significant covariate. The finding that gains in self-worth were magnified for students who valued sports competence suggests that educational interventions are not *one-size-fits-all*. Gamification seems particularly potent for students who already have a vested interest in the domain. However, this also raises a pedagogical challenge: future interventions may need to explicitly target the value students assign to physical activity (perhaps through cognitive engagement or reflection) to broaden the impact to those who are currently disengaged or devalue PE.

## Limitations and Strengths

The findings of the present study should be interpreted in light of several limitations. First, the intervention duration (12 weeks), while typical for school-based research, may be insufficient to observe changes in more stable constructs such as body image. Furthermore, it is not possible to definitively rule out a novelty response (e.g., the 'Hawthorne effect'), whereby the initial engagement elicited by the technological and gamified elements might wane over time. Longitudinal follow-ups are therefore essential to assess the long-term retention of the observed psychological and physiological gains. Second, although the cluster-randomized design is methodologically appropriate for educational settings, it inherently reduces statistical power compared to individual randomization. While the use of Linear Mixed Models (LMM) mitigated this risk by correctly modeling the nested variance (Li et al., 2017), the present sample size was not adequately powered to conduct complex three-way interactions (e.g., exploring gender-specific or maturity-specific trajectories), which remains an important avenue for future research. Third, despite the rigorous assessment of treatment fidelity, it was not possible to fully isolate a potential 'teacher effect'; the enthusiasm or specific pedagogical style of individual educators may have clustered within classes, thereby influencing the outcomes. Finally, the study was conducted in a specific region of Northern Italy, which may limit generalizability to other cultural or educational contexts.

Conversely, notable strengths of the study include the rigorous adherence to CONSORT guidelines extended for cluster trials, the explicit calculation of the design effect, and the high treatment fidelity verified via a standardized checklist (Li & Xiang, 2023). Furthermore, the concurrent use of objective health-related fitness testing (ALPHA-FIT) and valid, culturally adapted multidimensional psychological instruments (PSPP-R-IT, PIP-IT, PACES) provides a comprehensive, dual-domain evaluation that addresses the fragmentation frequently observed in previous gamification research (Arufe-Giráldez et al., 2022; Carl et al., 2022). This combined approach is particularly relevant, as recent systematic evidence confirms that physical fitness outcomes and motor competence variables continue to be largely absent from gamification research in secondary PE (Sal-de-Rellán et al., 2025).

## CONCLUSIONS

Gamification appears to be a promising and viable didactic strategy for high school PE. Within the parameters of this 12-week trial, prioritizing individual progress and autonomy demonstrated the potential to enhance students' perception of their sports competence and enjoyment of physical activity more effectively than traditional directive methods. While it may not immediately alter deep-seated body image perceptions, its ability to improve actual fitness and self-worth positions it as a valuable tool for fostering Physical Literacy. However, before widespread pedagogical implementation, multicentre replications with longer follow-ups and cost-benefit analyses regarding digital infrastructure are warranted. Educators are cautiously encouraged to explore student-centred, gamified formats, particularly for developing cardiorespiratory and agility skills.

## ACKNOWLEDGMENTS

The author gratefully acknowledges the commitment of the schools, teachers, and students whose participation made this research possible. Sincere appreciation is also extended to the graduate kinesiology students for their invaluable assistance with the blinded data collection and entry processes. All phases of this research, including the study's conceptualisation and design, statistical analysis, interpretation of the findings, and the preparation of this manuscript, were undertaken solely by the author.

## AI DISCLOSURE STATEMENT

During the preparation of this manuscript, the author used DeepL Write to improve linguistic clarity, grammar, and style, as English is not the author's native language. All AI-generated outputs were critically reviewed and thoroughly edited by the author to ensure factual accuracy, clarity of expression, and compliance with academic standards. The author takes full responsibility for the integrity and content of this manuscript.

## DATA AVAILABILITY

All data supporting the findings of this study are included in the article. Additional datasets are available from the corresponding author upon reasonable request.

## FUNDING

This research received no external funding.

## CONFLICT OF INTEREST

The author declares no conflict of interest.

## REFERENCES


Alfrey, L. (2024). An expansive learning approach to transforming traditional fitness testing in health and physical education: student voice, feelings and hopes. *Curriculum Studies in Health and Physical Education*, *15*(1), 24–39. https://doi.org/10.1080/25742981.2023.2183477

Alfrey, L., & Young, L. (2025). Student voice, expansive learning and transformative agency: an examination of the transformation of fitness testing pedagogies in Health and Physical Education. *Sport, Education and Society*, 1–17. https://doi.org/10.1080/13573322.2025.2592832

Arufe-Giráldez, V., Sanmiguel-Rodríguez, A., Ramos-Álvarez, O., & Navarro-Patón, R. (2022). Gamification in Physical Education: A Systematic Review. *Education Sciences*, *12*(8), 540. https://doi.org/10.3390/educsci12080540

Ashley, B. B., & Kawabata, M. (2023). Students' perceptions of fitness testing in physical education across primary, secondary, and pre-university school levels: a motivational profiles perspective. *Physical Education and Sport Pedagogy*, *28*(1), 76–93. https://doi.org/10.1080/17408989.2021.1953458

Barrientos Hernán, E. J., López-Pastor, V. M., Lorente-Catalán, E., & Kirk, D. (2023). Challenges with using formative and authentic assessment in physical education teaching from experienced teachers' perspectives. *Curriculum Studies in Health and Physical Education*, *14*(2), 109–126. https://doi.org/10.1080/25742981.2022.2060118

Blain, D. O., Standage, M., & Curran, T. (2022). Physical education in a post-COVID world: A blended-gamified approach. *European Physical Education Review*, 28(3), 757–776. https://doi.org/10.1177/1356336X221080372

Bourke, M., Wang, H. F. W., Wicks, H., Barnett, L. M., Cairney, J., & Fortnum, K. (2025). Children's and adolescents' actual motor competence, perceived physical competence and physical activity: A structural equation modelling meta-analysis. *Sports Medicine, 55*, 1923–1936. https://doi.org/10.1007/s40279-025-02233-2

Brian, A., Getchell, N., True, L., De Meester, A., & Stodden, D. F. (2020). Reconceptualizing and Operationalizing Seefeldt's Proficiency Barrier: Applications and Future Directions. *Sports medicine (Auckland, N.Z.)*, *50*(11), 1889–1900. https://doi.org/10.1007/s40279-020-01332-6

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J. P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., Lambert, E., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British journal of sports medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Burton, A. M., Cowburn, I., Thompson, F., Eisenmann, J. C., Nicholson, B., & Till, K. (2023). Associations between motor competence and physical activity, physical fitness and psychosocial characteristics in adolescents: A systematic review and meta-analysis. *Sports Medicine, 53*, 2191–2256. https://doi.org/10.1007/s40279-023-01886-1

Camacho-Sánchez, R., Manzano-León, A., Rodríguez-Ferrer, J. M., Serna, J., & Lavega-Burgués, P. (2023). Game-Based Learning and Gamification in Physical Education: A Systematic Review. *Education Sciences*, *13*(2), 183. https://doi.org/10.3390/educsci13020183

Campbell, M. K., Piaggio, G., Elbourne, D. R., Altman, D. G., & CONSORT Group (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ (Clinical research ed.)*, *345*, e5661. https://doi.org/10.1136/bmj.e5661

Carl, J., Barratt, J., Wanner, P., Töpfer, C., Cairney, J., & Pfeifer, K. (2022). The Effectiveness of Physical Literacy Interventions: A Systematic Review with Meta-Analysis. *Sports medicine (Auckland, N.Z.)*, *52*(12), 2965–2999. https://doi.org/10.1007/s40279-022-01738-4

Carraro, A., Young, M. C., & Robazza, C. (2008). A contribution to the validation of the Physical Activity Enjoyment Scale in an Italian sample. *Social Behavior and Personality: An International Journal,* *36*(7), 911–918. https://doi.org/10.2224/sbp.2008.36.7.911

Cereda, F. (2023a). A modern pedagogical approach in the subject of physical education. *Physical Education and Sport: Studies and Research*, *2*(2), 126–141. https://doi.org/10.56003/pessr.v2i2.265

Cereda, F. (2023b). Methods and models in the context of physical activity and physical education: Strength, weakness and gaps. *Journal of Physical Education and Sport*, *23*(5), 1064–1075. https://doi.org/10.7752/jpes.2023.05133

Cereda, F. (2025). The body under control: Unpacking students' affective experiences and the power dynamics of wearable technologies in physical education. *Curriculum Studies in Health and Physical Education*. https://doi.org/10.1080/25742981.2025.2566164

Cereda, F. (2026). Physical literacy development through interdisciplinary approaches in educational contexts: A critical review of theory, research, and practice. *Ricerche di Pedagogia e Didattica. Journal of Theories and Research in Education*. Advance online publication. https://doi.org/10.6092/issn.1970-2221/21536

Edwards, L. C., Bryant, A. S., Keegan, R. J., Morgan, K., & Jones, A. M. (2017). Definitions, Foundations and Associations of Physical Literacy: A Systematic Review. *Sports medicine (Auckland, N.Z.)*, *47*(1), 113–126. https://doi.org/10.1007/s40279-016-0560-7

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fernández-Bustos, J. G., Cuesta-Valera, P., Zamorano-García, D., & Simón-Piqueras, J. Á. (2024). Health-based physical education in an elementary school: effects on physical self-concept, motivation, fitness and physical activity. *Physical Education and Sport Pedagogy*, 1–15. https://doi.org/10.1080/17408989.2024.2342826

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: Development and preliminary validation. *Journal of Sport and Exercise Psychology, 11*(4), 408–430. https://doi.org/10.1123/jsep.11.4.408

Grauduszus, M., Koch, L., Wessely, S., & Joisten, C. (2024). School-based promotion of physical literacy: A scoping review. *Frontiers in Public Health*, 12, 1322075. https://doi.org/10.3389/fpubh.2024.1322075

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 16 million participants. *The Lancet. Child & adolescent health*, *4*(1), 23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Harte, N. P. A., Alfrey, L., Spray, C., & Cale, L. (2024). The if, why and how of fitness testing in secondary school physical education in the United Kingdom. *European Physical Education Review*, *30*(3), 475-492. https://doi.org/10.1177/1356336X231219937

Harter, S. (1999). *The construction of the self: A developmental perspective*. Guilford Press.

Iannaccone, A., Cudicio, A., Falese, L., Federico, B., Crotti, M., Lovecchio, N., Digennaro, S., & Agosti, V. (2025). Validation of the Italian Translation and Cultural Adaptation of the Canadian Assessment of Physical Literacy-2 (CAPL-2) Questionnaire for Children. *Children*, *12*(10), 1290. https://doi.org/10.3390/children12101290

Kirk, D. (2013). Educational value and models-based practice in physical education. *Educational Philosophy and Theory,* *45*(9), 973-986. https://doi.org/10.1080/00131857.2013.785352

Léger, L. A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 metre shuttle run test for aerobic fitness. *Journal of sports sciences*, *6*(2), 93–101. https://doi.org/10.1080/02640418808729800

Li, W., & Xiang, P. (2023). A pilot study of development and validation of treatment fidelity checklist for intervention research in K-12 school physical education. *International Journal of Physical Activity and Health,* *2*(3), Article 3. https://doi.org/10.18122/ijpah.020303.boisestate

Li, W., Chen, Y. J., Xiang, P., Xie, X., & Li, Y. (2017). Unit of analysis: Impact of Silverman and Solmon's article on field-based intervention research in physical education in the USA. *Journal of Teaching in Physical Education, 36*(2), 131–141. https://doi.org/10.1123/jtpe.2016-0169

Marín-Rubio, K., Lorente-Catalán, E., & Rubio-Campillo, X. (2024). 'Get a butt like a rock': cultivating adolescents' critical consciousness of physical self-concept through a physical education intervention. *Physical Education and Sport Pedagogy*, 1–21. https://doi.org/10.1080/17408989.2024.2425346

Mirwald, R. L., Baxter-Jones, A. D., Bailey, D. A., & Beunen, G. P. (2002). An assessment of maturity from anthropometric measurements. *Medicine and science in sports and exercise*, *34*(4), 689–694. https://doi.org/10.1097/00005768-200204000-00020

Mo, W., Saibon, J. B., Li, Y., Li, J., & He, Y. (2024). Effects of game-based physical education program on enjoyment in children and adolescents: a systematic review and meta-analysis. *BMC public health*, *24*(1), 517. https://doi.org/10.1186/s12889-024-18043-6

Nicolosi, S., Ortega Ruiz, R., Sgrò, F., Lipoma, M., & Benítez Sillero, J. D. (2024). Psychometric validation of the Revised Physical Self-Perception Profile: An Italian context study. *Behavioral Sciences, 14*(12), 1229. https://doi.org/10.3390/bs14121229

Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: a powerful marker of health. *International journal of obesity (2005)*, *32*(1), 1–11. https://doi.org/10.1038/sj.ijo.0803774

Piotrowski, T., Makaruk, H., Tekień, E., Feleszko, W., Kołodziej, M., Albrecht, K., Grela, K., Makuch, R., Werner, B., & Gąsior, J. S. (2025). Fundamental movement/motor skills as an important component of physical literacy and bridge to physical activity: A scoping review. *Children, 12*(10), 1406. https://doi.org/10.3390/children12101406

Quatman-Yates, C. C., Quatman, C. E., Meszaros, A. J., Paterno, M. V., & Hewett, T. E. (2012). A systematic review of sensorimotor function during adolescence: a developmental stage of increased motor awkwardness?. *British journal of sports medicine*, *46*(9), 649–655. https://doi.org/10.1136/bjsm.2010.079616

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor Competence and its Effect on Positive Developmental Trajectories of Health. *Sports medicine (Auckland, N.Z.)*, *45*(9), 1273–1284. https://doi.org/10.1007/s40279-015-0351-6

Rosenstiel, S., Volk, C., Schmid, J., Wagner, W., Demetriou, Y., Höner, O., Thiel, A., Trautwein, U., & Sudeck, G. (2022). Promotion of physical activity-related health competence in physical education: A person-oriented approach for evaluating the GEKOS intervention within a cluster randomized controlled trial. *European Physical Education Review*, 28(1), 279–299. https://doi.org/10.1177/1356336X211037432

Rosenstiel, S., Volk, C., Schmid, J., Wagner, W., Demetriou, Y., Höner, O., Thiel, A., Trautwein, U., & Sudeck, G. (2022). Promotion of physical activity-related health competence in physical education: A person-oriented approach for evaluating the GEKOS intervention within a cluster randomized controlled trial. *European Physical Education Review*, 28(1), 279–299. https://doi.org/10.1177/1356336X211037432

Ruiz, J. R., Castro-Piñero, J., España-Romero, V., Artero, E. G., Ortega, F. B., Cuenca, M. M., Jiménez-Pavón, D., Chillón, P., Girela-Rejón, M. J., Mora, J., Gutiérrez, Á., Suni, J., Sjöström, M., & Castillo, M. J. (2011). Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *British Journal of Sports Medicine, 45*(6), 518–524. https://doi.org/10.1136/bjsm.2010.075341

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sal-de-Rellán, A., Hernández-Suárez, Á., & Hernaiz-Sánchez, A. (2025). Gamification and motivation in adolescents. Systematic review from Physical Education. *Frontiers in psychology*, *16*, 1575104. https://doi.org/10.3389/fpsyg.2025.1575104

Salmon, J., Arundell, L., Cerin, E., Ridgers, N. D., Hesketh, K. D., Daly, R. M., ... & Crawford, D. (2023). Transform-Us! cluster RCT: 18-month and 30-month effects on children's physical activity, sedentary time and cardiometabolic risk markers. *British Journal of Sports Medicine*, 57(5), 311–319. https://doi.org/10.1136/bjsports-2022-105825

Salway, R., Jago, R., de Vocht, F., House, D., Porter, A., Walker, R., Kipping, R., Owen, C. G., Hudda, M. T., Northstone, K., van Sluijs, E., & International Children's Accelerometry Database (ICAD) Collaborators (2024). School-level intra-cluster correlation coefficients and autocorrelations for children's accelerometer-measured physical activity in England by age and gender. *BMC medical research methodology*, *24*(1), 179. https://doi.org/10.1186/s12874-024-02290-7

Saraiva Flôres, F., Paschoal Soares, D., Willig, R. M., Reyes, A. C., & Silva, A. F. (2024). Mastering movement: A Cross-sectional investigation of motor competence in children and adolescents engaged in sports. *PloS one*, *19*(5), e0304524. https://doi.org/10.1371/journal.pone.0304524

Seefeldt, V. (1980). Developmental motor patterns: Implications for elementary school physical education. *Psychology of Motor Behavior and Sport, 36*, 314-323.

Silverman, S., & Solmon, M. A. (1998). The unit of analysis in field research: Issues and approaches to design and data analysis. *Journal of Teaching in Physical Education, 17*(3), 270–284. https://doi.org/10.1123/jtpe.17.3.270

Slingerland, M., Weeldenburg, G., & Borghouts, L. (2024). Formative assessment in physical education: Teachers' experiences when designing and implementing formative assessment activities. *European Physical Education Review*, 30(4), 620–637. https://doi.org/10.1177/1356336X241237398

Sotos-Martínez, V. J., Tortosa-Martínez, J., Baena-Morales, S., & Ferriz-Valero, A. (2024). It's game time: Improving basic psychological needs and promoting positive behaviours through gamification in physical education. *European Physical Education Review*, 30(3), 435–457. https://doi.org/10.1177/1356336X231217404

Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A Developmental Perspective on the Role of Motor Skill Competence in Physical Activity: An Emergent Relationship. *Quest*, *60*(2), 290–306. https://doi.org/10.1080/00336297.2008.10483582

Wang, B., & Chen, S. (2022). Physical education curriculum interventions: A review of research patterns and intervention efficacy. *Quest, 74*(2), 205–220. https://doi.org/10.1080/00336297.2022.2032774

Wang, M., Xu, J., Zhou, X., Li, X., & Zheng, Y. (2025). Effectiveness of Gamification Interventions to Improve Physical Activity and Sedentary Behavior in Children and Adolescents: Systematic Review and Meta-Analysis. *JMIR serious games*, *13*, e68151. https://doi.org/10.2196/68151

Welk, G., Janz, K., Laurson, K., Mahar, M., Zhu, W., & Pavlovic, A. (2022). Development of Criterion-Referenced Standards for Musculoskeletal Fitness in Youth: Considerations and Approaches by the FitnessGram Scientific Advisory Board. *Measurement in Physical Education and Exercise Science*, *26*(4), 276–288. https://doi.org/10.1080/1091367X.2021.2014331

Whitehead, M. (Ed.). (2010). *Physical literacy: Throughout the lifecourse*. Routledge. https://doi.org/10.4324/9780203881903

World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behaviour*. World Health Organization.

Yang, A., Hamilton, X. D., Wang, Y., Smolianov, P., Castro-Piñero, J., Guan, J., Dolmatova, T., Zhang, X., Zhang, J., Zhan, E., & Hamilton, M. F. (2024). Results assessment methods of health-related fitness test batteries in school-based physical education programs: A global comparison. *Journal of Teaching in Physical Education, 44*(3), 564–572. https://doi.org/10.1123/jtpe.2024-0005

Zamorano-García, D., Infantes-Paniagua, Á., Cuevas-Campos, R., & Fernández-Bustos, J. G. (2023). Impact of Physical Activity-Based Interventions on Children and Adolescents' Physical Self-Concept: A Meta-Analysis. *Research Quarterly for Exercise and Sport*, *94*(1), 1–14. https://doi.org/10.1080/02701367.2021.1927945