



Reconciling model fidelity and teaching agency: How preservice teachers' implementation of a hybrid Sport Education-Step Game Approach relates to student learning

Cláudio Farias¹ , Ellen-Alyssa F Gambles² , Jacob Sierra-Diaz³ , Sixto González-Villora⁴ , and Isabel Mesquita¹ 

Abstract

This study examined how different levels of implementation fidelity to critical elements of a hybrid Sport Education–Step Game Approach (SE-SGA) were associated with students' game performance during volleyball teaching. Within a master's-level physical education teacher education program in Northern Portugal, nine preservice teachers (PSTs) and 189 10th-grade students across nine classes participated. Prior to school placement, the PSTs completed over 300 hours of coursework, including 60 hours of sport-specific pedagogy using the SE-SGA hybrid approach. A quasi-experimental pre–post design was employed. 108 lessons were video-recorded and coded using a 17-item fidelity checklist. Fidelity scores were averaged across 12 lessons per PST. Cluster analysis (Ward's method) identified three fidelity levels (high, moderate, low). One-way analysis of variance (ANOVA) and post hoc tests confirmed significant group differences. Students' game performance was assessed using validated measures: decision-making efficacy (0–3 scale), game performance index (mean across four actions), game involvement score (total actions), rate of play (actions/minute), and efficacy score (successful actions). Multivariate analysis of variance, ANOVA, and multiple mediation analyses were conducted. Results showed that students taught by high-fidelity PSTs achieved the greatest gains across all domains (game involvement, efficacy,

¹University of Porto, Portugal

²University of Sunderland, UK

³Universidad de Oviedo, Spain

⁴Universidad de Castilla-La Mancha, Spain

Corresponding author:

Cláudio Farias, Centre for Research, Education, Intervention and Innovation in Sport (CIFI2D), Faculty of Sport, University of Porto, Rua Doutor Plácido Costa 91, 4200-450 Porto, Portugal.

Email: claudiofarias@fade.up.pt

rate of play, performance index, and discrete actions: service, reception, setting, attack). Moderate- and low-fidelity groups showed improvements in specific indicators, particularly participation and technical actions. These findings challenge binary fidelity views and highlight how scaffolded, context-sensitive enactment may support learning. These results require further study through in situ analyses of PST–mentor interactions.

Keywords

Model-based pedagogy, physical education, game-based approaches, game performance

Introduction

Student-centered approaches, such as Sport Education (SE) (Siedentop et al., 2019) and game-based approaches (GBAs) (Farias et al., 2025), have become central to models-based practice in physical education (MBP-PE) (Casey and Kirk, 2020). While SE fosters holistic development, encouraging students to become motivated participants and stewards of a healthy sporting culture, it shares with GBAs the goal of developing competent sportspeople. These approaches have expanded PE's potential to promote autonomy and meaningful engagement, supporting cognitive, psychological, socio-affective, and motor outcomes, particularly in sport-based contexts (Farias et al., 2020; Hastie et al., 2011).

This study examines how preservice teachers (PSTs) implemented a hybridized student-centered approach, focusing on model fidelity and its relation to students' game performance. Previous research has examined how implementation within MBP-PE relates to instructional content and student learning outcomes (Hastie and Casey, 2014). The current study thus engages with two key tensions in the field. First, MBP-PE often assumes that only full implementation of a model's critical features ensures valid attribution of student learning outcomes (Iglesias and Fernandez-Rio, 2024). Yet learning to teach, especially during school placement, is a complex, non-linear process, often misaligned with the uncritical application of predefined pedagogical prescriptions (Farias et al., 2023). During placement, PSTs require a solid foundation, including content knowledge and pedagogical competence in models-based instruction, to navigate their first real teaching experience (Ward and Ayvazo, 2016); otherwise, they risk instructional drift (McMahon and MacPhail, 2007). As teaching is shaped by students' needs, PSTs' evolving competencies, and contextual factors, they benefit from physical education teacher education (PETE) support that scaffolds development through responsive mentoring and fosters adaptive decision-making (Goodyear et al., 2014).

Paradoxically, although pedagogical models are central to MBP-PE and aim to provide students with meaningful student-centered learning experiences (Casey and Kirk, 2020), rigid expectations from PETE programs can be counterproductive, particularly when PSTs feel underprepared for teaching (Silva et al., 2021). While methods courses preceding placement training may strengthen confidence and understanding, it is often cooperating teachers (CTs) and university supervisors who help PSTs interpret and adapt model-based strategies to real contexts (Farias et al., 2023). Providing sustainable student-centered PE may thus require PSTs to internalize and adapt core pedagogical practices to their emerging identities and contexts (Casey and Goodyear, 2015). Such professional learning benefits from mentoring balancing structure (support for models-based teaching) and flexibility (PSTs' agency), and encouraging intentional teaching without reducing models to fixed scripts (Valério et al., 2021). Inevitably, balancing model use with the need to foster agency creates tensions with the logic of strict model fidelity (Casey et al., 2021).

A second tension concerns whether PETE should prioritize PSTs' professional learning, effective student learning, or both. Although much attention has been paid to how PSTs learn to teach, studies assessing what students actually learn in PST-led PE lessons, whether through MBP-PE or traditional approaches, remain scarce. This raises important questions for PETE: what kind of learning do students achieve in PST-led lessons? Is it legitimate to prioritize teacher development while overlooking student outcomes? Does supporting PST development require strict model fidelity? Is it realistic to expect PSTs to fully implement models during early field experiences?

This study engages with these tensions by analyzing a PETE program designed to balance model fidelity and teacher agency in implementing a hybrid SE–Step Game Approach (SE-SGA) unit in volleyball. The approach combined a model-specific methods course with structured follow-up during placement. While PSTs received pedagogical scaffolds (Farias et al., 2022), such as workshops, shared resources, mentoring, and lesson-based feedback, the support was intentionally non-prescriptive. PSTs made autonomous decisions: when to seek support, how strictly to follow the model, and how to adapt to contextual and curricular demands.

Models-based practice in PE

Pedagogical models in MBP-PE are structured instructional frameworks designed to promote specific learning outcomes and meet curriculum goals. While many models associated with student-centered approaches share common pedagogical premises (Dyson et al., 2004), each model is grounded in pedagogical theory and defined by its core idea, learning aspirations, and a set of critical elements that guide teaching, learning, and assessment (Casey and Kirk, 2020). Namely, SE aims to replicate authentic sport experiences through inclusive peer-teaching interactions and participation in competition events across a full “season.” It provides an organizational framework in which students embody varied roles, such as player, coach, or referee, and work systematically within persistent teams, in a festive, cooperative climate. According to research, gains in student performance emerge from students' enhanced sense of ownership, cooperative learning dynamics, prosocial responsibility, sense of belonging, and meaningful engagement with sport culture, traditions, and values (Farias et al., 2020). These features aim to support students' engagement, participation, and social interaction, creating conditions that may facilitate learning within gameplay contexts (Hastie et al., 2011).

While SE sets the organizational structure and learning context, GBAs, particularly those based on Teaching Games for Understanding (TGfU) (Bunker and Thorpe, 1982), offer specialized development of sport content through representative, developmentally appropriate tasks tailored to students with variable ability. By organizing content around tactical principles and game categories (e.g. net or invasion games), GBAs aim to progressively develop students' tactical awareness, decision-making, and gameplay (Morales-Belando et al., 2022). These instructional features emphasize how content is structured, represented, and progressively developed through tasks, aligning with dimensions of content knowledge and pedagogical content knowledge in teaching (Ward and Ayzazo, 2016).

To expand MBP-PE's educational impact, hybrid units combining SE and GBAs have been proposed to integrate the complementary strengths of both, merging the tactical progression and content structure of GBAs with the authentic, student-centered context of SE grounded in real sport experiences. Although comparative research remains scarce, theoretical (Casey and Kirk, 2020) and empirical evidence (González-Víllora et al., 2019) suggests that hybrid approaches have been associated with stronger outcomes than single-model implementations. While they enhance student motivation, enjoyment, and responsibility, game performance gains are also greater

when SE is combined with GBAs (Araújo et al., 2016) than when SE is used in isolation (Farias et al., 2019; Gonzalez-Villora et al., 2019) or GBAs alone (Pan et al., 2023). From this perspective, hybrid units can be conceptualized as integrating organizational conditions for participation (SE) with structured opportunities for content development (GBAs), reflecting the combination of structural-related and content-related dimensions of implementation, potentially strengthening both engagement and learning in game-based contexts.

The SE-SGA combination has gained particular relevance in net games like volleyball, a compulsory team sport in upper secondary PE in countries such as Portugal, where it is seen as central to students' physical literacy (DGE, 2018). The SGA brings a structured tactical progression rooted in TGfU and the Skill Development Approach (Rink, 1993), specifically designed for volleyball. This progression employs increasingly complex modified games to foster tactical awareness, decision-making, and skill adaptability, while preserving representativeness of the mature game and high student engagement (Mesquita et al., 2005). Core pedagogies include rule and space modifications, varying player numbers (e.g. 2v2 to 4v4), and progression across stages (acquisition, structuring, adaptation), enabling responsiveness to students' diverse ability levels. Most research on SE-SGA has focused on the teacher as facilitator (Silva et al., 2021), the development of student-coaches' pedagogical knowledge (Araújo et al., 2017), and improvements in students' game performance (Araújo et al., 2016; Mesquita et al., 2005). However, all studies involved experienced PE teachers with strong volleyball backgrounds. To date, no study has examined how PSTs implement SE-SGA during school placement, nor how this affects instructional fidelity and student learning, key dimensions for assessing the feasibility and educational value of hybrid models in PETE.

Model fidelity

Pedagogical models are defined by a set of critical elements that guide teaching and learning (Casey and Kirk, 2020). Model fidelity refers to the extent to which these elements are implemented as intended in practice (Hastie and Casey, 2014). In real PE settings, however, fidelity of implementation often varies. Among PSTs, research reveals contrasting outcomes (Silva et al., 2021): some demonstrate strong commitment to model-based teaching, while others abandon it early, reverting to teacher-centered approaches when facing practical challenges. Despite its relevance, model fidelity among PSTs remains underexplored. An exception is Curtner-Smith et al. (2021), who, in a study with early-career in-service teachers, identified distinct levels of SE fidelity (full version, watered-down, and cafeteria style). These patterns were associated with contextual, individual, and mentoring-related factors shaping how the model was interpreted and enacted. In addition, existing research suggests that PSTs' implementation of pedagogical models is variable and also highly dependent on contextual support, opportunities for practice, and mentoring processes (Hastie et al., 2011; McMahon and MacPhail, 2007).

The findings above illustrate the complexity of real-world implementation and highlight ongoing tensions in how fidelity is conceptualized and assessed. While some authors present a model's critical elements as "non-negotiables," recent perspectives argue for context-sensitive adaptations, provided core pedagogical intentions are preserved (Casey and Kirk, 2020). There is also divergence regarding the role of fidelity in research: some defend strict fidelity as necessary to validly link student outcomes to the model, while others caution that relying solely on limited qualitative impressions may obscure the actual nature of implementation (Fernandez-Rio and Iglesias, 2024). Further concerns arise when fidelity assessments are based on partial sampling of lessons, raising questions about their representativeness (Iglesias and Fernandez-Rio, 2024).

Despite the multiple positions in the fidelity debate, one key aspect remains underexplored: how varying levels of fidelity, regardless of definition, relate to student learning, particularly in terms of game performance within MBP-PE. This is especially relevant in the context of PSTs' training, where variability in implementation is expected but rarely examined in relation to student outcomes (i.e. game performance improvement). Therefore, this study addressed this issue by analyzing the implementation of a hybrid SE-SGA unit in volleyball within a PETE program. Specifically, it examined: (i) the degree of implementation of SE-SGA elements by nine PSTs; (ii) students' game performance across nine classes; and (iii) whether differences in implementation levels, including variation across structural-related and content-related elements, are associated with differences in student outcomes.

Methods

Setting and participants

Setting. This study was conducted within a master's-level PETE program at a university in Northern Portugal recognized for its expertise in sport pedagogy and MBP-PE curriculum (Farias et al., 2023). The program emphasizes student-centered teaching in PE, particularly through instructional models such as SE and GBAs (e.g. SGA).

Data were collected during the school placement component in year 2 of the PETE program. The host school held a long-standing partnership with the PETE program and was open to pedagogical innovation. Located in an urban inland city, the school served around 1300 students, mostly from middle socioeconomic backgrounds, and included 130 international students of diverse origins.

PSTs and PE students. The researchers sought to recruit a cohort of PSTs placed at the same school. Nine PSTs (four female, five male; mean age $\approx 23.5 \pm 3.6$), organized into triads, volunteered to participate. Each PST taught one fixed 10th-grade class throughout the unit. Classes averaged 21 students, totaling 189 (98 boys, 91 girls; mean age $\approx 16.2 \pm 1.78$). Students had no prior experience with pedagogical models (single or hybrid).

University supervisor, non-participant observer, CTs, and teacher educator. The university supervisor held a dual role: institutional supervisor for the PSTs' placement and principal investigator leading SE-SGA implementation. Aged 47, he had over 15 years of experience conducting and publishing research on pedagogical models in PE and was the supervisor of the postdoctoral student.

A 27-year-old postdoctoral student acted as a non-participant observer. She had several publications which focused on MBP-PE and four years of experience as a teacher educator. Over two years, she was deeply immersed in the PETE program: in the first year, she observed all SE-SGA volleyball teaching methods unit sessions to document content and delivery; in the second year, she was embedded in the school placement, collecting data during all PST-taught lessons and attending all mentoring sessions. Her responsibilities included writing field notes on the presence and enactment of SE-SGA critical elements and coordinating student performance data collection.

Three experienced CTs supported mentoring, each supervising a PST triad. Their backgrounds reflected distinct specializations: CT1 (female), aged 56, was a former volleyball athlete and school coach; CT2 (female), aged 58, had a 30-year background in dance education; and CT3 (male), aged 62, was an experienced handball coach. The CTs had 29–37 years of PE teaching experience and over 20 years' experience of mentoring PSTs.

One teacher educator was indirectly involved. She taught the first-year volleyball methods unit introducing SE-SGA and had 12 years of higher education experience and extensive MBP-PE publications.

The SE-SGA program

Year 1: Volleyball SE-SGA teaching methods unit. The volleyball methods unit delivered in the first year of the PETE course focused on an SE-SGA. Supplemental Material S1 outlines the unit content and PSTs' preparation. This aligns with Casey and Kirk's (2020) view that hybridization should preserve core features of each model while fostering synergistic aspirations: supporting students in achieving skillful game participation through engaged decision-making, instruction, and social collaboration.

The teacher educator employed a pedagogical scaffolding framework developed by Farias and Mesquita (2022) to support the gradual transfer of responsibility to students across lesson planning, instruction, and assessment. This enabled the alignment of SGA's tactical progression and skill complexity with SE features such as student ownership, role-playing, and shared responsibility in decision-making.

Year 2: School placement—the volleyball SE-SGA unit taught. At the start of the school placement, the university supervisor met with the PSTs and CTs to introduce the study. While SE-SGA use was encouraged for the volleyball unit, it was not mandatory. Three key principles were clarified: (i) PSTs had full autonomy in choosing their pedagogical approach; (ii) their placement grade would not depend on full adherence to the hybrid SE-SGA features; and (iii) ongoing support was available, though PSTs were not expected to follow university supervisor suggestions if they felt unprepared or misaligned with students' pressing needs.

The university supervisor and CTs worked in differentiated mentoring roles. The university supervisor oversaw intervention coherence, equal access to support, and research integrity. This structure offered PSTs consistent mentoring and freedom to decide whether and how to implement SE-SGA in their school context. In his role, the university supervisor ensured fidelity to the study design and equal PST access to support. This included:

- **Fieldwork:** Each PST had one observed lesson, preceded by a collaborative pre-lesson discussion (PSTs, CTs, university supervisor) to examine the pedagogical rationale. Post-lesson reflections addressed: (i) alignment between planned and enacted teaching; (ii) expected versus observed student behaviors; (iii) challenges; (iv) strengths and improvements; and (v) their follow-up action plan.
- **Workshops:** Six sessions were held at the host university during unit implementation. The first two covered: (i) SE-SGA critical elements; (ii) supporting evidence; and (iii) hybrid strategies (e.g. scaffolding, modified tasks, inclusive techniques, adapted assessment like a simplified Game Performance Assessment Instrument (GPAI; Oslin et al., 1998)). The remaining four workshops were problem-based, drawing on lesson observations or PST input. Sessions included sample plans with varying SE-SGA fidelity and addressed class organization and peer-teaching scaffolding.

CTs provided situated pedagogical support, addressing PSTs' day-to-day instructional needs without enforcing or discouraging model compliance. They encouraged reflection and supported decisions regarding: (i) alignment of lesson planning with curricular goals; (ii) management of

routines and logistics (e.g. attendance, equipment); and (iii) representation, sequencing, and adaptation of instructional tasks, including feedback on task progression, anticipated tactical-technical errors, and the appropriateness of game modifications for students' needs.

As a stabilizing element of the program and in line with the curricular guidelines set by the school's PE subject group for 10th grade, the volleyball content covered in all classes was consistent across PSTs. Variation was therefore expected to arise from how content was structured, sequenced, and enacted within SE-SGA, rather than from curricular differences. Supplemental Material S2 summarizes the instructional progression of the unit, including the main game forms, game modifications, tactical and technical foci, and the gradual redistribution of instructional responsibility (PSTs/students) across the season.

Ethical procedures

Ethical approval was granted by the host university's Ethics Committee and authorized by the school board, following Portuguese national guidelines for research involving human participants. Prior to data collection, PSTs and CTs were fully informed of the study's aims and procedures. Participation was voluntary, and written informed consent was obtained in line with the Declaration of Helsinki, with anonymity and confidentiality emphasized. Additionally, a meeting with students' legal guardians explained the study and secured parental consent. Consent was also obtained from students, ensuring they understood their role and the study procedures. All student data were reported in aggregate, without identification. To minimize perceived obligation, PSTs were assured that participation, or non-participation, would not affect their academic standing or final evaluation.

Data collection

A total of 108 PE lessons (12 × 90-minute sessions taught by nine PSTs) were video-recorded using GoPro cameras positioned in the gym to capture wide-angle footage of student interactions, gameplay, and PST interventions. The recordings were analyzed to assess fidelity of SE-SGA implementation. Verbal interactions were captured via audio devices worn by PSTs, enabling identification of activities related to critical elements (e.g. role-playing such as peer-teaching tasks). Although verbal interactions were audio-recorded, they were not coded as a separate analytical layer. Instructional content was operationalized through the design and enactment of game-based elements (e.g. guided discovery, tactical complexity, and problem-solving progression), which structure students' engagement with learning tasks within gameplay, thereby enabling the identification of systematic differences in how instructional content was structured and enacted across lessons (Farias et al., 2022).

(Teaching) fidelity of critical elements implementation. A 17-item observation checklist, informed by Metzler (2011) and prior research on SE, GBAs, and hybrid volleyball units (e.g. Araújo et al., 2016; Farias et al., 2019; González-Villora et al., 2019), was used to assess the presence of key pedagogical elements from an SE-SGA unit. The checklist was developed for this study by adapting and integrating elements reported in this literature. Figure 1 presents the checklist of critical elements categorized under SE features (season, affiliation, formal competition, record keeping, festivity, and roleplay) and game-based teaching elements (game-based tasks, guided discovery, game modifications, problem-solving content development, and tactical complexity).

SE critical elements (structural-related)	
1. Season	The season is extended over a long period, including a practice period.
2.	The season is divided into pre-season, regular season, and post-season, with a culminating event.
3. Affiliation	Teams work together at all times during class.
4.	Students remain members of the same team throughout the season.
5. Formal competition	The competition schedule provides teams and players with a continuous focal point for their practice and preparation, allowing them to make short-term and long-term decisions.
6.	There is an appropriate and formal competition where team involvement is evident during these moments.
7.	The game rules may be modified to promote fairness and better participation.
8. Record keeping	The games provide many opportunities to record performances that can be used to outline the team's strategy.
9.	During class, important records are kept for the season.
10. Festivity	Seasons and competitions are as festive and celebratory as possible, without compromising the event itself.
11.	Teams have names, drawing a parallel with traditional sports.
12. Roleplay	Students perform roles and responsibilities different from the player's role.
SGA critical elements (content-related)	
13. Game-based tasks	Students engage in at least 30 minutes of practice with modified games.
14. Guided discovery	The teacher uses tactical questioning or designs tasks that allow students to discover solutions on their own.
15. Game modifications	Modifications to the full game are performed.
16. Problem-solving content development	All tasks are related to the small-sided game being taught.
17. Tactical complexity	The content of the task is developmentally appropriate (i.e. it is aligned with the students' learning stage).

Figure 1. Critical elements of the SE-SGA unit.
SE-SGA: Sport Education–Step Game Approach.

While monitored by the university supervisor, fidelity scoring was conducted by the postdoctoral student and an experienced PE teacher educator, also a published author in MBP-PE with no other involvement in the study. Prior to coding, they completed training delivered by the university supervisor using unrelated PE lesson videos to ensure shared understanding of procedures. Of the 108 lessons, 72 were video-coded; the remaining 36 were observed in person by the lead researcher

Decision-making				
Score	Service	Reception	Setting	Attack
0	Error	Error	Error	Error
1	Easily received by opponents	Poor conditions for the setter	Poor conditions for the attacker	Easily received by opponents
2	Difficult to receive	Good conditions for the setter	Good conditions for the attacker	Difficult to receive
3	Point	—	—	Point

Figure 2. Scoring matrix of the gameplay variables.

and later discussed with the second observer. Observer training followed a structured protocol involving familiarization, pilot coding, and consensus-building discussions.

To assess reliability, 20% of lessons (22 in total) were re-coded independently after a three-week interval. Intra-observer agreement ranged from 89% to 96%, and inter-observer agreement from 88% to 92%, exceeding established thresholds (Baumgartner and Jackson, 1995; Fleiss et al., 2003), confirming coding consistency.

(Student) game performance variables. Video data from 3v3 games were analyzed in Lesson 1 (pre-test) and Lesson 12 (post-test). Each student participated in four 8-minute games per time point, with consistent team composition (same opponent teams). For analysis, two 5-minute segments per student (totaling 10 minutes) were coded, enabling comparison under standardized conditions (Farias et al., 2019). In total, 5218 actions were coded over 3600 minutes.

An adapted version of the GPAI (Manso-Lorenzo et al., 2024) was used, aligned with volleyball's internal logic (Godbout and Gréhaigne, 2020) and game specificity (Mesquita, 2006). It assessed service, reception, setting, and attack using a 0–3 qualitative scale. These indicators (e.g. reception: provides good conditions for the setter) embed both decision-making and execution elements. Figure 2 shows how tactical adequacy and action outcomes were rated.

The postdoctoral student coded the students' gameplay under the supervision of the university supervisor. A second coder, a PE teacher and volleyball coach not involved in the study, was trained over two weeks. Inter- and intra-observer reliability were tested on 20% of the dataset. Intra-observer agreement ranged from 87% to 98%, and inter-observer agreement from 89% to 91%, surpassing accepted reliability standards (Fleiss et al., 2003), confirming consistent and valid coding of students' game performance.

Data analysis

(Teaching) fidelity of critical elements implementation. All 108 PE lessons were coded using a 17-item checklist. Fidelity scores were calculated based on the proportion of lessons in which each element was present, as agreed by two independent observers. A critical element was considered achieved when both observers consistently marked it as present. Implementation fidelity was classified as high, moderate, or low (e.g. ≥ 13 elements = high) (Stylianou et al., 2016). To identify natural groupings among PSTs, a hierarchical cluster analysis (Ward's method, squared Euclidean distance) was performed using SPSS 30.0. The analysis yielded a three-cluster solution, reflecting high, moderate, and low implementation levels, with three PSTs in each group. These groupings aligned with the mentoring structure, as each cluster corresponded to a different CT. This alignment

was interpreted as reflecting the potential contribution of distinct mentoring contexts, including CTs' pedagogical orientations and content-related guidance, on PSTs' implementation decisions.

Each PST received a fidelity score (1–17) for each lesson. These scores were converted to percentages and averaged over the 12 lessons. For analytical purposes, implementation fidelity was further disaggregated into two complementary dimensions. Structural-related fidelity reflected the implementation of SE elements (items 1–12), capturing the organizational and pedagogical architecture of the model (e.g. season structure, affiliation, formal competition, record keeping, festivity, and roles). Content-related fidelity reflected the enactment of SGA elements (items 13–17), capturing instructional processes directly associated with game understanding and performance development (i.e. game-based tasks, guided discovery, game modifications, problem-solving progression, and tactical complexity). For each PST, separate fidelity scores were calculated for structural-related and content-related dimensions by averaging the proportion of lessons in which the corresponding elements were implemented across the 12-lesson unit. These scores were then aggregated within each cluster to compute mean (M) and standard deviation (SD) values.

Assumptions of normality (Shapiro–Wilk, $ps > .19$) and homogeneity of variances (Levene's test, $p = .644$) were met. Descriptive statistics (M, SD) were computed. A one-way analysis of variance (ANOVA) examined differences in fidelity levels across groups, followed by Tukey HSD tests. To further examine differences in implementation profiles, one-way ANOVAs were conducted separately for structural-related and content-related fidelity across clusters. Effect sizes were reported using η^2 . This disaggregation enabled examination of whether differences between clusters were associated with organizational features or instructional content enactment. Effect sizes were reported using η^2 , ϵ^2 , and ω^2 .

(Student) game performance variables. Gameplay scores were cumulatively calculated for each gameplay action (service, reception, setting, attack) using a 0–3 scale depicting growing sophistication in decision-making (see Figure 2). A game performance index was calculated based on: (service score + reception score + setting score + attack score)/4.

Three complementary metrics were derived: (1) game involvement score—total number of actions per player; (2) rate of play—actions per minute; and (3) efficacy score—number of actions rated as successful (point or set continuity).

Data were analyzed using SPSS 27.0. Normality was assessed via Kolmogorov–Smirnov tests ($ps > .050$). Descriptive statistics (M, SD) were calculated.

To analyze differences across fidelity levels, a one-way ANOVA was applied with follow-up post hoc comparisons using the Tukey HSD test. To verify the progression of each model fidelity group regarding student game performance (i.e. three-level factor) during the intervention, multivariate analysis of variance, ANOVA, multivariate analysis of covariance (MANCOVA), and analysis of covariance (ANCOVA) were conducted. Eta-squared (η^2) values were reported to indicate effect sizes. Statistical significance was set at $p \leq .050$ (95% confidence interval). Lastly, a parallel multiple mediation analysis was conducted to determine direct and indirect effects on overall game performance.

Results

Fidelity of critical elements implementation

Based on the analysis of descriptive measures, the nine PSTs were distributed across the three fidelity levels. Notably, the three triads, each supervised by a different CT, aligned with these three levels of fidelity.

As presented in Table 1, PSTs in the high implementation group demonstrated an average percentage of critical elements usage of 82.4% (± 3.1), those in the moderate implementation group averaged 58.2% (± 3.8), and those in the low implementation group averaged 27.8% (± 2.4). Shapiro–Wilk tests indicated that the data were normally distributed across all groups ($ps > .19$), and Levene’s test confirmed the homogeneity of variances, $p = .644$, based on the mean. There was a significant effect of implementation group on the percentage of critical elements used across lessons, $F(2, 6) = 225.60$, $p < .001$, $\eta^2 = 0.987$. Post hoc comparisons indicated that PSTs in the high implementation group ($M = 82.37\%$, $SD = 3.07\%$) used significantly more critical elements than those in the moderate ($M = 58.17\%$, $SD = 3.83\%$, $p < .001$) and low implementation groups ($M = 27.77\%$, $SD = 2.41\%$, $p < .001$). PSTs in the moderate implementation group used significantly more critical elements than the low implementation group ($M = 58.17\%$, $SD = 3.83\%$ vs. $M = 27.77\%$, $SD = 2.41\%$, $p < .001$).

Effect sizes were large, with $\eta^2 = 0.987$, $e^2 = 0.983$, and ω^2 (fixed effect) = 0.980, all indicating substantial effects of the implementation level on the use of critical elements.

To further examine the nature of fidelity differences, implementation was disaggregated into structural-related (SE; items 1–12) and content-related (SGA; items 13–17) dimensions (Table 2). A one-way ANOVA revealed significant differences across clusters for both dimensions. Structural-related fidelity was significantly higher in the higher implementation group ($M = 82.4$, $SD = 3.1$) compared to the moderate ($M = 58.2$, $SD = 3.8$) and lower groups ($M = 27.8$, $SD = 2.4$), $F(2, 6) = 225.60$, $p < .001$, $\eta^2 = 0.987$. Similarly, content-related fidelity differed significantly across clusters, with the higher group ($M = 84.0$, $SD = 4.5$) showing significantly greater enactment than the moderate ($M = 52.0$, $SD = 6.5$), and lower groups ($M = 21.0$, $SD = 5.0$), $F(2, 6) = 110$, $p < .001$, $\eta^2 \approx 0.96$.

A more detailed breakdown of implementation by element (Table 3) showed that differences between clusters were evident across all structural-related and content-related benchmarks. Higher implementation was characterized by consistently high levels across both dimensions, whereas moderate and lower implementation groups showed progressively reduced enactment, particularly for content-related elements such as guided discovery, problem-solving content development, and tactical complexity.

Interaction between the critical elements implementation and students’ game performance

The means and standard deviations of the gameplay scores of the four gameplay components assessed (service, reception, distribution, and attack) are shown in Table 4.

After inspecting the assumptions for the MANCOVA, the four components of game performance showed statistically significant differences among the three implementation groups at the end of the implementation [Wilks’ Lambda = 0.866; $F(4, 96) = 3.704$; $p = .008$]. The effect size was considered large ($\eta^2 = 0.134$). Regarding ANCOVA results, all groups showed significant improvements ($p < .001$) from pre- to post-test. Regarding the post-test inter-group scores, significant differences were observed among implementation groups. Specifically, the high-implementation group outperformed the moderate- and low-implementation groups in service [$F(2) = 415.53$, $p < 0.001$], reception [$F(2) = 25.992$, $p < 0.001$], and attack [$F(2) = 18.679$, $p < 0.001$]. However, no significant differences were observed in distribution [$F(2) = 2.174$; $p = .117$] among the three groups.

Table 1. Percentage in each lesson of the use of the 17 critical elements of the SE-SGA.

Implementation level	PSTs	Lesson-by-lesson percentage of implementation of critical components of SE-SGA												Mean (SD)			
		1	2	3	4	5	6	7	8	9	10	11	12				
Higher	A	76.5	82.4	88.2	82.4	82.4	82.4	82.4	88.2	76.5	70.6	70.6	82.4	88.2	70.6	70.6	81.4 (7.0)
	B	82.4	88.2	88.2	88.2	94.1	100.0	76.5	82.4	70.6	88.2	88.2	82.4	88.2	88.2	82.4	85.8 (7.7)
	C	76.5	70.6	70.6	76.5	82.4	82.4	82.4	70.6	76.5	82.4	82.4	82.4	94.1	94.1	94.1	79.9 (8.1)
Moderate	D	70.6	58.8	64.7	64.7	58.8	52.9	58.8	47.1	64.7	64.7	64.7	64.7	47.1	58.8	47.1	59.3 (7.3)
	E	70.6	64.7	70.6	64.7	64.7	47.1	52.9	52.9	64.7	64.7	64.7	64.7	52.9	64.7	52.9	61.3 (7.7)
	F	64.7	47.1	47.1	52.9	58.8	58.8	64.7	47.1	41.2	47.1	47.1	47.1	52.9	64.7	52.9	53.9 (8.3)
Lower	G	41.2	29.4	35.3	29.4	23.5	23.5	23.5	29.4	29.4	29.4	29.4	23.5	29.4	23.5	23.5	29.4 (5.6)
	H	41.2	29.4	23.5	29.4	23.5	29.4	29.4	35.3	23.5	17.6	17.6	35.3	29.4	35.3	29.4	28.9 (6.4)
	I	35.3	23.5	17.6	23.5	23.5	29.4	11.8	17.6	23.5	29.4	29.4	29.4	29.4	29.4	29.4	25.0 (7.1)

Notes. PSTs: preservice teachers; SD: standard deviation; SE-SGA: Sport Education-Step Game Approach.

Table 2. Mean (SD) percentage of implementation of structural-related and content-related fidelity across implementation clusters (higher, moderate, lower), with ANOVA results.

Fidelity dimension	Higher	Moderate	Lower	<i>F</i> (2,6)	<i>p</i>	η^2
Structural-related fidelity	82.4 (3.1)	58.2 (3.8)	27.8 (2.4)	225.60	< .001	0.987
Content-related fidelity	84.0 (4.5)	52.0 (6.5)	21.0 (5.0)	95–130	< .001	≈ 0.96

Notes. SD: standard deviation; ANOVA; analysis of variance. Scores represent mean and standard deviation.

Table 3. Descriptive breakdown of the percentage of implementation of structural-related (SE) and content-related (SGA) critical elements across implementation clusters (higher, moderate, lower).

Element	Higher (%)	Moderate (%)	Lower (%)
Structural-related benchmarks			
Season (items 1–2)	90	65	30
Affiliation (items 3–4)	88	60	28
Formal competition (items 5–7)	85	58	25
Record keeping (items 8–9)	82	55	22
Festivity (items 10–11)	80	52	20
Roleplay (item 12)	88	60	25
Mean (SD)	82.4 (3.1)	58.2 (3.8)	28.8 (2.4)
Content-related benchmarks			
Game-based tasks (item 13)	92	65	35
Guided discovery (item 14)	85	50	20
Game modifications (item 15)	88	55	25
Problem-solving content development (item 16)	90	48	18
Tactical complexity (item 17)	87	52	20
Mean (SD)	84.1 (4.6)	51.8 (6.3)	20.9 (5)

Note. SD: standard deviation; SE: Sport Education; SGA: Step Game Approach.

Overall game performance indices. Table 5 shows the means and standard deviations of the gameplay composite scores (game involvement, efficacy, rate of play, and game performance index).

Multivariate analysis revealed significant differences among implementation groups [Wilks’ Lambda = 0.478; $F(8,334) = 18.63$; $p < .001$]. The effect size was considered very large ($\eta^2 = 0.316$). All outcomes showed significant improvements ($p < .001$) from pre- to post-test. Effect sizes demonstrated medium effects for game performance index ($\eta^2 = 0.575$), large effects for game involvement ($\eta^2 = 0.902$) and rate of play ($\eta^2 = 0.681$), and very large effects for efficacy ($\eta^2 = 0.376$).

Regarding post-test inter-group ANOVA, results indicated significant differences among implementation groups. Specifically, the moderate application group achieved significantly higher game involvement [$F(2) = 29.241$; $p < .001$]. The higher implementation group scored significantly better results in efficacy [$F(2) = 521.707$; $p < .001$] and rate of play [$F(2) = 13.636$; $p < .001$]. Although the higher implementation group also showed higher mean scores in game performance index [$F(2) = 22.701$; $p = .117$], this difference was not statistically significant.

Parallel multiple mediator model. Figure 3 presents the mediation analysis using overall game performance indices. The total effect indicated a significant relationship between efficacy and game

Table 4. Descriptive statistics of the evolution of the decision-making indices.

	Service		Reception		Distribution		Attack	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Higher implementation	1.70 (1.56)	2.49* (1.34)	0.92 (.60)	1.39* (.43)	0.74 (.62)	1.29* (.49)	0.97 (.70)	1.29* (.60)
Moderate implementation	1.13 (.42)	1.10 (.36)	0.93 (.53)	1.11* (.42)	0.94 (.76)	1.19* (.62)	1.28 (.69)	1.38* (.46)
Lower implementation	0.87 (.56)	1.05* (.54)	0.75* (.59)	0.79 (.39)	0.53 (.56)	0.60* (.64)	0.97 (.77)	1.14* (.61)

Note. Scores represent mean (standard deviation).

*Significance (p-value) throughout the time.

Table 5. Descriptive statistics of the evolution of the overall game performance.

	Game involvement		Efficacy score		Rate of play		Game performance index	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Higher implementation	13.64 (5.31)	15.38* (7.02)	3.71 (1.55)	5.49* (1.34)	1.70 (.66)	1.92* (.87)	1.16 (.55)	1.41* (.48)
Moderate implementation	25.45* (9.64)	17.14 (5.51)	1.09 (.30)	1.19* (.26)	2.54* (.96)	1.71 (.55)	1.07 (.36)	1.19* (.27)
Lower implementation	12.82 (5.41)	17.31* (6.53)	0.89 (.37)	1.01* (.27)	0.53 (.56)	1.28* (.54)	1.73* (.65)	0.90 (.34)

Note. Mean (standard deviation).

*Significance (p-value) throughout the time.

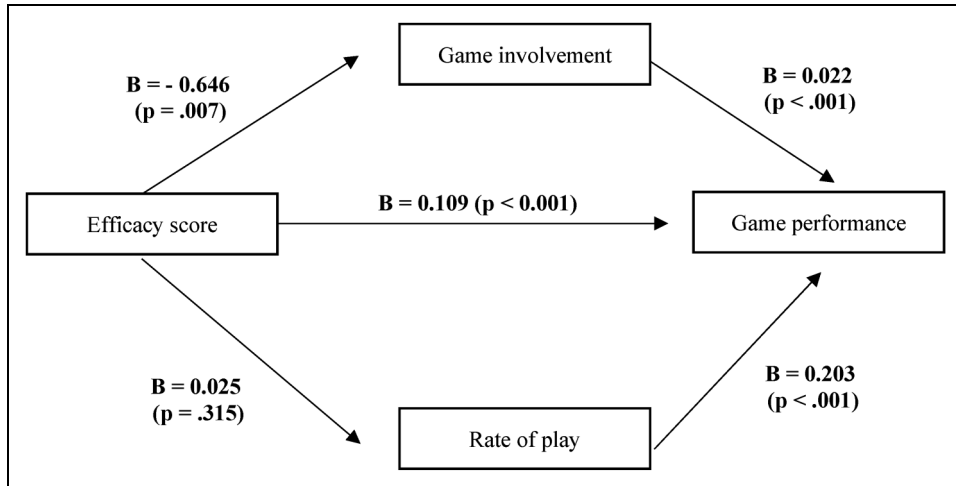


Figure 3. Multiple mediation role for the overall game performance.

performance index ($B = 0.109$; $p < .001$). All indices contributed to explaining overall game performance. However, the relationship between efficacy and rate of play was not statistically significant ($p > .050$).

Discussion

The PSTs demonstrated markedly different levels of fidelity when implementing the hybrid SE-SGA unit. Those in the high-fidelity group applied, on average, 82.4% of the 17 critical elements per lesson, significantly more than the moderate (58.2%) and low-fidelity (27.8%) groups. These differences were statistically significant, with large effect sizes, showing substantial variation even under comparable pedagogical conditions (Baumgartner and Jackson, 1995).

Importantly, differences in implementation were not only quantitative but also qualitative, as fidelity varied across structural- and content-related dimensions. Content-related elements (e.g. guided discovery, tactical progression, game modifications, and problem-solving structures) varied across clusters, as reflected in element-level implementation percentages and mean fidelity scores. Higher implementation showed more consistent enactment, whereas moderate and lower levels showed reduced implementation, particularly in guided discovery and tactical complexity. In part, this contrasts with prior research indicating PSTs often struggle to sustain pedagogical models during early teaching (Fernandez-Rio and Iglesias, 2024; Loflin, 2015; Silva et al., 2021). While not implying full model mastery, the consistent enactment observed suggests that meaningful model implementation is attainable within a supportive PETE instructional ecology (Araújo et al., 2017; Evangelio et al., 2018; González-Villora et al., 2020).

Indeed, the PETE program's structure appears to have supported PSTs' capacity to implement an SE-SGA unit. In year 1, PSTs completed a volleyball methods course organized around the hybrid model, introducing its pedagogical rationale through structured practice and reflection. Experiencing the model as both student and teacher (living the curriculum) fosters pedagogical intentionality (Casey and Kirk, 2020; Curtner-Smith et al., 2008; Deenihan et al., 2011). When combined with scaffolded planning, peer observation, and dialogic feedback, as in this study, such experiential learning supports model fidelity (Casey and MacPhail, 2018; Ward et al., 2018).

During school placement, a dual mentoring structure provided continued support. The university supervisor ensured program coherence through observation cycles, while CTs offered situated mentoring, assisting with routines, curriculum alignment, and instructional challenges (Farias et al., 2023). Implementation of the SE-SGA features was encouraged but not mandatory, preserving PSTs' autonomy, aligned with actor-oriented views on curriculum enactment (Penuel et al., 2014). The alignment between fidelity clusters and CT supervision strongly suggests that mentoring was structurally influential. While the university supervisor ensured overall coherence and equal access to pedagogical support, CTs provided situated guidance more closely connected to instructional enactment. CTs' pedagogical orientations and content expertise may shape PSTs' interpretations of pedagogical priorities, influencing which elements were treated as accountable in practice. Accordingly, PSTs' decisions regarding content-related elements such as task design, guided discovery, and tactical progression may have been shaped by CTs' subject-matter knowledge, particularly in how volleyball content was represented, sequenced, and adapted for learning (González-Villora et al., 2019).

Still, the variability observed highlights the non-linear nature of implementation. Though PSTs shared the same program design and formal supports, fidelity differences may reflect individual factors (motivation, prior socialization, instructional confidence) (Curtner-Smith et al., 2008; Ward and Ayvazo, 2016). At the meso-level, each fidelity level corresponded to a different CT triad. Despite their mentoring experience, CTs' disciplinary backgrounds (volleyball, dance, handball) appear to have shaped PSTs' engagement with SE-SGA (Silva et al., 2021). For instance, the volleyball-specialist CT supervised the high-fidelity group, possibly reinforcing tactical clarity and confidence. These patterns require further study through in situ analyses of PST-mentor interactions.

Students taught by PSTs in the higher-fidelity group showed the most consistent improvements. Statistically significant gains were observed across all four discrete gameplay actions (service, reception, setting, attack) and in all composite indicators: game involvement, rate of play, efficacy, and overall game performance. These results reinforce the pedagogical argument that high-fidelity SE-SGA implementation, fully enacting its critical elements, is associated with skillful gameplay and student engagement (Fernandez-Rio and Iglesias, 2024; Hastie et al., 2017), aligning with the more consistent enactment of content-related elements. In the moderate-fidelity group, students improved in three gameplay actions (excluding service) and two performance indicators (efficacy and game performance index), but declined in participation measures, game involvement and rate of play. This suggests a partial learning trajectory with technical gains (e.g. setting, attack) but insufficient support for sustained game engagement (Farias et al., 2019). Students in the lower-fidelity group also improved in three gameplay actions (excluding reception). Notably, assessed actions reflect both decision-making and execution. For instance, setting up an effective attack requires cue recognition and anticipation of ball trajectory. However, no improvement in the game performance index suggests simpler play patterns. For example, "providing good conditions for the setter/attacker" (2 points) reflects greater tactical soundness than "poor conditions" (1 point) (Zhang et al., 2024). Still, gains in game involvement and rate of play indicate that these PSTs fostered active student participation (Farias et al., 2019).

Differences in student outcomes may be associated not only with the overall level of fidelity, but with distinct profiles of implementation. Reduced gains in some groups may reflect less consistent enactment of content-related elements (e.g. guided discovery, tactical progression, problem-solving structures), suggesting that higher-quality implementation may depend on alignment between structural organization and instructional content. Descriptively, the higher implementation group showed more balanced levels of structural- and content-related fidelity, whereas the moderate

and lower groups showed reduced content-related implementation relative to structural elements. In practical terms, the high-fidelity cluster consistently implemented guided discovery, task modification, and tactical complexity, whereas the moderate and low clusters used these elements discontinuously or in a simplified form. Although this pattern was not formally tested, it suggests that higher-quality implementation may involve a more balanced enactment of organizational structure and instructional content.

Furthermore, the present study did not quantify or systematically code teacher–student instructional interactions (e.g. feedback, questioning, cueing), not considering these elements as a separate analytical layer. However, instructional content was operationalized at the task-design, sequencing, and adaptation levels through elements such as game-based tasks, guided discovery, game modifications, and tactical progression. In line with research on pedagogical content knowledge, effective teaching is reflected not only in moment-to-moment interactions but also in how tasks are selected, structured, and progressively developed to support learning (Ward and Ayvazo, 2016), which in this study enabled the identification of systematic differences in instructional content across implementation profiles without relying on micro-level interaction coding. These elements align with specialized content knowledge, including the ability to design tasks, anticipate learner difficulties, and adapt instruction.

These findings offer two key implications. First, greater adherence to the hybrid model’s critical elements was correlated with stronger outcomes. Second, learning was not exclusive to high-fidelity contexts: moderate- and low-fidelity implementations also yielded improvements, though more limited in scope. This challenges fidelity “orthodoxy,” suggesting that effectiveness may also be associated with responsive adaptation to contextual demands (Casey and Kirk, 2020; Silva et al., 2021). Indeed, improvements under partial fidelity may reflect pedagogical decisions prioritizing engagement, task simplification, or alignment with curricular/time constraints or student readiness (McCaughy et al., 2006; Penuel et al., 2014). When grounded in student-centered principles, such adaptations can still uphold core pedagogical goals (Casey, 2024).

Implications for research and practice

This study supports a flexible view of model fidelity, indicating that pedagogical effectiveness, as evidenced by student learning, may occur along a continuum of implementation, even without full replication of all model components (Casey et al., 2021). The three-tier fidelity analysis showed that even partial yet coherent enactment of a hybrid SE-SGA unit may support meaningful learning, reinforcing the view that fidelity should not be reduced to a count of elements, but also consider pedagogical coherence and intentional alignment embedded in practice (Casey, 2024; Penuel et al., 2014). Two PSTs may score equally on fidelity yet differ pedagogically. For example, team affiliation may be superficially addressed (e.g. assigning names and colors) or deeply embedded through student-designed logos, chants, or rituals (Hastie et al., 2011). Similarly, tactical questioning may span from factual recall elements (lower-order: how do you perform the forearm pass?) to strategic problem-solving prompts (higher-order: when do you move up to the net or stay down court?) (Farias et al., 2019). These enactments reflect different levels of alignment with the model’s core aims and illustrate diverse ways to activate the pedagogical potential of critical elements. These examples underscore the need to assess both the presence and the depth of implementation. Variation in implementation may also reflect individual and contextual factors, including PST and CT characteristics, which were not explicitly examined in the present study. Prior research in PETE and models-based practice suggests that teachers’ prior socialization, content expertise, pedagogical beliefs, and mentoring contexts play a critical role in shaping how pedagogical models are

interpreted and enacted in practice (Casey and MacPhail, 2018; Curtner-Smith et al., 2008; Ward and Ayvazo, 2016). In this sense, differences in fidelity and instructional emphasis may not solely reflect adherence to model elements, but also how these elements are mediated by teachers' backgrounds and the supervisory environment in which they are embedded. Here, systematizing phased enactments aligned with pedagogical complexity, across (e.g. festivity vs. game modifications) and within elements (e.g. evolving record keeping), may prove useful.

From this perspective, PETE programs might frame pedagogical models not as rigid scripts, but as principled frameworks oriented toward student-centered learning, where teachers act as facilitators, students are actively engaged, and learning tasks are collaboratively structured, adapted, and progressively challenging (Dyson et al., 2004). Within such an approach, fidelity may be better understood as alignment with pedagogical goals rather than strict adherence to predefined model elements, allowing PSTs to progressively enact and refine model elements as their pedagogical knowledge and confidence develop. This requires investment in conceptual clarity and pedagogical reasoning, equipping PSTs to adapt instructional decisions purposefully while maintaining the educational intent of the model (Deenihan et al., 2011).

Limitations

A key limitation of the study is that teacher–student instructional interactions were not analyzed as a separate analytical layer, as fine-grained coding of feedback, questioning, and cueing was not undertaken. Future research should incorporate systematic coding of teacher–student instructional interactions (e.g. feedback, questioning, and cueing) to examine how content-related practices are enacted across lessons and how they relate to students' decision-making, engagement, and game performance. Equally important is examining PST–mentor interactions, particularly how instructional decisions are negotiated in practice, including the design, sequencing, and adaptation of tasks, as well as the nature of instructional guidance provided to students. Such analyses may clarify how pedagogical responsibility is distributed and how content-related decisions are shaped within the mentoring process (Farias et al., 2023).

Given the quasi-experimental design and absence of a control group, causal interpretations should be made with caution. Nevertheless, the consistent patterns observed across implementation profiles provide a basis for examining associations between model enactment and student outcomes. Future studies could further isolate the effects of individual elements (e.g. guided discovery, task modifications, and tactical progression) (Morales-Belando et al., 2022; Zhang et al., 2024) or adopt longitudinal approaches to track how PSTs' decisions evolve.

Conclusions


The three-tier model fidelity analysis in this study suggests that student learning may emerge along a continuum of model implementation, not being exclusively contingent on the full replication of all model components. Fidelity research is thus called to move beyond verifying *what* is implemented, to also examine *how* and *to what pedagogical depth* practices are enacted. The findings also support the view that fidelity to student-centered models is attainable early in PETE. Concurrently, they highlight the importance of moving beyond surface-level metrics to consider how structural conditions, individual dispositions, and contextual dynamics may influence instructional decisions. A more comprehensive approach to fidelity would consider both the presence of critical elements and the pedagogical intentions and conditions that guide their implementation. These insights encourage PETE programs and future research to reframe fidelity not as rigid adherence, but as

principled, context-responsive enactment, seeking alignment with broader student-centered goals through flexible pedagogical practice.


Acknowledgments

We thank the PSTs, their CTs, the children, and the school board for their participation and support.


ORCID iDs

Cláudio Farias  <https://orcid.org/0000-0002-1364-0539>

Ellen-Alyssa F Gambles  <https://orcid.org/0000-0002-5931-136X>

Jacob Sierra-Díaz  <https://orcid.org/0000-0002-1098-4870>

Sixto González-Víllora  <https://orcid.org/0000-0003-2473-5223>

Isabel Mesquita  <https://orcid.org/0000-0001-5985-9686>

Author contributions

Cláudio Farias: conceptualization, investigation methodology, writing—original draft preparation, and writing—review and editing; Ellen-Alyssa Gambles: conceptualization, writing—original draft preparation, and writing—review and editing; Jacob Sierra-Díaz: methodology, and writing—original draft preparation; Sixto González-Víllora: methodology, writing—original draft preparation, and writing—review and editing; and Isabel Mesquita: conceptualization, investigation, and writing—review and editing.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

Data cannot be shared and made public under the Data Protection Legislation governing the host institution.

Supplemental material

Supplemental material for this article is available online.

References

- Araújo R, Mesquita I, Hastie P, et al. (2016) Students' game performance improvements during a hybrid sport education—step-game-approach volleyball unit. *European Physical Education Review* 22(2): 185–200.
- Araújo RMF, Hastie PA, de Assunção Bessa Pereira CH, et al. (2017) The evolution of student-coach's pedagogical content knowledge in a combined use of sport education and the step-game-approach model. *Physical Education and Sport Pedagogy* 22(5): 518–535.
- Baumgartner TA and Jackson AS (1995) *Measurement for Evaluation in Physical Education and Exercise Science*. 5th ed. Dubuque: Wm. C. Brown.
- Bunker D and Thorpe R (1982) A model for the teaching of games in secondary schools. *Bulletin of Physical Education* 18: 5–8.
- Casey A (2024) The ghosts of research past, present and future: Understanding the past to inform a future of models-based practice research. *Physical Education and Sport Pedagogy*: 1–14.

- Casey A and Goodyear V (2015) Can cooperative learning achieve the four learning outcomes of physical education? A review of literature. *Quest (Grand Rapids, Mich)* 67(1): 56–72.
- Casey A and Kirk D (2020) *Models-based Practice in Physical Education*. London: Routledge.
- Casey A and MacPhail A (2018) Adopting a models-based approach to teaching physical education. *Physical Education and Sport Pedagogy* 23(3): 294–310.
- Casey A, MacPhail A, Larsson H, et al. (2021) Between hope and happening: Problematizing the M and the P in models-based practice. *Physical Education and Sport Pedagogy* 26(2): 111–122.
- Curtner-Smith MD, Hastie PA and Kinchin GD (2008) Influence of occupational socialization on beginning teachers' interpretation and delivery of sport education. *Sport, Education and Society* 13(1): 97–117.
- Curtner-Smith MD, Kinchin GD, Hastie PA, et al. (2021) 'It's a lot less hassle and a lot more fun': Factors that sustain teachers' enthusiasm for and ability to deliver sport education. *Journal of Teaching in Physical Education* 40(2): 312–321.
- Deenihan JT, MacPhail A and Young AM (2011) 'Living the curriculum': Integrating sport education into a physical education teacher education programme. *European Physical Education Review* 17(1): 51–68.
- Direção-Geral Da Educação [DGE] (2018) Ano letivo de 2017–2018. Available at: <https://dge.mec.pt/ano-letivo-de-2017-2018> (accessed 5 July 2025)
- Dyson B, Griffin LL and Hastie PA (2004) Sport education, tactical games, and cooperative learning: Theoretical and pedagogical considerations. *Quest* 56(2): 226–240.
- Evangelio C, Sierra-Díaz J, González-Villora S, et al. (2018) The sport education model in elementary and secondary education: A systematic review. *Movimento* 24(3): 931–946.
- Farias C, Fernández-Río J, Martins J, et al. (2023) Multi-system influences on physical education preservice teachers' teaching practice in pandemic times. *Quest (Grand Rapids, Mich)* 75(4): 325–343.
- Farias C and Mesquita I (eds) (2022) *Learner-Oriented Teaching and Assessment in Youth Sport*. London: Routledge.
- Farias C, Mesquita I and Hastie PA (2019) Student game-play performance in invasion games following three consecutive hybrid sport education seasons. *European Physical Education Review* 25(3): 691–712.
- Farias C, Pill S and Griffin L (2025) *Game-Based Approaches in Physical Education: International Applications*. New York: Routledge.
- Farias C, Teixeira J, Ribeiro E, et al. (2022) Effects of a two-stage physical education teacher education programme on preservice teachers' specialised content knowledge and students' gameplay in a student-centered sport education–step game approach. *European Physical Education Review* 28(3): 816–834.
- Farias C, Wallhead T and Mesquita I (2020) "The project changed my life": Sport education's transformative potential on student physical literacy. *Research Quarterly for Exercise and Sport* 91(2): 263–278.
- Fernandez-Río J and Iglesias D (2024) What do we know about pedagogical models in physical education so far? An umbrella review. *Physical Education and Sport Pedagogy* 29(2): 190–205.
- Fleiss JL, Levin B and Paik MC (2003) *Statistical Methods for Rates and Proportions*. 3rd ed. Hoboken: Wiley.
- Godbout P and Gréhaigine JF (2020) Regulation of tactical learning in team sports – The case of the tactical-decision learning model. *Physical Education and Sport Pedagogy* 27(3): 215–230.
- González-Villora S, Evangelio C, Sierra-Díaz J, et al. (2019) Hybridizing pedagogical models: A systematic review. *European Physical Education Review* 25(4): 1056–1074.
- González-Villora S, Fernandez-Río FJ, Guijarro E, et al. (2020) *The Game-Centred Approach to Sport Literacy*. Abingdon: Routledge.
- Goodyear VA, Casey A and Kirk D (2014) Hiding behind the camera: Social learning within the cooperative learning model to engage girls in physical education. *Sport, Education and Society* 19(6): 712–734.
- Hastie PA and Casey A (2014) Fidelity in models-based practice research in sport pedagogy: A guide for future investigations. *Journal of Teaching in Physical Education* 33(3): 422–431.

- Hastie PA, de Ojeda DM and Luquin AC (2011) A review of research on sport education: 2004 to the present. *Physical Education and Sport Pedagogy* 16(2): 103–132.
- Hastie PA, Ward JK and Brock SJ (2017) Effect of graded competition on student opportunities for participation and success rates during a season of sport education. *Physical Education and Sport Pedagogy* 22(3): 316–327.
- Iglesias D and Fernandez-Rio J (2024) A model fidelity check in cooperative learning research in physical education. *Physical Education and Sport Pedagogy* 31(3): 364–379.
- Loflin JW (2015) Relationship between teacher fidelity and physical education student outcomes. *The Physical Educator* 72: 359–383.
- Manso-Lorenzo V, Evangelio C, Ruiz-Tendero G, et al. (2024) Game performance assessment instrument in physical education: A systematic review from 2015 to 2024. *Physical Education and Sport Pedagogy*: 1–16.
- McCaughy N, Barnard S, Martin J, et al. (2006) Teachers' perspectives on the challenges of teaching physical education in urban schools: The student emotional filter. *Research Quarterly for Exercise and Sport* 77(4): 486–497.
- McMahon E and MacPhail A (2007) Learning to teach sport education: The experiences of pre-service teachers. *European Physical Education Review* 13: 229–246.
- Mesquita I (2006) Ensinar bem para aprender melhor o jogo de voleibol. In: Tani G, Bento J and Petersen R (eds) *Pedagogia do Desporto*. Rio de Janeiro: Guanabara Koogan, pp. 327–343.
- Mesquita I, Graça A, Gomes AR, et al. (2005) Examining the impact of a step game approach to teaching volleyball on student tactical decision making and skill execution during game play. *Journal of Human Movement Studies* 48(6): 469–492.
- Metzler M (2011) *Instructional Models for Physical Education*. 3rd ed. Scottsdale: Holcomb Hathaway.
- Morales-Belando MT, Kirk D and Arias-Estero JL (2022) A systematic review of teaching games for understanding intervention studies from a practice-referenced perspective. *Research Quarterly for Exercise and Sport* 93(4): 670–681.
- Oslin JL, Mitchell SA and Griffin LL (1998) The game performance assessment instrument (GPAI): Development and preliminary validation. *Journal of Teaching in Physical Education* 17(2): 231–243.
- Pan YH, Huang CH and Hsu WT (2023) A comparison of the learning effects between TGfU-SE and TGfU on learning motivation, sport enjoyment, responsibility, and game performance in physical education. *Frontiers in Psychology* 14: 1165064.
- Penuel WR, Phillips RS and Harris CJ (2014) Analysing teachers' curriculum implementation from integrity and actor-oriented perspectives. *Journal of Curriculum Studies* 46(6): 751–777.
- Rink J (1993) *Teaching Physical Education for Learning*. 2nd ed. Missouri: Mosby.
- Siedentop D, Hastie P and Van der Mars H (2019) *Complete Guide to Sport Education*. Champaign: Human Kinetics.
- Silva R, Farias C and Mesquita I (2021) Challenges faced by preservice and novice teachers in implementing student-centred models: A systematic review. *European Physical Education Review* 27(4): 798–816.
- Stylianou M, Kloeppel T, Kulinna P, et al. (2016) Teacher fidelity to a physical education curricular model and physical activity outcomes. *Journal of Teaching in Physical Education* 35(4): 337–348.
- Valério C, Farias C and Mesquita I (2021) Pre-service teachers' learning and implementation of student-centred models in physical education: A systematic review. *Journal of Physical Education and Sport* 21(6): 3326–3338.
- Ward P and Ayvazo S (2016) Pedagogical content knowledge: Conceptions and findings in physical education. *Journal of Teaching in Physical Education* 35(3): 194–207.
- Ward P, Chen YJ, Higginson K, et al. (2018) Teaching rehearsals and repeated teaching: Practice-based physical education teacher education pedagogies. *Journal of Physical Education, Recreation & Dance* 89(6): 20–25.

Zhang J, Geok Soh K, Bai X, et al. (2024) Optimizing learning outcomes in physical education: A comprehensive systematic review of hybrid pedagogical models integrated with the sport education model. *PLoS ONE* 19(12): e0311957.

Author biographies

Cláudio Farias is an Assistant Professor at the Centre for Research, Education, Intervention and Innovation in Sport (CIFID2D), Faculty of Sport, University of Porto, Portugal.

Ellen-Alyssa F Gambles is a Lecturer in Secondary Physical Education Initial Teacher Training at the University of Sunderland, UK and Treasurer and Communications Co-ordinator for the AIESEP Teaching Games for Understanding Special Interest Group.

Jacob Sierra-Diaz is an Assistant Professor in Ciencias de la Educación, University of Oviedo, Spain and a member of the EDAFIDES Educación, Actividad Física, Deporte y Salud.

Sixto González-Villora is a Full Professor at the Faculty of Education of Albacete at UCLM, Spain and UCLM's representative on the Network of Universities for Research on Children and Adolescents (RUIA).

Isabel Mesquita is a Full Professor at the Centre for Research, Education, Intervention and Innovation in Sport (CIFID2D), Faculty of Sport, University of Porto, Portugal and Director of the Sport Pedagogy department.