REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity

www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333

Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence

Ensayo aleatorizado por conglomerados de educación física gamificada integrada al trabajo en equipo sobre las habilidades motrices y la competencia socioemocional de estudiantes urbanos

Dewi Ayu Lestari[1]; Soni Nopembri[1]; Pamuji Sukoco[1]; Oktaviarini Yahya Rahmadhanty[2]; Nevitaningrum[3]

[1]Department of Physical Education, Faculty of Sport and Health Sciences, Universitas Negeri Yogyakarta, Indonesia
[2]Departement of Physical Education, Health, and Recreation, Faculty of Teacher Training and Education, Singaperbangsa Karawang University, Indonesia
[3]Department of Physical Education, Faculty of Teacher Training and Education, Siliwangi University, Indonesia.

****Correspondence Author:** Oktaviarini Yahya Rahmadhanty, oktaviarini.yahya@fkip.unsika.ac.id

Editorial Schedule: *Article received: 11/07/2026 Accepted:28/09/2026 Published: 01/10/2026*

To cite this article use the following reference:
Lestari, D.A.; Nopembri, S.; Sukoco, P.; Rahmadhanty, O.Y; Nevitaningrum (2026). Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence. *Sportis Sci J, 12*(6), 1-34, https://doi.org/10.55040/6qhzb671

Author contribution: Author contribution: Study design (DAL, SN, PS, OYR), data collection (DAL), statistical analysis (DAL, OYR, NN), manuscript preparation (SN, PS, OYR, NN).

Funding: No funding was provided for this project.

Conflict of interest: The authors declare no conflict of interest.

Ethical aspects: Ethical clearance for this study was approved by the Ethics Committee for Research, Universitas Negeri Yogyakarta, Indonesia.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Abstract
Physical education should be designed not only to develop fundamental movement skills but also to foster social emotional competence of elementary school students. This study examined the effectiveness of team-building gamified physical education, compared with conventional instruction, on urban elementary students' fundamental movement skills and social-emotional competence. A cluster randomized controlled trial involved 240 fourth-grade students from eight public elementary schools in urban Karawang, with 120 students per group. The six-week intervention comprised 12 sessions conducted twice weekly, each session lasting 35 to 45 minutes. Fundamental movement skills were assessed using the TGMD-3, while social-emotional competence was measured using the Indonesian version of the Strengths and Difficulties Questionnaire (SDQ). Data were analyzed using ANCOVA, with pre-test scores included as the covariate; 95% confidence intervals were reported for the adjusted mean differences. The experimental group showed significantly higher locomotor skills, ball skills, GMQ, and prosocial behavior than the control group (all $p < .001$). For Total Difficulties, the difference favored the experimental group but was not statistically significant ($d = -0.23$, $p = .078$). The largest effect sizes were for locomotor skills ($\eta^2p = 0.339$, $d = 1.43$, 95% CI [1.53, 2.18]) and prosocial behavior ($\eta^2p = 0.243$, $d = 1.13$, 95% CI [1.06, 1.67]); for GMQ, the adjusted mean difference was 5.05 points (95% CI [3.61, 6.48], $d = 0.90$). Team-building gamified physical education appears promising for improving fundamental movement skills and prosocial behavior, but not psychosocial difficulties. Given the limited number of schools, the short intervention, and the absence of follow-up, generalization should be cautious.

Keywords: gamification; team building; physical education; fundamental movement skills; social-emotional competence

Resumen
La educación física debe diseñarse no solo para desarrollar las habilidades motrices fundamentales, sino también para fomentar la competencia socioemocional en la educación primaria. Este estudio examinó la efectividad de una educación física gamificada basada en la construcción de equipo frente a la instrucción convencional, sobre dichas habilidades y la competencia socioemocional en contexto urbano. Se realizó un ensayo controlado aleatorizado por conglomerados con 240 estudiantes de cuarto curso de ocho escuelas primarias públicas urbanas de Karawang, con 120 por grupo. La intervención, de seis semanas, comprendió 12 sesiones dos veces por semana, de 35 a 45 minutos cada una .Las habilidades motrices fundamentales se evaluaron mediante el TGMD-3, mientras que la competencia socioemocional se midió con la versión indonesia del Strengths and Difficulties Questionnaire (SDQ). Los datos se analizaron mediante ANCOVA, incluyendo las puntuaciones de pretest como covariable; se informaron intervalos de confianza del 95 % para las diferencias de medias ajustadas. El grupo experimental mostró puntuaciones significativamente superiores en habilidades locomotoras, habilidades con balón, GMQ y conducta prosocial que el grupo de control (todas las $p<0,001$). En Dificultades Totales, la diferencia favoreció al grupo experimental, pero no fue significativa ($d=-0,23$; $p = 0,078$). Los mayores tamaños del efecto se observaron en las habilidades locomotoras ($\eta^2p = 0,339$; $d =1,43$; IC del 95 %[1,53, 2,18]) y

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

en la conducta prosocial (η^2p=0,243; d=1,13; IC del 95 % [1,06, 1,67]); para el GMQ, la diferencia de medias ajustadas fue de 5,05 puntos (IC del 95 %[3,61, 6,48];d=0,90). La educación física gamificada orientada al trabajo en equipo parece prometedora para mejorar las habilidades motrices fundamentales y el comportamiento prosocial, pero no las dificultades psicosociales. Dado el número limitado de escuelas,la corta duración y la ausencia de seguimiento, la generalización debe hacerse con cautela.

Palabras clave: gamificación; trabajo en equipo; educación física; habilidades motrices fundamentales; competencia socioemocional

Introduction

Fundamental movement skills (FMS) and social-emotional competence, often conceptualized through social-emotional learning (SEL), are two essential foundations of elementary school children's development that significantly influence their physical health, psychological well-being, social participation, and school readiness. FMS not only reflects children's ability to perform locomotor and manipulative movements but also serves as the foundation for participation in a broader range of physical activities. Numerous systematic reviews and meta-analyses have demonstrated that FMS competence is positively associated with both overall physical activity and moderate-to-vigorous physical activity, and that it can be enhanced through well-designed physical education interventions (Engel et al., 2018; García-Hermoso et al., 2020; Jones et al., 2020; Yin et al., 2025). Therefore, inadequate FMS should be understood not merely as a limitation in motor performance but also as an indicator of vulnerability to lower participation in physical activity and limited movement experiences that support children's holistic development.

Social-emotional competence is a critical prerequisite for children's ability to regulate emotions, build positive relationships, demonstrate prosocial behavior, and adapt to school demands. Meta-analytic evidence indicates that universal school-based SEL interventions are consistently associated with improvements in students' social-emotional skills, behavioral outcomes, peer relationships, and academic functioning (Cipriano et al., 2023; Chen & Yu, 2022). These competencies have become increasingly relevant amid children's digitalized lifestyles, particularly in urban settings such as Karawang, where high screen exposure combined with limited opportunities for active play has been linked to psychosocial and

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

behavioral problems among school students (Stiglic & Viner, 2019; Sutomo et al., 2022). Physical education is therefore one of the most strategic pedagogical settings for fostering children's motor and social-emotional development simultaneously.

Physical education practices in elementary schools continue to focus primarily on the acquisition of motor techniques or competitive games, while social-emotional development is not always explicitly incorporated as a pedagogical objective, even though instructional quality appears to matter more than instructional time alone. Within this context, gamification has emerged as a promising instructional strategy, with evidence that well-designed game elements enhance students' motivation, engagement, and learning experiences in physical education (Arufe-Giráldez et al., 2022; Camacho-Sánchez et al., 2023; El-Tanahi et al., 2024). Nevertheless, its implementation often remains limited to superficial game elements, such as points, leaderboards, and competition, without a pedagogical design that deliberately integrates motor challenges, social interaction, reflection, and collective responsibility.

Team building and cooperative learning approaches offer a more comprehensive alternative by emphasizing positive interdependence, shared responsibility, meaningful interaction, and collaborative knowledge construction, with particularly strong effects on the social domain (Fernández-Río & Méndez-Giménez, 2016; Schulze & Huth, 2022; Böke et al., 2025). The combination of gamification and team building is therefore theoretically complementary rather than redundant: gamification provides the motivational structure (clear goals, immediate feedback, progression) that sustains engagement, whereas team building channels this engagement toward shared goals, cooperative interaction, and collective accountability, so that points, levels, and badges reward collaborative success instead of individual performance alone. In addition, the Sport Education model has been reported to enhance team cohesion, sustained cooperation, and students' interpersonal interactions (Kao, 2019). A recent study conducted in Indonesia found that cooperative circuit games effectively improved both fundamental movement skills and social-emotional learning among urban elementary school students in Yogyakarta (Kuspratiwi et al., 2026). However, these findings have not fully addressed the need for a regular physical education model that explicitly integrates gamification and team building within a unified instructional framework designed to

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

simultaneously promote FMS and SEL, rather than simply incorporating cooperative activities or engaging games.

This gap warrants further investigation because physical education has not consistently been designed to promote motor and social-emotional development at the same time. Consequently, instructional practices often address these domains separately and may be insufficiently responsive to the developmental needs of children living in increasingly digital urban environments. Accordingly, this study examined the effects of a Team Building-Integrated Gamification Model implemented in regular physical education on the fundamental movement skills and psychosocial functioning of urban elementary school students. We positioned social-emotional competence as the broader developmental framework rather than a directly measured construct. Compared with conventional instruction, we hypothesized that the intervention would lead to greater improvements in locomotor skills, ball skills, and gross motor development, as well as higher prosocial behavior. We examined its effect on psychosocial difficulties exploratorily.We used a cluster-randomized controlled design, with schools as the unit of randomization and eight public elementary schools participating.

Materials and Methods

Study Participants

This cluster-randomized controlled trial was conducted in eight public elementary schools in urban Karawang, Indonesia. Schools were randomly allocated in a 1:1 ratio by an independent researcher using a computer-generated random sequence, with four schools assigned to the experimental group and four to the control group.

An a priori sample-size calculation using G*Power 3.1 indicated that 128 individual participants were required to detect a medium effect size ($f = 0.25$, $\alpha = .05$, power = .80). To account for clustering, the sample was inflated using a design effect of 1.87, based on an average cluster size of 30 students and an assumed ICC of 0.03. Therefore, a minimum of approximately 240 students was targeted and recruited (120 per group).

Participants were fourth-grade students aged 9–10 years who regularly attended physical education classes and whose parents or legal guardians provided written informed

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

consent; children provided verbal assent. Students with conditions that precluded safe participation in group-based moderate-intensity physical activity were excluded. All participating physical education teachers held a bachelor's degree in Physical Education and had 5–12 years of teaching experience.

All 240 enrolled students completed the pre-test, intervention/control activities, and post-test assessments. No schools or students withdrew, were excluded after randomization, or were lost to follow-up; thus, all participants were included in the final analyses. No missing outcome data were identified. Minor non-serious incidents during assessment, including brief rest or water breaks due to temporary fatigue, were managed according to the study procedures and did not affect assessment completion, scoring, or data availability.

Ethical Considerations

The study protocol was approved by the Research Ethics Committee of Universitas Negeri Yogyakarta (No. B/261/UN34.16LT/2025) and the study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from parents or legal guardians using a consent form that explained the study objectives, procedures, duration, potential minor physical risks, safety measures, the voluntary nature of participation, and participants' right to withdraw at any time without academic consequences. Verbal assent was also obtained from each child before the baseline assessment and intervention.

Safety procedures implemented throughout the intervention included: (1) health readiness screening at the beginning of each session; (2) continuous teacher supervision during activities; (3) inspection of the practice area and use of safety equipment when necessary; (4) monitoring of exercise intensity using a target Borg Rating of Perceived Exertion (RPE) of 11-13; and (5) documentation of any adverse events occurring during the learning sessions. Following the post-test assessment, a debriefing session was conducted with participating schools, and the control group was provided access to the intervention protocol after all outcome measurements had been completed.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Research Design

This study employed a cluster-randomized controlled trial design, with schools serving as the unit of randomization. Baseline (pre-test) assessments were conducted one week before the intervention and included measures of fundamental movement skills and indicators of social-emotional competence. The intervention was implemented during regular physical education classes over a six-week period, comprising 12 sessions delivered twice weekly, with each session lasting 35-45 minutes for both groups. Post-test assessments were conducted one week after the final intervention session using the same procedures and instruments as those employed during the pre-test.

An independent researcher who was not involved in data collection or intervention delivery generated a computer-based random-number sequence and randomly allocated the eight schools to the experimental or control condition in a 1:1 ratio before the intervention began (four schools per condition). All participating physical education teachers held Bachelor's degrees in Physical Education and had comparable teaching experience (5–12 years). Baseline equivalence across school clusters was examined using pre-test outcome scores, with no statistically significant differences observed among clusters (all $p > .05$). Baseline intraclass correlation coefficients were also estimated using linear mixed models with random school intercepts and were low for all outcomes (ICC $\leq .021$). Nevertheless, because only eight clusters were included, school-level differences could not be fully eliminated; this limitation affects statistical precision and generalizability and is considered in the interpretation of the findings.

An a priori sample-size calculation was conducted using G*Power 3.1 for the primary group comparison, assuming a medium effect size ($f = .25$), $\alpha = .05$, and power of .80. The individual-level requirement was 128 participants. To account for clustering, this value was adjusted using the design effect, $DEFF = 1 + (m - 1) \times ICC$, assuming an average cluster size of 30 students and a conservative ICC of .03. The resulting design effect was 1.87, yielding a required sample of approximately 240 students (120 per condition). Therefore, the final sample of 240 students across eight schools met the planned sample-size requirement.

REVISTA INTERNACIONAL SPORTIS
Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

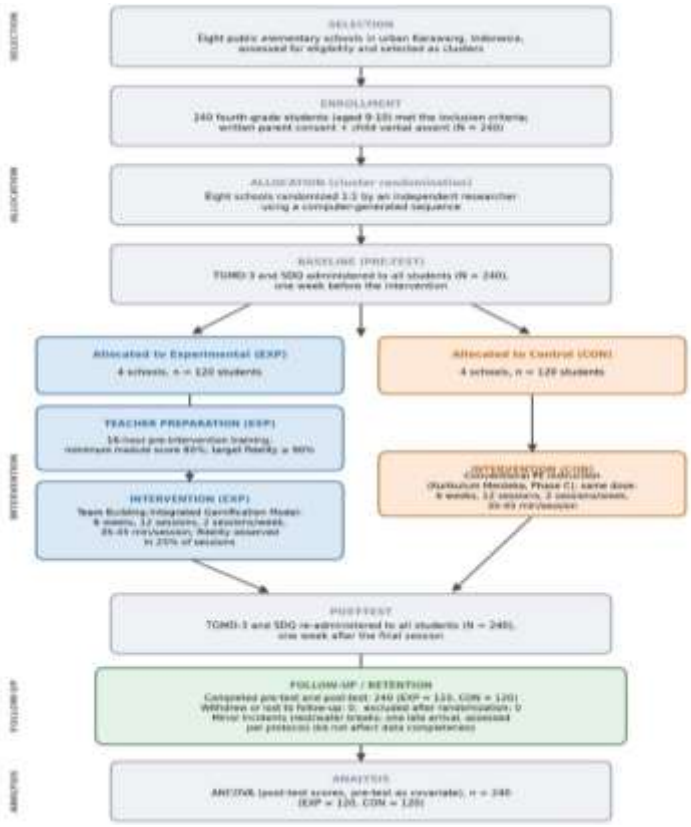


FIGURE 1. Research flow

Procedure

Experimental Group (EXP)

Learning Program Intervention

The experimental group received the Team Building-Integrated Gamification Model intervention over a six-week period, comprising 12 sessions delivered twice weekly during regular physical education classes. Each session lasted approximately 35-45 minutes. The program was systematically designed to enhance both fundamental movement skills and social-emotional competence by integrating gamification elements and team-building principles. The complete session-by-session protocol of the intervention, including the specific activities, team missions, gamification elements, and equipment of all 12 sessions, is provided in Appendix A.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Each intervention session followed a standardized four-part structure. First, a 5-10-minute warm-up was conducted, consisting of dynamic stretching, light locomotor activities, and simple cooperative exercises. Second, the main activity lasted approximately 20-25 minutes and consisted of four to five stations integrating locomotor tasks, ball skills, coordination, and basic stability exercises. Each station involved 20-40 seconds of activity followed by 20-40 seconds of rest, with three to five sets or repetitions per station, targeting a moderate exercise intensity corresponding to a Borg Rating of Perceived Exertion (RPE) of 11-13.

Third, each session incorporated a gamification challenge integrated with team-building activities in the form of group missions requiring strategic planning, role allocation, cooperative communication, peer support, and shared decision-making. Gamification was operationalized through the use of points, levels, badges, and educational leaderboards, whereas team building was facilitated through permanent heterogeneous teams of four to five students, shared group goals, rotating roles, positive interdependence, and individual accountability. Briefly, points were awarded at the team level through session-specific mechanisms defined for each of the 12 sessions in Appendix A (e.g., points counted only if every member completed the mission, bonuses for strategy, leadership, and creativity); levels marked team progression across the three program phases; and badges were cards awarded to teams for collective behaviors such as communication, technique, peer support, and mastery. The "educational leaderboard" was a cumulative team-points display, updated on a scoreboard visible to all students at the end of each session; because only team totals were shown and individual students were never ranked, it functioned as feedback on cooperative success rather than individual competition. All four experimental schools received the identical manual, session plans, teacher training, and supporting materials, and fidelity monitoring, which included checklist items on points, levels, badges, and leaderboards, confirmed 94.2% overall adherence. Finally, each session concluded with a 5-10-minute cool-down and guided reflection, consisting of light stretching followed by group reflection on cooperation, effort, sportsmanship, and emotional regulation throughout the learning activities.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

TABLE 1. Program Dose Parameters

Component	**Specification**
Program duration	6 weeks
Total sessions	12 sessions
Frequency	2 sessions per week
Duration per session	35-45 minutes
Stations per session	4-5 stations
Time per station	4-5 minutes
Transition time	30-60 seconds
Sets per station	3-5 sets/tasks
Repetitions per set	8-15 repetitions
Work duration per set	20-40 seconds
Rest duration per set	20-40 seconds
Target intensity	Moderate (Borg RPE for children: 11-13)
Team size	4-5 students per team

TABLE 2. Phase Progression with Activity Examples

Phase	**Sessions**	**Main Focus**
Phase 1: Foundation	1-4	Introduction to the model, rules, point system, and basic movement patterns
Phase 2: Consolidation	5-8	Strengthening locomotor skills, ball skills, team coordination, and peer support
Phase 3: Integration and semi-autonomy	9-12	Skill integration, student leadership, creativity, and group reflection

*The complete intervention program is presented in Appendix A.

Program Content Validity Assessment

The content validity of the Team Building-Integrated Gamification Model was evaluated through expert judgment before the intervention outcomes were analyzed. Four experts in physical education pedagogy and training participated in the validation process. Each expert held at least a master's degree in physical education or sports pedagogy and had 10-18 years of professional experience in curriculum development, elementary physical education teaching, or motor development.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

The experts independently assessed 20 predetermined indicators covering the alignment of the program with physical education learning objectives, integration of fundamental movement skills and social-emotional learning, team-building and gamification procedures, session structure, practicality, safety, and motivational quality. Each indicator was rated using a five-point Likert scale, ranging from 1 = not relevant/appropriate to 5 = highly relevant/appropriate. Aiken's V coefficient was calculated for each indicator and summarized as a mean value across the 20 indicators. A value of $V \geq 0.80$ was specified a priori as the criterion for acceptable content validity.

The internal consistency of the 20-item expert-validation instrument was assessed using Cronbach's alpha, with a coefficient of $\alpha \geq 0.70$ considered acceptable. These procedures evaluated the relevance, appropriateness, feasibility, and safety of the program design according to expert judgment. They were not intended to demonstrate the effectiveness of the intervention in improving students' motor or social-emotional outcomes.

Facilitators, Training, and Fidelity Monitoring

The physical education teachers responsible for delivering the intervention in the experimental schools completed a 16-hour pre-intervention training program before implementation. The training covered: (1) the design principles and implementation procedures of the Team Building-Integrated Gamification Model; (2) instructional techniques for teaching fundamental movement skills, including demonstration, scaffolding, and corrective feedback; (3) strategies for facilitating social-emotional learning within physical activity contexts; and (4) procedures for assessment, documentation, safety, and implementation fidelity. Teachers were required to achieve a minimum score of 80% on each training module before delivering the intervention.

Implementation fidelity was monitored using a structured observation checklist, with adherence to the protocol considered satisfactory when at least 90% of the fidelity items were implemented as planned. Fidelity monitoring included session structure, application of gamification elements, implementation of team-building strategies, monitoring of exercise intensity, quality of teacher feedback, safety procedures, and group reflection activities.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Approximately 25% of the intervention sessions were independently observed by two trained raters to evaluate inter-rater reliability.

Concrete fidelity results were as follows. Overall implementation fidelity across the observed sessions reached 94.2%, indicating high protocol adherence. Inter-observer reliability for the fidelity checklist demonstrated high inter-rater consistency (ICC = 0.89, 95% CI [0.82, 0.94])

With respect to possible differences between teachers or centers, the eight schools were randomized by an independent researcher using a computer-generated sequence (1:1 allocation), as described in the Research Design section, and baseline equivalence across school clusters was confirmed prior to the intervention (all pre-test comparisons $p > .05$; see Results). All participating teachers in both conditions held a Bachelor's degree in Physical Education and had comparable teaching experience (5-12 years), so teacher- or center-level differences are unlikely to confound the observed intervention effects.

Detailed implementation fidelity procedures and the teacher training curriculum are provided in Appendices B and D.

Control Group (CON)

The control group participated in conventional physical education instruction based on the Kurikulum Merdeka for fourth-grade students (Phase C). To ensure methodological consistency with the study design, the control group received the same instructional dose as the experimental group, consisting of six weeks, 12 sessions, two sessions per week, and 35-45 minutes per session. The primary difference between the two groups was the instructional model rather than the duration of exposure.

Instruction in the control group primarily consisted of: (1) teacher-centered direct instruction; (2) repetitive skill practice; (3) simple modified games and general physical fitness activities; and (4) minimal and unstructured integration of social-emotional learning. Concretely, each session typically began with the teacher demonstrating a movement skill, followed by whole-class imitation and repetitive individual or paired practice of locomotor skills (straight running, zig-zag running, forward jumping, galloping), ball skills (passing,

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

catching, kicking toward a target), and basic coordination and stability exercises (one-leg standing, simple bridge, coordinated stepping). Simple modified games, including mini soccer, throw-and-catch, and simple relay races, were then played under the teacher's direction, with the teacher providing verbal instruction and correction throughout. Any social-emotional elements were limited to the brief, general reflection on the day's learning during the cool-down. The control group did not use points, levels, badges, or leaderboards and did not implement an explicit team-building framework as a core instructional component. Although cooperation and sportsmanship could naturally emerge during activities, they were not systematically structured as they were in the experimental group.

A typical session therefore followed the structure presented in Table 3 and detailed in Appendix C: a 5-10-minute warm-up of walking, light jogging, light jumping, and dynamic stretching; a 20-25-minute main activity devoted to locomotor skills, ball skills, coordination, and simple modified games delivered through teacher demonstration, verbal instruction, and repetitive practice; and a 5-10-minute cool-down with light stretching and brief reflection.

Teachers in the control schools did not receive additional training related to the Team Building-Integrated Gamification Model and were instructed to maintain their usual teaching practices to preserve the ecological validity of the control condition as a representation of conventional physical education instruction in urban elementary schools.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

TABLE 3. Conventional Physical Education Session Format for the Control Group

Component	Description
Program duration	6 weeks, 12 sessions, 2 sessions/week, 35 to 45 minutes/session
Warm-up	5 to 10 minutes of walking, light jogging, light jumping, and dynamic stretching
Main activity	20 to 25 minutes of locomotor skills, ball skills, coordination, and simple modified games
Cool-down	5 to 10 minutes of light stretching and brief reflection (
Main strategy	Teacher demonstration, verbal instruction, and repetitive practice
Activities	Straight running, zig-zag running, forward jumping, galloping, passing, catching, kicking toward a target, mini soccer, throw-and-catch, and simple relay
Special features	No gamification elements, no explicit team-building structure, and socio-emotional reflection not systematically applied

*The control group instructional protocol is presented in Appendix C.

Instrument

Test of Gross Motor Development-Third Edition (TGMD-3)

Fundamental movement skills were assessed using the Test of Gross Motor Development-Third Edition (TGMD-3), a standardized instrument widely used to evaluate children's fundamental motor skills (Ulrich, 2019; Zhu et al., 2025). The TGMD-3 comprises two primary subtests: Locomotor Skills and Ball Skills (Carballo-Fazanes et al., 2021; Valentini et al., 2022). The locomotor subtest includes six movement tasks: running, galloping, hopping, skipping, horizontal jumping, and sliding. The ball skills subtest consists of seven object-control tasks: two-hand striking a stationary ball, one-hand forehand striking of a self-bounced ball, one-hand stationary dribbling, two-hand catching, kicking a stationary ball, overhand throwing, and underhand throwing (Carballo-Fazanes et al., 2021; Valentini et al., 2022). In this study, the TGMD-3 was administered and scored in its standard form: all 13 tasks, the original performance criteria, and the scoring rules of the examiner's manual (Ulrich, 2019) were applied without any modification.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Each skill was evaluated using specific performance criteria. Each criterion received a score of 1 if performed correctly and 0 if performed incorrectly. Criterion scores were summed to obtain the raw locomotor and ball skills subtest scores, which were subsequently converted into age-standardized scores and used to calculate the Gross Motor Quotient (GMQ). Specifically, raw subtest scores were converted into age-based standard scores using the normative tables of the examiner's manual (Ulrich, 2019), and the GMQ was obtained by summing the locomotor and ball skills standard scores, yielding an overall index of motor competence with a population mean of 100 and SD of 15. In the present study, subtest-level analyses (Locomotor and Ball Skills) were based on raw subtest scores, whereas overall motor competence was expressed as the age-standardized GMQ. Higher GMQ scores indicate better fundamental motor skill proficiency.

In this study, all assessors were physical education teachers who completed structured training in TGMD-3 administration and scoring procedures before data collection. To evaluate measurement consistency, approximately 20-30% of the sample was randomly selected and independently assessed by two trained raters who were blinded to both group allocation and assessment time point. Inter-rater reliability was analyzed using the intraclass correlation coefficient (ICC) based on a two-way random-effects model with absolute agreement. Concrete reliability results were as follows. Inter-rater reliability between the two independent raters was excellent, yielding an ICC of 0.92 (95% CI [0.88, 0.95]) for Locomotor skills and 0.90 (95% CI [0.85, 0.94]) for Ball skills.

Social-Emotional Assessment (Strengths and Difficulties Questionnaire [SDQ] - Indonesian Version)

Social-emotional competence was measured using the Indonesian version of the Strengths and Difficulties Questionnaire (SDQ), specifically the Teacher Report form for school-aged children. The SDQ is a widely used screening instrument for assessing children's psychosocial functioning. The questionnaire consists of 25 items distributed across five subscales: Prosocial Behavior, Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, and Peer Relationship Problems. Each item is rated on a three-point

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Likert scale (0 = Not True, 1 = Somewhat True, and 2 = Certainly True). Scores for each subscale range from 0 to 10.

The Total Difficulties Score (TDS) was calculated by summing the four problem-oriented subscales, Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, and Peer Relationship Problems, yielding a total score ranging from 0 to 40. The Prosocial Behavior subscale was analyzed separately and was not included in the TDS calculation. Higher Prosocial Behavior scores indicate better social competence, whereas higher TDS values indicate greater psychosocial difficulties.

For descriptive purposes, TDS values were categorized as normal (0-13), borderline (14-16), and abnormal (17-40). The SDQ was completed by classroom teachers at both the pre-test and post-test assessments. Teachers received standardized written instructions before data collection and completed the questionnaire independently. To minimize assessment bias, teachers were not informed of the study's specific analytical focus. Because the SDQ was completed by the classroom teachers of their own students, complete blinding to the learning condition implemented in each school was not feasible in this school-based design; teachers in the experimental schools had directly implemented the program. Expectation bias was limited as far as possible by withholding the study hypotheses and analytical comparisons from the teachers, by using standardized written instructions, and by having each teacher complete the questionnaire independently. Nevertheless, residual rater expectancy bias cannot be excluded, particularly for the teacher-rated Prosocial Behavior and Total Difficulties scores, and this limitation is acknowledged in the Discussion. Inter-rater agreement for a subset of participants was evaluated using the intraclass correlation coefficient (ICC) based on a two-way random-effects model with absolute agreement.

To minimize reporting bias, classroom teachers completed the assessments without knowledge of the specific study hypotheses or analytical comparisons. Concrete reliability results were as follows. Inter-rater reliability for the SDQ Teacher Report showed strong agreement, with ICCs of 0.86 (95% CI [0.80, 0.91]) for Total Difficulties and 0.84 (95% CI [0.77, 0.89]) for Prosocial Behavior

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Statistical Analysis

Data were analyzed using IBM SPSS Statistics version 26. Data were screened for completeness and entry accuracy, and descriptive statistics were calculated for all gross-motor and social-emotional outcomes at pre-test and post-test. Program content validity was assessed using Aiken's V based on ratings from four experts across 20 predetermined indicators, while the internal consistency of the expert-validation instrument was assessed using Cronbach's alpha. These analyses evaluated the quality of the program design and validation instrument and were not interpreted as evidence of intervention effectiveness.

Baseline equivalence between the experimental and control groups was examined using pre-test scores. Intervention effects were analyzed using Analysis of Covariance (ANCOVA), with post-test scores as dependent variables, group as the fixed factor, and the corresponding pre-test score as a covariate. The assumptions of residual normality, homogeneity of variance, and homogeneity of regression slopes were assessed using the Shapiro-Wilk test with visual inspection of residual plots, Levene's test, and group × pre-test interaction terms, respectively. Effect sizes were reported as partial eta squared η_p^2ηp2 and Cohen's dd, with statistical significance set at $\alpha = 0.05$α=0.05.

Because schools were the unit of randomization, potential clustering was examined before interpreting the individual-level ANCOVA results. Baseline intraclass correlation coefficients (ICCs) were estimated using linear mixed models with random school intercepts. The ICCs were 0.000 for TGMD-3 Gross Motor Quotient, 0.000 for Locomotor Skills, 0.021 for Ball Skills, 0.000 for SDQ Total Difficulties, and 0.000 for SDQ Prosocial Behavior. The largest ICC, 0.021 for Ball Skills, yielded a design effect of $1 + (30 - 1) \times 0.021 = 1.611$+(30−1)×0.021=1.61, corresponding to an effective sample size of approximately 149 students. This remained above the a priori individual-level requirement of 128 participants. Moreover, the original sample-size calculation used the more conservative assumption of an ICC of 0.03 and a design effect of 1.87 (Hedges & Hedberg, 2007).

Given the negligible observed ICCs, the conservative sample-size calculation, and the limited number of randomized schools, individual-level ANCOVA was retained as the primary

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

analysis. Nevertheless, the 240 students were not treated as completely independent observations, and the possibility of slightly anti-conservative standard errors and inflated statistical significance cannot be entirely excluded. The reported η_p^2 and d values therefore represent individual-level intervention effects; a separate cluster-level effect size was not estimated in the present analysis. This issue is acknowledged as an important limitation, and future studies should include a larger number of schools and use multilevel or cluster-level analyses as the primary analytical approach.

Results

Program Content Validity and Reliability

Content validity was assessed by four experts in physical education pedagogy and training. The experts held at least a master's degree in physical education or sports pedagogy and had 10-18 years of professional experience in curriculum development, elementary physical education teaching, or motor development. They independently rated 20 program indicators using a five-point Likert scale ranging from 1 = not relevant/appropriate to 5 = highly relevant/appropriate.

The mean Aiken's V was 0.93, with indicator-level values ranging from 0.85 to 1.00. All 20 indicators met the predefined acceptance criterion of $V \geq 0.80$, indicating favorable expert judgment regarding the content relevance, developmental appropriateness, feasibility, and safety of the program. The internal consistency of the expert-validation instrument was acceptable, with Cronbach's alpha of $\alpha = 0.738$.

These findings demonstrate the content validity and internal consistency of the program-validation instrument only. They do not constitute evidence that the intervention improves fundamental movement skills or social-emotional competence. Intervention effectiveness was examined separately through the pre-test and post-test outcome analyses and the ANCOVA results reported in the subsequent sections.Detailed validation and reliability results are provided in Tables 4 and 5.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

TABLE 4. Validation of the Team Building-Integrated Gamification Model Program

No.	Rated Aspect	Σs	V	Description
CONTENT/MATERIAL ALIGNMENT				
1	Alignment of the Team Building-Integrated Gamification Model program with Physical Education learning objectives and Kurikulum Merdeka Phase C for Grade 4 elementary students	20	1.00	Accepted
2	Integration of fundamental movement skill development (locomotor, ball skills, stability) with gamification elements (points, levels, badges, educational leaderboard) in program design	19	0.95	Accepted
3	Inclusion of social-emotional learning components through team-building strategies (prosocial behavior, emotion regulation, cooperative communication, problem-solving) appropriate to students' developmental stages	19	0.95	Accepted
4	Appropriateness of team-based activity selection (heterogeneous teams, specific role assignment, positive interdependence, individual accountability) with gamification mechanics for Grade 4 students	18	0.90	Accepted
5	Potential of the program to enhance locomotor skills (running, jumping, galloping, hopping, sliding) through structured circuit activities with progressive team missions	18	0.90	Accepted
6	Potential of the program to enhance ball skills (passing, catching, throwing, kicking) through cooperative stations with accuracy-based point systems	19	0.95	Accepted
7	Potential of the program to enhance prosocial behavior (cooperation, helping, sharing, peer support) through team goals and collective scoring mechanisms	20	1.00	Accepted
8	Potential of the program to enhance cooperative communication skills (verbal instruction, encouragement, simple feedback) through team building roles and structured reflection phases	17	0.85	Accepted
9	Potential of the program to enhance cooperative problem-solving skills through strategic team challenges and decision-making tasks in circuit activities	19	0.95	Accepted
10	Potential of the program to enhance emotion regulation and resilience through retry challenges and supportive team environments in physical activity situations	17	0.85	Accepted
PROGRAM CONSTRUCTION				
11	Appropriateness of session structure (5-10 min warm-up with cooperative ice breaking, 20-25 min core circuit activity, 3-5 min cool-down, 2-5 min group reflection) for Grade 4 students within 35-45 min per session	20	1.00	Accepted
12	Design of four to five circuit stations with appropriate station duration (4-5 min), work-rest ratio (1:1, 20-40 seconds each), 30-60 second transitions, and 3-5 sets per station based on child training principles	19	0.95	Accepted

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

No.	Rated Aspect	Σs	V	Description
13	Systematic program progression from Foundation (Sessions 1-4: model introduction, game rules, point system, basic movement patterns), Consolidation (Sessions 5-8: locomotor reinforcement, ball skills, team coordination, peer support), to Integration and Semi-Autonomy (Sessions 9-12: skill integration, student leadership, creativity, group reflection)	18	0.90	Accepted
14	Suitability of the implementation guide (activity instructions, team management strategies, gamification mechanics, RPE monitoring at Borg 11-13) and implementation fidelity checklist (minimum 90% core components) for elementary PE teachers	17	0.85	Accepted
15	Provision of opportunities for students to observe demonstrations of motor skills, cooperative strategies, and team missions at each circuit station	19	0.95	Accepted
16	Provision of opportunities for student questioning, discussion, and group reflection about effort, cooperation, and prosocial behavior during reflection phases and inter-station transitions	18	0.90	Accepted
17	Provision of opportunities for students to actively practice motor skills within cooperative team contexts supported by gamification feedback (points, badges, levels)	18	0.90	Accepted
PROGRAM ERGONOMICS/PRACTICALITY				
18	Use of equipment (cones, markers, small/medium/soft balls, coordination ladders, bean bags, ropes, low balance beams, floor lines, mats, stopwatch, score boards, badge cards, observation sheets) that is easily obtainable or adaptable to school conditions	17	0.85	Accepted
19	Safety of equipment and activities, including mat use, safety zones, supervision procedures, and RPE 11-13 intensity monitoring through structured work-rest intervals	20	1.00	Accepted
20	Attractiveness and motivational quality of gamified team-based circuit activities (points, badges, team missions, leadership rotation, educational leaderboard) for active participation among Grade 4 students	19	0.95	Accepted
	Mean	18.55	0.93	Accepted

TABLE 5. Reliability of the Team Building-Integrated Gamification Model

Cronbach's Alpha	N of Items
0.738	20

These content validity findings reflect expert judgments regarding the relevance, appropriateness, and feasibility of the program design. They should not be interpreted as evidence of intervention effectiveness. The effectiveness of the Team Building-Integrated

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Gamification Model was evaluated separately through the pre-test and post-test outcome analysis presented in the subsequent sections.

Measurement Reliability and Intervention Fidelity

Inter-rater reliability for the TGMD-3 was excellent. The ICC was 0.92 (95% CI [0.88, 0.95]) for Locomotor Skills and 0.90 (95% CI [0.85, 0.94]) for Ball Skills. Inter-rater reliability for the SDQ Teacher Report was also strong, with an ICC of 0.86 (95% CI [0.80, 0.91]) for Total Difficulties and 0.84 (95% CI [0.77, 0.89]) for Prosocial Behavior. These ICC values indicate strong agreement between independent raters for the outcome measures.

Implementation fidelity was monitored using a structured 20-item observation checklist. Approximately 25% of the intervention sessions were independently observed by two trained raters. Overall implementation fidelity across the observed sessions was 94.2%, indicating high adherence to the intervention protocol. Inter-rater reliability for the fidelity checklist was high, with an ICC of 0.89 (95% CI [0.82, 0.94]). These results indicate high consistency of measurement and high adherence to the intervention protocol across the observed sessions.

Descriptive Results

TABLE 6. Descriptive Statistics of Study Variables

Variable	Group	Pre-test M ± SD	Post-test M ± SD
Locomotor	EXP	7.89 ± 1.90	9.48 ± 2.52
	CON	5.75 ± 1.65	8.87 ± 2.11
Ball Skills	EXP	9.04 ± 1.84	10.97 ± 2.33
	CON	6.69 ± 1.68	10.00 ± 2.05
GMQ	EXP	96.28 ± 5.31	105.47 ± 5.87
	CON	90.36 ± 5.16	97.23 ± 5.42
SDQ Total Difficulties	EXP	14.18 ± 3.79	11.32 ± 4.25
	CON	16.64 ± 3. 93	13.38 ± 3.84
SDQ Prosocial	EXP	6.18 ± 2.00	7.53 ± 2.23
	CON	4.32 ± 1.81	7.07 ± 2.11

Note. EXP = experimental group; CON = control group; GMQ = Gross Motor Quotient; SDQ = Strengths and Difficulties Questionnaire.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Table 6 presents descriptive statistics for all study variables at pre-test and post-test for each group. Overall, the experimental group improved across all indicators of fundamental movement skills and prosocial behavior, along with a reduction in Total Difficulties scores. In contrast, changes in the control group were smaller. All 240 students completed the study (experimental group = 120; control group = 120), with no attrition during the intervention period. No missing data were identified, and descriptive inspection did not indicate values outside the theoretical ranges of the instruments. Descriptive baseline comparisons showed no statistically significant differences between the experimental and control groups on the measured pre-test variables (all $p > .05$).

To verify baseline equivalence across the eight participating school clusters, we performed a series of one-way Analyses of Variance (ANOVA) on pre-test scores. The findings showed no statistically significant differences among the school clusters across all baseline measures: TGMD-3 Locomotor ($F(7, 232) = 0.362$, $p = .923$), TGMD-3 Ball Skills ($F(7, 232) = 1.633$, $p = .127$), TGMD-3 GMQ ($F(7, 232) = 0.924$, $p = .488$), SDQ Total Difficulties ($F(7, 232) = 0.872$, $p = .529$), and SDQ Prosocial ($F(7, 232) = 0.784$, $p = .602$). The results confirm that initial motor skills and behavioral characteristics were homogeneous across all school clusters before the intervention.

Baseline Cluster Equivalence and Intraclass Correlations

Baseline comparisons across the eight school clusters showed no statistically significant differences for any outcome: TGMD-3 Locomotor, $F(7, 232) = 0.362$, $p = .923$; TGMD-3 Ball Skills, $F(7, 232) = 1.633$, $p = .127$; TGMD-3 GMQ, $F(7, 232) = 0.924$, $p = .488$; SDQ Total Difficulties, $F(7, 232) = 0.872$, $p = .529$; and SDQ Prosocial Behavior, $F(7, 232) = 0.784$, $p = .602$.

Baseline ICCs estimated from random-intercept linear mixed models ranged from 0.0001 to 0.021, indicating that little outcome variability was attributable to differences between schools. The largest ICC was observed for Ball Skills ($ICC = 0.021$). With an average cluster size of 30 students, this ICC yielded an estimated design effect of 1.61 and an effective sample size of approximately 149 students ($240/1.61$). This effective sample size exceeded the

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

a priori individual-level requirement of 128 participants. In addition, the original sample-size calculation assumed a more conservative ICC of 0.03, corresponding to a design effect of 1.87.

Although the observed baseline ICCs were small, the 240 students could not be considered fully independent observations because schools were the unit of randomization. Therefore, the individual-level ANCOVA findings should be interpreted cautiously, with consideration of the ICC estimates and the limited number of randomized schools ($n = 8$).

Normality Results

Residual normality in the ANCOVA models (post-test scores controlling for pre-test scores) was assessed using the Shapiro-Wilk test, which produced test statistics ranging from 0.898 to 0.992 and significance values from .006 to .863. Several residuals yielded statistically significant results ($p < .05$), indicating deviations from normality. Visual inspection of Q-Q plots and histograms indicated that these deviations were not severe and remained within acceptable limits. Given the relatively large sample size in each group (n = 120) and established evidence that parametric procedures, including ANCOVA, are generally robust to moderate violations of normality, especially in the presence of large F values for the primary outcomes, subsequent analyses were deemed appropriate. This deviation is nevertheless acknowledged as a limitation.

Homogeneity of Regression Slopes

TABLE 7. Homogeneity of Regression Slopes

Variable	F Group × Pretest	p
TGMD-3 Locomotor	0.001	0.977
TGMD-3 Ball Skills	1.873	0.176
TGMD-3 GMQ	0.087	0.769
SDQ Total Difficulties	0.176	0.675
SDQ Prosocial	3.021	0.088

NOTE. No significant interaction was found between group and pre-test scores for any variable (all $p > .05$), indicating that the assumption of homogeneity of regression slopes was satisfied.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Table 7 presents the results of the homogeneity of regression slopes test. No significant interaction was found between group and pre-test scores for any outcome variable (all p >.05), indicating that the assumption of homogeneity of regression slopes was met. Therefore, the use of ANCOVA to compare post-test scores between groups while controlling for pre-test scores was considered methodologically appropriate.

Homogeneity of Variance

TABLE 8. Results of the Homogeneity of Variance Test

Variable	F (1, 238)	p	Interpretation
TGMD-3 Locomotor	3.763	0.054	Homogeneous
TGMD-3 Ball Skills	0.684	0.409	Homogeneous
TGMD-3 GMQ	0.531	0.467	Homogeneous
SDQ Total Difficulties	0.441	0.507	Homogeneous
SDQ Prosocial	0.637	0.426	Homogeneous

NOTE. GMQ = Gross Motor Quotient; SDQ = Strengths and Difficulties Questionnaire. The assumption of homogeneity of variances was considered satisfied when p > 0.05.

Table 8 presents the results of Levene's test for homogeneity of variances. No statistically significant evidence of unequal variances was observed for any of the five outcome variables (all p >.05). The TGMD-3 Locomotor outcome approached the conventional significance threshold, F(1, 238) = 3.763, p =.054; however, this value remained above the predefined α =.05 criterion. Given the balanced sample sizes between groups, the homogeneity of variance assumption was considered satisfied for the ANCOVA analyses.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

TABLE 9. Individual-Level ANCOVA Results with Adjusted Mean Differences

Variable	Group F (1,237)	p	η^2p	Adj. Mean EXP	Adj. Mean CON	Cohen's d	95% CI for adjusted mean difference
TGMD-3 Locomotor	121.718	< .001	0.339	10.099	8.243	1.43	[1.53, 2.18]
TGMD-3 Ball Skills	58.756	< .001	0.199	11.194	9.740	1.00	[1.08, 1.83]
TGMD-3 GMQ	47.957	< .001	0.168	103.874	98.826	0.90	[3.61, 6.48]
SDQ Total Difficulties	3.135	0.078	0.013	12.177	12.523	-0.23	[-0.73, 0.04]
SDQ Prosocial	75.996	< . 001	0.243	7.979	6.613	1.13	[1.06, 1.67]

NOTE. EXP = experimental group; CON = control group; GMQ = Gross Motor Quotient; SDQ = Strengths and Difficulties Questionnaire. Adjusted Mean = mean adjusted for pre-test scores. The negative Cohen's d value for SDQ Total Difficulties indicates the expected direction of change (i.e., a reduction in psychosocial difficulties).

Table 9 summarizes the primary ANCOVA results, reporting F values, significance levels, partial eta-squared effect sizes, and adjusted means for each group. After adjusting for pre-test scores, ANCOVA identified significant group effects for TGMD-3 Locomotor, $F(1, 237) = 121.718$, $p < .001$, $\eta^2p = .339$; TGMD-3 Ball Skills, $F(1, 237) = 58.756$, $p < .001$, $\eta^2p = .199$; TGMD-3 GMQ, $F(1, 237) = 47.957$, $p < .001$, $\eta^2p = .168$; and SDQ Prosocial Behavior, $F(1, 237) = 75.996$, $p < .001$, $\eta^2p = .243$. In each of these outcomes, the experimental group obtained significantly higher adjusted scores than the control group. For SDQ Total Difficulties, the group effect was not statistically significant, $F(1, 237) = 3.135$, $p = .078$, $\eta^2p = .013$, and the 95% confidence interval for the adjusted mean difference (−0.73 to 0.04) included zero. While the difference favored the experimental group (indicating lower difficulties), this should not be interpreted as evidence of a reliable intervention effect.

The adjusted mean differences between groups were 1.86 points for Locomotor (95% CI [1.53, 2.18]), 1.45 points for Ball Skills (95% CI [1.08, 1.83]), 5.05 points for GMQ (95% CI [3.61, 6.48]), and 1.37 points for Prosocial Behavior (95% CI [1.06, 1.67]). For SDQ Total Difficulties, the adjusted mean difference was −0.35 points (95% CI [−0.73, 0.04]). The corresponding individual-level Cohen's d values were 1.43 for Locomotor, 1.00 for Ball Skills, 0.90 for GMQ, 1.13 for Prosocial Behavior, and −0.23 for SDQ Total Difficulties. The

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

confidence intervals presented in Table 9 pertain to adjusted mean differences rather than Cohen's d values.

The largest partial eta-squared value was found for TGMD-3 Locomotor (η^2p = .339), followed by SDQ Prosocial Behavior (η^2p = .243), both of which represent large effects according to conventional individual-level ANCOVA benchmarks. However, as the primary ANCOVA was conducted at the student level and did not account for the school-level hierarchy, these effect sizes should not be interpreted as cluster-adjusted estimates. It is important to distinguish statistical significance from practical or educational relevance. Specifically, the difference in SDQ Total Difficulties did not reach statistical significance, and even if it had, a between-group difference in questionnaire scores alone would not indicate clinical improvement or a change in students' diagnostic status.

In summary, the Team Building-Integrated Gamification Model was associated with statistically significant improvements in fundamental movement skills and prosocial behavior. In contrast, the effect on total psychosocial difficulties was not statistically significant. These effects should be interpreted with caution, as the reported ANCOVA effect sizes are individual-level estimates, whereas the school served as the unit of randomization

Discussion

This cluster randomized trial evaluated whether a Team Building-Integrated Gamification Model would surpass conventional physical education instruction in enhancing elementary students' fundamental movement skills and psychosocial functioning. The findings supported the first hypothesis and partially supported the second: after adjusting for pre-test scores, the experimental group demonstrated significantly greater locomotor skills, ball skills, GMQ, and prosocial behavior. However, the reduction in psychosocial difficulties did not achieve statistical significance. Three interpretive points warrant consideration: the mechanisms underlying motor gains, the implications and limitations of the psychosocial results, and the credibility of the estimates given the cluster design.

The observed motor improvements align with previous research. Interventions targeting fundamental movement skills that integrate structured, repeated practice with enjoyable

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

learning experiences are more likely to promote children's motor development (Koolwijk et al., 2023; Li et al., 2021; Ha et al., 2021). Additionally, gamification in physical education has been shown to increase motivation, engagement, and commitment to physical activity (Arufe-Giráldez et al., 2022; Sotos-Martínez et al., 2023). The current model incorporated both elements: students engaged in movement tasks individually and through group missions, role rotation, and collaborative challenges, which likely enhanced the volume and quality of task-relevant practice. The substantial effects observed (η^2p = .168 to .339) surpass those typically reported for conventional instruction, although direct comparisons are constrained by variations in study design and measurement instruments.

The prosocial finding extends this pattern to the psychosocial domain and is coherent with evidence that sport participation promotes prosocial behavior in children and adolescents (Li & Shao, 2022; Li, 2021; Carreres-Ponsoda et al., 2021) and with physical education models emphasizing personal and social responsibility (Carreres-Ponsoda et al., 2021; Sindiani et al., 2025). The shared goals, cooperative communication, peer support, and guided reflection embedded in the design plausibly created structured opportunities for prosocial behavior, a mechanism also highlighted in reviews of cooperative learning (Zhou & Colomer, 2024). By contrast, the effect on SDQ Total Difficulties was not significant ($p = .078$, $d = -0.23$), and its confidence interval included zero. Three explanations are plausible: six weeks may be too short to shift broad psychosocial difficulties; Total Difficulties is a composite of emotional, conduct, hyperactivity, and peer problems and is therefore less sensitive to a movement-based intervention than specific prosocial indicators; and the small effect suggests limited statistical power for this outcome. More fundamentally, the SDQ is a screening instrument for psychosocial functioning, not a comprehensive measure of social-emotional competence, so these findings speak to prosocial behavior and psychosocial difficulties rather than social-emotional competence in general.

Two critical considerations qualify these interpretations. First, because the teachers who implemented the intervention also completed the SDQ and were not blinded to group allocation, expectation bias cannot be ruled out. This potential bias primarily affects the teacher-rated psychosocial outcomes, whereas the TGMD-3 was evaluated using standardized performance

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

criteria. Second, the pedagogical advantage of the model likely results from the intentional integration of gamification with structured team building, aligning with cooperative learning and Sport Education approaches that promote positive interdependence and collaboration (Fernández-Río & Méndez-Giménez, 2016; Kao, 2019; Schulze & Huth, 2022; Sotos-Martínez et al., 2023; Sindiani et al., 2025), rather than from gamification elements alone. While the program's previously reported content validity (Aiken's V) supports the relevance of its components, it does not establish effectiveness. Claims of effectiveness are based solely on the outcome differences observed in this study, and the current design does not permit causal attribution to specific design components.

The primary limitation pertains to the cluster-randomized design: schools served as the unit of randomization, with only eight clusters available, and the individual-level ANCOVA did not fully account for school-level clustering. As a result, standard errors and effect sizes may be biased, and statistical significance may be overstated. Although baseline intraclass correlation coefficients (ICCs) were very low (ICC $\leq$ 0.021) and the effective sample size exceeded the a priori requirement, ICC estimates based on eight clusters are imprecise. Additional limitations include the lack of blinding for SDQ raters, the study's restriction to a single urban area with predominantly upper-middle socioeconomic backgrounds (limiting generalizability to other urban or rural populations), the six-week intervention duration, and the absence of follow-up assessments. Future research should employ multilevel analyses with a greater number of clusters, utilize blinded or independent outcome assessors, implement longer interventions, and include follow-up assessments. Further investigation should also address underlying mechanisms such as motivation, engagement, quality of peer interaction, and implementation fidelity.

Conclusions

This study provides preliminary evidence that the Team Building-Integrated Gamification Model may improve elementary students' fundamental movement skills and prosocial behavior compared with conventional instruction, whereas the effect on psychosocial difficulties was not statistically significant. The main contribution is a pedagogically integrated

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

design, combining gamification with structured team building, that simultaneously targets motor and prosocial outcomes within regular physical education. However, the evidence derives from only eight schools in one urban area with specific socioeconomic characteristics, a six-week intervention, and no follow-up, so the findings apply to the studied population and should not be broadly generalized. Future research should employ multilevel analyses with more clusters, blinded outcome assessment, longer intervention periods, and follow-up assessments to confirm and extend these findings.

Acknowledgements

The authors express their gratitude to the participating elementary schools, teachers, students, and parents/guardians for their support and involvement in this study. Appreciation is also extended to school administrators and colleagues who granted permission, facilitated the research, and provided valuable feedback throughout the research process.

References

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

Aygün, Y., Böke, H., Yağın, F. H., Tüfekçi, Ş., Murathan, T., Gençay, E., Prieto-González, P., & Ardigò, L. P. (2024). Emotional and social outcomes of the teaching personal and social responsibility model in physical education: A systematic review and meta-analysis. Children, 11, 459. https://doi.org/10.3390/children11040459

Böke, H., Aygün, Y., Tüfekçi, Ş., Yağın, F. H., Canpolat, B., Norman, G., Prieto-González, P., & Ardigò, L. P. (2025). Effects of cooperative learning on students' learning outcomes in physical education: A meta-analysis. *Frontiers in Psychology, 16*, 1508808. https://doi.org/10.3389/fpsyg.2025.1508808

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

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

Carballo-Fazanes, A., Rey, E., Valentini, N. C., Rodríguez-Fernández, J. E., Varela-Casal, C., Rico, J., Barcala-Furelos, R., & Abelairas-Gómez, C. (2021). Intra-rater (live vs. video assessment) and inter-rater (expert vs. novice) reliability of the *Test of Gross Motor Development-Third Edition*. *International Journal of Environmental Research and Public Health, 18*(4), 1652. https://doi.org/10.3390/ijerph18041652

Carreres-Ponsoda, F., Escartí, A., Jiménez-Olmedo, J. M., & Cortell-Tormo, J. M. (2021). Effects of a Teaching Personal and Social Responsibility Model intervention in competitive youth sport. *Frontiers in Psychology, 12*, 624018. https://doi.org/10.3389/fpsyg.2021.624018

Chen, H., & Yu, Y. (2022). The impact of social-emotional learning: A meta-analysis in China. *Frontiers in Psychology, 13*, 1040522. https://doi.org/10.3389/fpsyg.2022.1040522

Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., Sehgal, K., Zieher, A. K., Eveleigh, A., McCarthy, M., Funaro, M., Ponnock, A., Chow, J. C., & Durlak, J. A. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. Child Development, 94(5), 1181-1204. https://doi.org/10.1111/cdev.13968

El-Tanahi, N., Soliman, M., Abdel Hady, H., Alfrehat, R., Faid, R., Abdelmoneim, M., Torki, M., & Hamoudah, N. (2024). The effectiveness of gamification in physical education: A systematic review. *International Journal of Education in Mathematics, Science and Technology, 12*(2), 406-417. https://doi.org/10.46328/ijemst.4005

Engel, A. C., Broderick, C. R., van Doorn, N., Hardy, L. L., & Parmenter, B. J. (2018). Exploring the relationship between fundamental motor skill interventions and physical activity levels in children: A systematic review and meta-analysis. Sports Medicine, 48(9), 2211-2226. https://doi.org/10.1007/s40279-018-0923-3

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

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-Río, J., & Méndez-Giménez, A. (2016). El aprendizaje cooperativo: Modelo pedagógico para educación física. Retos, 29, 201-206. https://doi.org/10.47197/retos.v0i29.38721

García-Hermoso, A., Alonso-Martínez, A. M., Ramírez-Vélez, R., Pérez-Sousa, M. Á., Ramírez-Campillo, R., & Izquierdo, M. (2020). Association of physical education with improvement of health-related physical fitness outcomes and fundamental motor skills among youths. JAMA Pediatrics, 174(6),Article. https://doi.org/10.1001/jamapediatrics.2020.0223

Ha, A. S., Lonsdale, C., Lubans, D. R., Ng, F. F.-Y., & Ng, J. Y. Y. (2021). Improving children's fundamental movement skills through a family-based physical activity program: Results from the "Active 1 + FUN" randomized controlled trial. *International Journal of Behavioral Nutrition and Physical Activity, 18*, 99. https://doi.org/10.1186/s12966-021-01160-5

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis, 29*(1), 60–87. https://doi.org/10.3102/0162373707299706

Jones, D., Innerd, A., Giles, E. L., & Azevedo, L. B. (2020). Association between fundamental motor skills and physical activity in the early years: A systematic review and meta-analysis. Journal of Sport and Health Science, 9(6), 542-552. https://doi.org/10.1016/j.jshs.2020.03.001

Kao, C.-C. (2019). Development of team cohesion and sustained collaboration skills with the sport education model. Sustainability, 11(8), 2348. https://doi.org/10.3390/su11082348

Khan, A., Uddin, R., & Burton, N. W. (2018). Insufficient physical activity in combination with high screen time is associated with adolescents' psychosocial difficulties. International Health, 10(4), 246-251. https://doi.org/10.1093/inthealth/ihy019

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Koolwijk, P., Hoeboer, J., Mombarg, R., Savelsbergh, G. J. P., & de Vries, S. (2023). Fundamental movement skill interventions in young children: A systematic review. International Journal of Sport and Exercise Psychology, 22(7), 1661-1683. https://doi.org/10.1080/1612197X.2023.2210597

Kuspratiwi, F. W., Suherman, W. S., Komari, A., & Nevitaningrum, N. (2026). The effects of Cooperative Circuit Games on fundamental motor skills and social-emotional learning in elementary school students. Pedagogy of Physical Culture and Sports, 30(2), 93-110. https://doi.org/10.15561/26649837.2026.0203

Li, B., Liu, J., & Ying, B. (2021). Physical education interventions improve the fundamental movement skills in kindergarten: A systematic review and meta-analysis. *Food Science and Technology, 42*, e46721. https://doi.org/10.1590/fst.46721

Li, J., & Shao, W. (2022). Influence of sports activities on prosocial behavior of children and adolescents: A systematic literature review. International Journal of Environmental Research and Public Health, 19(11), 6484. https://doi.org/10.3390/ijerph19116484

Liu, W., Wu, X., Huang, K., Yan, S., Ma, L., Cao, H., Gan, H., & Tao, F. (2021). Early childhood screen time as a predictor of emotional and behavioral problems in children at 4 years: A birth cohort study in China. Environmental Health and Preventive Medicine, 26, 3. https://doi.org/10.1186/s12199-020-00926-w

Mery, M., Martono, M., Halidjah, S., & Hartoyo, A. (2022). Sinergi peserta didik dalam Proyek Penguatan Profil Pelajar Pancasila. Jurnal Basicedu, 6(5), 7840-7849. https://doi.org/10.31004/basicedu.v6i5.3617

Muppalla, S. K., Vuppalapati, S., Pulliahgaru, A. R., & Sreenivasulu, H. (2023). Effects of excessive screen time on child development: An updated review and strategies for management. *Cureus, 15*(6), e40608. https://doi.org/10.7759/cureus.40608

Schulze, C., & Huth, M. V. (2022). Effects of cooperative learning structures in physical education: A systematic review. International Journal of Physical Education, Fitness and Sports, 11(4), 1-14. https://doi.org/10.34256/ijpefs2241

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Sindiani, M., Schroeder, H. B., & Dunsky, A. (2025). Social-emotional learning in physical education classes at elementary schools. *Frontiers in Psychology, 16*, 1499240. https://doi.org/10.3389/fpsyg.2025.1499240

Sotos-Martínez, V. J., Tortosa-Martínez, J., Baena-Morales, S., & Férriz-Valero, A. (2023). Boosting student's motivation through gamification in physical education. *Behavioral Sciences, 13*(2), 165. https://doi.org/10.3390/bs13020165

Stevanović, D., Urbán, R., Atilola, O., Vostanis, P., Balhara, Y. P. S., Avicenna, M., Kandemır, H., Knez, R., Franić, T., & Petrov, P. (2014). Does the Strengths and Difficulties Questionnaire self report yield invariant measurements across different nations? Data from the International Child Mental Health Study Group. Epidemiology and Psychiatric Sciences, 24(4), 323-334. https://doi.org/10.1017/S2045796014000201

Stiglic, N., & Viner, R. M. (2019). Effects of screentime on the health and well-being of children and adolescents: A systematic review of reviews. BMJ Open, 9(1),Article. https://doi.org/10.1136/bmjopen-2018-023191

Sutomo, R., Ramadhani, F. P. R., & Hanifa, I. N. (2022). Prevalence and associated factors of psychosocial and behavioral problems in Indonesian adolescent students during the COVID-19 pandemic. Frontiers in Pediatrics, 10, 908384. https://doi.org/10.3389/fped.2022.908384

Ulrich, D. A. (2019). *Test of Gross Motor Development, Third Edition: Examiner's manual* (3rd ed.). PRO-ED. https://proedinc.com/products-14805.html

Valentini, N. C., Duarte, M. G., Zanella, L. W., & Nobre, G. C. (2022). Test of Gross Motor Development-3: Item difficulty and item differential functioning by gender and age with Rasch analysis. *International Journal of Environmental Research and Public Health, 19*(14), 8667. https://doi.org/10.3390/ijerph19148667

Wagner, M., Webster, E. K., & Ulrich, D. A. (2016). Psychometric properties of the Test of Gross Motor Development, Third Edition (German translation): Results of a pilot study. Journal of Motor Learning and Development, 4(2), 293-304. https://doi.org/10.1123/jmld.2016-0006

Original article. Cluster randomized trial of team-building gamified physical education for urban students' movement and social-emotional competence.
Vol. 12, n.º **6**; p. 1-34, October 2026. https://doi.org/10.55040/6qhzb671

Wimbarti, S., Siregar, J., Oktaviana, M., & Regiastri, R. (2019). Strengths and Difficulties Questionnaire Parent Report (SDQ-PR) as screening instrument of children mental health in Indonesia. Jurnal Psikologi, 46(2), 101-116. https://doi.org/10.22146/jpsi.46698

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*, 1522. https://doi.org/10.1186/s12889-025-22696-2

Zhu, Y., Wang, J., Ding, Y., Qian, Y., Korivi, M., Chen, Q., & Ye, W. (2025). Assessing the measurement properties of the *Test of Gross Motor Development-3* using the COSMIN methodology: A systematic review. *Behavioral Sciences, 15*(1), 62. https://doi.org/10.3390/bs15010062

Zhou, T., & Colomer, J. (2024). Cooperative learning promoting cultural diversity and individual accountability: A systematic review. Education Sciences, 14, 567. https://doi.org/10.3390/educsci14060567