Article

EPER

European Physical Education Review
1–16

Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1356336X17730307
journals.sagepub.com/home/epe
SAGE

The long-term development of volleyball game play performance using Sport Education and the Step-Game-Approach model

Rui Araújo
CIFI[2]D, Faculty of Sport, University of Porto, Porto, Portugal

Peter Hastie
School of Kinesiology, Auburn University, Auburn, AL, USA

Keith R. Lohse
University of Utah, Department of Health, Kinesiology, and Recreation, UT, USA

Cristiana Bessa
CIFI[2]D, Faculty of Sport, University of Porto, Porto, Portugal

Isabel Mesquita
CIFI[2]D, Faculty of Sport, University of Porto, Porto, Portugal

Abstract

The purpose of this study was to analyse 18 Portuguese high school students' game play performance improvements across three hybrid Sport Education-Step-Game-Approach volleyball seasons. Students' play performance at the entry and exit points of each season was evaluated using the Game Performance Assessment Instrument during 2*vs*2 games. A series of hierarchical linear models was then constructed in order to quantify the impact of gender, skill and time on the students' Game Performance Index scores over the three seasons. The best predictive model showed a nonlinear effect of time on student performance such that all participants' levels improved from their first experience at the seventh-grade through to the end of the ninth-grade season. This study has shown the value of implementing multiple seasons of the same sport within Sport Education, as the implementation of three seasons seemed to produce a fading in the gaps between skill levels.

Keywords

GPAI, gender, longitudinal design, physical education, skill level

Corresponding author:
Rui Araújo, CIFI[2]D, Faculty of Sport, University of Porto, Rua Dr. Plácido Costa, 91 – 4200-450, Porto, Portugal.
Email: raraujo@fade.up.pt

Introduction

Researchers of Sport Education, perhaps the most studied of the instructional models listed by Metzler (2011), have identified the need for longitudinal data that extend beyond a single season (Araújo et al., 2014; Wallhead and O'Sullivan, 2005). Specifically, while researchers have recognized the potential of Sport Education for developing students' skill and tactical awareness, to date those studies have involved but a single season (e.g. Araújo et al., 2016; Cho et al., 2012; Farias et al., 2015; Hastie, 1998b; Hastie et al., 2013; Mesquita et al., 2012; Pritchard et al., 2008).

The call for longitudinal studies is grounded in the notion of "time" as a critical element in Sport Education. Given the complexity of the organization of activities within Sport Education (e.g. distributing roles, establishing formal competition, allocating students to teams, among others) and the instructional shift from a teacher-directed to a student-driven model, a number of authors have reinforced the critical nature of longer rather than shorter seasons (see Brunton, 2003; Hastie, 1998a, 1998b; Hastie et al., 2013; Mesquita et al., 2012; Pritchard et al., 2008). By consequence, a study of Sport Education implementation beyond a single season might allow the examination of the critical elements which have the most significant impact on the teaching and learning process (Hastie and Mesquita, 2016). In addition, an application of Sport Education over repeated seasons (where team composition and the sport played remain constant) might promote positive aspects of the social and instructional systems within working groups and consequently improve student learning (Wallhead and O'Sullivan, 2005). The goal of this study was to respond to calls for more longitudinal data collection protocols and the analysis of students' achievement of volleyball game play over multiple Sport Education seasons.

Another gap identified by Sport Education researchers has been the examination of the learning tasks and content development during seasons (Araújo et al., 2014). Specifically, Sport Education is an "outward-focused" model (Hastie and Curtner-Smith, 2006: 23), in which the primary concern is the promotion of a more democratic and inclusive pedagogy and the organization of the lesson, such as teacher and students' roles, formal competition and affiliation to a team (Siedentop et al., 2011). To provide a more inward-focus on the didactics of instruction, Araújo et al. (2014) as well as Hastie and Mesquita (2016) have called for alliances of Sport Education and other instructional models and research designs that allow for the specific examination of the content and learning tasks taught during a season. While there are examples of these hybrid models (e.g. Teaching Games for Understanding: Hastie and Curtner-Smith, 2006; Invasion Games Competence Model: Mesquita et al., 2012; Tactical Games: Pritchard et al., 2014), none of these have been employed with net/wall games (Araújo et al., 2014; Harvey and Jarret, 2014).

The Step-Game-Approach (SGA: Mesquita et al., 2005) may have the potential to fill this gap through the provision of an appropriate framework for the development of game play ability in net/wall sports such as volleyball, badminton or tennis. Framed upon didactical ideas derived from the Teaching Games for Understanding Model (Bunker and Thorpe, 1982), and the Skill Development Approach (Rink, 1993), Mesquita et al. (2005) have introduced an instructional approach in which three types of instructional tasks are implemented. Specifically, these include acquisition tasks (which emphasize the development of a specific skill), structuring tasks (which focus on comprehending the tactical and technical skills of the game but without opposition) and adaptation tasks (in which the goal, action structure and basic tactical features are identical to the full volleyball game) (Mesquita et al., 2015; Pereira et al., 2011). From acquisition tasks to adaptation tasks, the variability of practice is gradually increased to better resemble the actual game conditions (Williams and Hodges, 2005).

In addition, in the SGA, students are presented with progressive (step-by-step) game problems that challenge their capacity for understanding and current performance profiles (Mesquita et al., 2005). More specifically, this model utilizes four stages of learning volleyball beginning from a 1*vs*1 game through to a 4*vs*4 game. The sequence of the game form evolves in such a way that both the formal and functional structure complexity remains aligned with students' tactical understanding and skill level. In this way, a hybrid Sport Education–SGA season may have the potential to facilitate all the affective and social goals of Sport Education, without underestimating the content to be taught during the season.

In addition to the aforementioned gaps, Sport Education researchers have also identified differential effects of gender and skill level on students' achievement (Araújo et al., 2014). While some studies have reported superior learning opportunities for boys and higher skill-level students (e.g. Alexander and Luckman, 2001; Brock et al., 2009; Hastie, 1998a, 1998b), others have shown advantages for girls and lower skill-level students (Carlson and Hastie, 1997; Mesquita et al., 2012). Nonetheless, all of these studies are limited to a single-season experience and only one study has considered the application of hybrid models (Mesquita et al., 2012).

The limitations of previous Sport Education research related to season design and the fact that volleyball is a mandatory element of the physical education curriculum in Portugal contributed to the purpose of the present study. More specifically, this study sought to analyse students' game play volleyball performance improvements across three hybrid Sport Education–SGA seasons. Within the study itself, a subsequent goal was to determine the extent to which students' gender and skill level might have a differentiating impact on students' learning.

Method

Design of the study

The aim of this study was to determine if a particular curriculum intervention would lead to improvement in students' volleyball game play performance across time, and as such the study followed a longitudinal design that included three hybrid Sport Education–SGA seasons over three years. The study began when the students were in the seventh-grade and ended during their ninth-grade. The students had no previous experience with either of the models. Despite some minor changes and withdrawals, the students remained in the same class throughout the entire study. Given the Portuguese physical education curriculum allowed for teaching only one volleyball unit per school year (seventh-grade, eighth-grade and ninth-grade) the three seasons were separated by approximately 12 months. In the time between the three seasons, none of the physical education lessons were taught using Sport Education or SGA models. In addition, during this time students were not involved in extra-curricular volleyball activities. It is important to appreciate that the goal of this study was *not* to determine the superiority of this hybrid model over other forms of instruction, but solely to establish if improvements were possible in this context. Therefore, the longitudinal design did not involve the concurrent application of a control group.

Participants

The participants in this study were 18 students (eight girls and 10 boys) from one physical education class in a school in northern Portugal. These participants were selected using both purposive and convenience sampling criteria (Patton, 2002). They were chosen because of their availability to remain in the same class throughout three school years, as well as their capacity and willingness to participate in the project. At the commencement of the study, these students were aged between 11 and 13 ($M = 11.8$). The school board approved the study, participants were informed of the aims of the study and confidentiality of data was assured. The Ethics Committee of the researchers' university approved the study.

The Sport Education–SGA seasons

The three Sport Education–SGA seasons ranged from 20 to 25 volleyball lessons, with three 45-minute lessons being scheduled per week. Each season followed all the key organization features of Sport Education (i.e. seasons, persisting teams, formal competition, record keeping, festivity and a culminating event), while the learning tasks followed the didactical framework of the SGA. Table 1 provides a complete outline of the season plans for each of the three seasons.

The Sport Education processes. All the seasons were divided into phases of skill and tactical practice, non-consequential practice matches, and formal competition. In the earlier lessons of each season, the procedures embedded in Sport Education were addressed and students were allocated to teams based on a pre-season assessment test. Although teams were balanced in order to have a fair competition, within each team different students had different skill levels. This was important in order to encourage the higher skill-level students to help the lower skill-level students. In the following phase of the season students participated in within-team practices, learning about officiating responsibilities, scrimmage game play, and formal competition. During the second and third seasons, given the differentiating effect of students' skill level on their improvements observed during the first season, the concept of "graded competition" (Siedentop et al., 2011) was implemented. That is, during competitive matches, each team created sub-teams who competed against other students of similar skill levels. In the last lessons of each season a culminating event was organized.

During the three seasons students were allocated to different within-team (e.g. equipment manager and student-coach) and within-match roles (e.g. referees and statisticians) on a rotating basis. The student-coach was chosen by the teacher based on their ability to perform that role and to prevent potential imbalanced power relations between students based on status and gender, the latter feature being found important in research on Sport Education (Brock et al., 2009; Hastie, 1998a). In addition, the formal competition schedule ensured the equitable participation of all students (e.g. same playing time for all students) and students were held formally accountable for their fair-play behaviours during competition. Particularly, teams scored extra points for behaviours regarding fair play during competition, inclusive practice and peer engagement. On the contrary, teams accumulated penalties within the lessons if they showed actions that compromised these behaviours (e.g. not accepting referees' decisions).

The teacher took most of the instructional leadership responsibilities during the first season. Notwithstanding, student-coaches were progressively called upon to take more responsibility for instruction. Between lessons two and five of the first season, student-coaches were only responsible for warm-up tasks. Between lessons seven and ten, the student-coaches began to lead instruction during peer-assisted tasks, and from the eleventh onwards, they began to select learning tasks they deemed appropriate for their teams' performance improvements. During the second season, the student-coaches took responsibility for almost all the peer-assisted tasks earlier in the season, and by the third season all lessons were student-driven.

Table 1. Season plans for the three Sport Education seasons.

First season	Second season	Third season
Lesson 1 and 2 • Explanation of competition format • Allocation of teams and individual roles • Teacher-directed instruction: overhead pass, 1*vs*1 Lesson 3–6 • Teacher-directed instruction: overhead pass, 1*vs*1 Lesson 7–10 • Student-directed instruction: warm-up, overhead pass and 1*vs*1 • Lesson planned by the teacher • Shared teacher- and student-directed monitoring Lesson 11–16 • Student-directed instruction: warm-up, underhand serve and 1*vs*1 (opposition) • Lesson planned by the teacher and students had the opportunity to plan some learning tasks • Formal competition: 1*vs*1 Lesson 17–24 • Formal Competition: 1*vs*1 and 2*vs*2 • Student-directed instruction: warm-up, 1*vs*1, and 2*vs*2 • Lesson planned by the teacher with students having the opportunity to plan some learning tasks Lesson 25 • Culminating event	Lesson 1 • Explanation of competition format and handbooks • Allocation of teams • Student-directed instruction: warm-up • Teacher-directed instruction: overhead pass and 2*vs*2 Lesson 2 and 3 • Student-directed instruction: warm-up, overhead pass and 2*vs*2 • Lesson planned by the teacher • Shared teacher- and student-directed monitoring Lesson 4–6 • Student-directed instruction: warm-up, overhead pass and 2*vs*2 • Lesson planned by the teacher Lesson 7–12 • Student-directed instruction: warm-up, overhead pass and 2*vs*2 • Lesson planned by the teacher Lesson 13–16 • Formal Competition: 2*vs*2 • Student-directed instruction: overhead pass and 2*vs*2 • Lesson planned by the teacher Lesson 17–19 • Formal Competition: 2*vs*2 • Student-coaches and their teams planned	Lesson 1 • Explanation of the competition format and handbooks • Allocation of teams • Friendly competition: 2*vs*2 Lesson 2–4 • Student-directed instruction: 2*vs*2 game and overhead pass in a linear trajectory • Shared teacher- and student-directed monitoring • Formal competition: 2*vs*2 with specific rules • Statistics and refereeing Lesson 5–8 • Student-directed instruction: warm-up, overhead pass in a nonlinear trajectory, verbal communication and 2*vs*2 • Shared teacher- and student-directed monitoring • Formal competition: 2*vs*2 with specific rules to promote verbal communication • Statistics and refereeing Lesson 9–11 • Student-directed instruction: warm-up, transition and opposition • Formal competition: 2*vs*2 with specific rules to promote the approach of the setter to the net in the 2nd touch • Statistics and refereeing (more than one touch for player and three for team) Lesson 12–14 • Student-directed instruction: warm-up and forearm pass • Student-directed monitoring • Formal competition: 2*vs*2 with specific rules to promote opposition • Statistics and refereeing

(continued)

Table 1. (continued)

First season	Second season	Third season
	content and learning tasks to the lesson Lesson 20 • Culminating event	Lesson 15–21 • Student-coaches and their teams planned content and learning tasks to the lesson • Student-directed monitoring • Formal competition: 2*vs*2 with larger courts • Statistics and refereeing Lesson 22 • Culminating event

The SGA approach. The 2*vs*2 game form was considered the most suitable for diagnosing the skill level of the participants (Mesquita et al., 2005). On one hand, it allows for the identification of basic performance levels when players are unable to minimally sustain the ball (step one). On the other hand, it also identifies performance levels above that range from the minimum capacity to support the ball (step two) to the capacity of organizing game actions (early third step), or even the ability to differentiate the most suitable solutions for the game situation (early fourth step). Based on the students' performance on a 2*vs*2 assessment test conducted prior to each season, the first step was applied. The main goal of this step was to understand the logic of the simplest game form (1*vs*1 game), namely two main tactical skills were taught in this step: *intervention* (place the body where the ball is) and *opposition* (play the ball to the vulnerable place of the opponent's court, both in the serve and in the attack). In addition, two technical skills were taught, namely the overhead pass and the underhand serve.

The second step of the SGA was taught during the second season, namely the ability to cooperate with a partner while challenging others in the 2*vs*2 game. Three specific tactical skills were introduced: (i) watching the opponents' placement; (ii) verbal communication; and (iii) assigning accountability zones. The forearm pass was also introduced, since the students needed other technical skills to play the ball with lower trajectories.

It was during the third season that students tended to separate into two different skill levels within the class. By consequence, although 2*vs*2 games were maintained, variations to the learning tasks and game rules were implemented in order to adjust the difficulty of learning tasks across different skill levels. Graded competition was maintained in terms of matching similarly skilled opponents, but the games themselves differed in terms of rules, court dimensions and scoring.

Instructional and treatment validity

Itemizing teacher and learner processes. A 10-item checklist (Table 2) was adapted from Hastie et al. (2013) and Pereira et al. (2011) in order to confirm the behavioural fidelity of the teacher's instruction according to both Sport Education and the SGA model. Two trained observers with extensive research in instructional models observed five randomly selected lessons of each season (in a total of 15 lessons observed) and recorded the presence of those items. Items one, three, four, five, seven, eight and nine are characteristics of Sport Education, while items two, six and 10 are

Table 2. Instructional checklist.

1. Groups of students go to designated home areas and begin warming up with their group.
2. The tasks under observation are basic game forms, game-like tasks or acquisition tasks, and the time spent in acquisition tasks is reduced to the minimum necessary.
3. Students practice together with their group/team under the direction of a peer leader.
4. The content of the task is related with the stage of the SGA that is being taught during the season.
5. Students remain a part of easily identifiable groups throughout the lesson and throughout different tasks.
6. All the tasks are related with the small-sided game that is being taught.
7. Performance records are kept by students.
8. Students perform specialized tasks within their group/team.
9. Student performance scores count toward a formal and public scoring system.
10. Modifications to the full game are performed.

Adapted from Hastie et al. (2013) and Pereira et al. (2011).

related to the SGA model. A 100% agreement between the two observers confirmed the absence of doubt regarding the instructional model used in each lesson.

Contextual and operational requirements. In order to guarantee a proper implementation of an instructional model, some favourable contextual conditions should be guaranteed, such as teacher expertise and operational requirements (Hastie and Casey, 2014; Metzler, 2011). With regard to teacher expertise, despite his experience as a physical educator for over 20 years, together with two years' experience implementing Sport Education, the teacher participated in a workshop during the school year prior to this study. This workshop comprised lectures that focused on several instructional models used in physical education, discussions of the conceptualization and purpose of Sport Education, and the specific application of Sport Education to volleyball as well as highlights of the outcomes of Sport Education research. In addition, during the year prior to the study the teacher accomplished his master's degree in which he applied a hybrid Sport Education–SGA volleyball season with classes not involved in this study. With specific regard to operational requirements, the space (at least three volleyball courts in all lessons) and material available (balls, cones, score sheets, whistles, etc.) to practice were guaranteed in order to create the best possible conditions to enhance students' learning throughout the implementation of the study.

Data collection

A systematic observation of the video records of students' behaviours while playing a 10 minute 2*vs*2 game was used in order to analyse students' play performance at the entry (PreT) and exit (PosT) of each season. For the purposes of the present study, the Game Performance Assessment Instrument (GPAI; Oslin et al., 1998) adapted to volleyball requirements (Mesquita et al., 2005) was used. More specifically, the tally scoring method recommended by Mitchell et al. (2006) was applied, in which students' game play actions were assessed as appropriate or inappropriate responses. When the amounts of those actions were totalled, the Game Performance Index was calculated as follows: (Decision Making+Adjustment+Skill Efficiency+Skill Efficacy)/4. In particular, Decision Making refers to the ability to choose the most appropriate response to a tactical problem (e.g. sending the ball to the vulnerable space on the opponent's court). Adjustment involves the ability to change position to play the ball. Skill Efficiency refers to the quality of the technical execution of a skill (e.g. raise and cup hands during the set), while Skill Efficacy is the outcome of the player's action (e.g. winning a point, continuity of play, or an error). A total of 39,208 game actions over 3,360 minutes of observations were coded across the three seasons.

In order to specifically and reliably identify appropriate and inappropriate actions, narrow definitions of these behaviours were developed (Memmert and Harvey, 2008). Beyond that, the reliability of the data was examined through intra-observer (15 days after the first observation) and inter-observer testing procedures (performed by a second observer). Following the recommendations of Tabachnick and Fidell (2007), two researchers who were experienced in volleyball as both coaches and players assessed 10% of the complete data set. Intraclass Correlation Coefficients (ICC; Atkinson and Nevill, 1998; Baumgartner and Jackson, 1995) ranged between 87% and 98%, which are above the percentages suggested by Fleiss et al. (2003) as acceptable levels.

Data analysis

In order to quantify changes in the Game Performance Index over time, a series of hierarchical linear models were constructed. For these models, Time (measured in pre-/post-season assessments over three years for six total time points) was nested within students. These mixed-effect regressions treated Time, Skill Level (–0.5 = low skill; 0.5 = high skill), and Gender (–0.5 = Male; 0.5 = Female) as fixed effects.

Mixed-effect linear regressions was conducted using R and R Studio (R Core Team, 2014) using the "lme4" (Bates et al., 2014) and "dplyr" (Wickham and Francois, 2014) packages. To model changes in the Game Performance Index a "step-up" procedure was used in which variables were added to successive models. Models were compared based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Wald Tests for the change in deviance with $\alpha = 0.05$ (Long, 2012). The first model was a "random intercepts" model that establishes a baseline level of deviance against which subsequent models were compared. Subsequent models added factors of Time, Skill Level, the quadratic effect of time, gender, and their interactions.

Assessment of power. With respect to the use of hierarchical linear models, Snijders (2005) notes that the power in any multi-level model is largely based on the total number of observations. Since squared standard errors are inversely proportional to sample sizes, the required sample size for a multi-level design will be determined by the sample size that would be required for a simple random sample design multiplied by the design effect. This is particularly true for "level-one" variables (the power of which is almost completely determined by the total number of observations, regardless of how they are nested).

For a level-one variable (in this study, time) the total size is given by $m*n$ where m is the number of clusters (students) and n is the number of observations per cluster (time points) leaving us with a total level-one sample size of 108. The required conventional sample size for a "simple random sample design" using a program like G*Power (3.1.9.2) would be "n" of 55, if one sought to detect an f-squared = 0.15, assuming alpha = 0.05, power = 0.80, and the number of tested predictors = 1. Thus, even assuming that the ICC is 0, our $m*n = 108$ (which is greater than 55), results in a power greater than the 80% necessary to detect a moderate effect-size for the effect of time.

It is recognized, however, that this power is reduced when the examination moves to level-two variables (i.e. main effect of skill) or interactions between level-two and level-one variables (i.e.

Table 3. Game play performance across three seasons.

	Pre-test *M* (SD)	Post-test *M* (SD)	Δ
Season 1			
Population	.51 (.17)	.73 (.09)	.22
Males	.60 (.12)	.76 (.08)	.16
Females	.38 (.14)	.69 (.10)	.31
Higher skilled	.66 (.06)	.77 (.09)	.11
Lower skilled	.40 (.12)	.70 (.09)	.30
Season 2			
Population	.71 (.11)	.77 (.08)	.06
Males	.74 (.11)	.80 (.07)	.06
Females	.67 (.11)	.76 (.10)	.09
Higher skilled	.74 (.11)	.80 (.07)	.06
Lower skilled	.68 (.10)	.76 (.09)	.08
Season 3			
Population	.75 (.10)	.81 (.07)	.06
Males	.79 (.11)	.84 (.05)	.05
Females	.71 (.10)	.77 (.09)	.06
Higher skilled	.78 (.11)	.83 (.05)	.05
Lower skilled	.73 (.10)	.79 (.09)	.05

M: mean; SD: standard deviation.

the interaction of skill and time). However, these are not the primary analyses of interest, and in this study are labelled more as exploratory. Furthermore, while this lack of power greatly increases the chance of missing a level-two effect or a level-two by level-one interaction, it would not undermine statistically significant interactions that might be found. In summary then, the design of this study has more than adequate power to look at the level-one effect of time, but we need to be cautious in interpreting any level-two effects or interactions. Such interactions may suggest directions for future research, but could hardly be considered "confirmatory" research findings.

Results

The results of the students' game play performances are presented in Table 3, while Table 4 shows the outcomes of the various steps in the multi-level modelling step-up procedure. As can be seen from Table 3, students made improvements from pre- to post-test during each season, and showed solid retention, with only marginal decrements between seasons. Figure 1 shows the change curves for each participant as well as the mean (coloured line) at each assessment point.

An examination of Table 4 shows that Model 4 provided the best fit of all tested models. Within Table 4, the *p*-value refers to the statistical comparison of the given model to the model above it (starting at Model 0 as a baseline measure of error). Model 4 predicted game performance as a function of the continuous variables of Time, Time^2, the categorical variable of Skill (lower = –0.5; higher = 0.5), and the interaction of Skill with Time. The addition of the interaction of Skill

Table 4. Model comparisons.

	df	AIC	BIC	logLik	Deviance	χ^2	χdf	*p*
M0	3	767.36	775.02	−380.68	761.36			
M1	6	729.66	744.99	−358.83	717.66	43.69	3	1.75e-09
M2	7	724.90	742.78	−355.83	710.90	6.76	1	0.0093
M3	8	719.57	740.00	−351.78	703.57	7.33	1	0.0068
M4	12	704.89	735.54	−340.45	680.89	22.67	4	0.0001
M5	16	709.24	750.10	−338.62	677.24	3.65	4	0.4550

AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion.

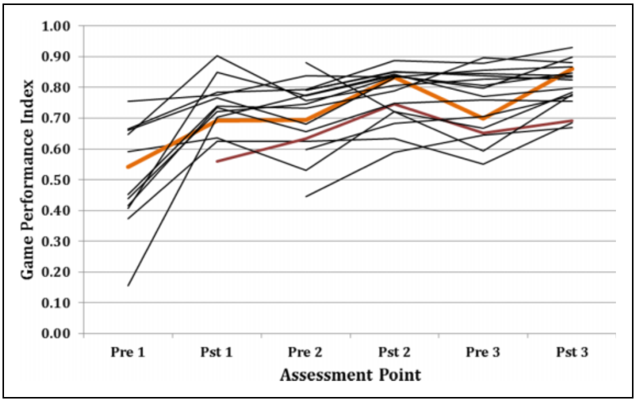

Figure 1. Individual and mean change curves. Pre: pre-test; Pst: post-test.

and Time in Model 4 was a statistically significant improvement beyond Model 3, which contained all other variables except the interaction (Wald Test $\chi^2 = 22.67$, $p = 0.0001$; note also the decreased AIC and BIC). This model suggests that there was a *nonlinear* effect of Time on student performance, controlling for Skill, such that all participants improved, but the rate of improvement declined over time (see Table 5). The fixed effects in this model should be interpreted the same as coefficients in linear regression.

There was also a statistically significant effect of Skill, in which lower-skilled players had significantly lower game performance scores at the first assessment than higher-skilled players (17.21 points lower at the intercept). Furthermore, there was a significant interaction of the Time and Skill, such that the rate of improvement for the lower-skilled players was steeper than the rate of improvement for the higher-skilled players. It should be noted that adding students' Gender to the model (along with the interactions of Gender with Time and Skill) did not explain any additional variance (see Model 5 in Table 4). By all model fit indices (AIC, BIC, and change in deviance), adding Gender as a predictor did not improve the model controlling for Skill and Time, but it should be noted that given the limited number of participants, this study was underpowered to detect all but the largest between-subject effects, such as Skill and Gender.

Given that the nonlinear model provided the best fit, Table 5 shows the values of the various fixed effects within the model, the predictions of which are shown in Figure 2. Examination of this

Table 5. Model summary.

Fixed Effects				
	Estimate	Std. Error	*t* (17)	*p*
Intercept	57.64	2.14	26.91	< 0.001
Time	13.28	3.19	4.16	< 0.001
Skill	17.21	3.82	4.51	< 0.001
$Time^2$	−1.26	0.42	−3.04	.003
Time: Skill	−2.39	0.99	−2.42	.014
Random Effects				
Groups	Name	Variance	SD	Cor
Subject	Intercept	22.57	4.75	
	Time	92.26	9.61	−0.31
	$Time^2$	1.46	1.21	0.30
Residual		50.38	7.10	

Note. Estimates and standard error terms are presented as percentages. Cor = Correlation between the slope and intercept random-effects.

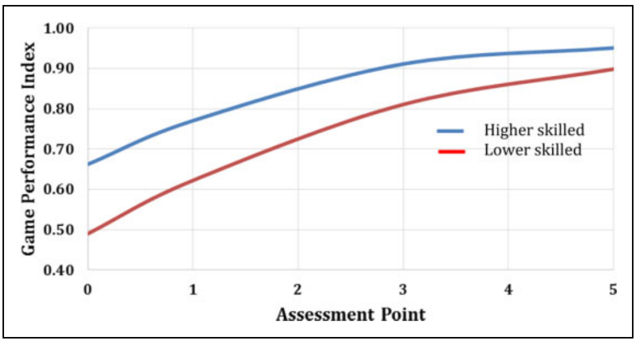

Figure 2. Predictions from the nonlinear model effect of time.

figure shows that while lower-skilled students began at a lower entry point in the beginning of the seventh-grade, by the end of the ninth-grade season they were approaching the game play performances of the higher-level students.

Discussion

The aim of the present study was to analyse students' game play performance improvements across three hybrid Sport Education–SGA seasons. In addition, this analysis took into account students' gender and skill levels. The findings of the study showed that students of both genders and skill levels improved from their first experience with the hybrid model during the seventh-grade through to the third iteration during the ninth-grade. The extension of Sport Education implementation beyond a single-season experience would therefore be seen as fundamentally valuable in terms of student learning of a specific sport.

The sporadic application of Sport Education seasons or the application of a single season over time provides little time to control all the variables that could interfere with the teaching and learning process (Brunton, 2003; Hastie and Mesquita, 2016). The complexity of the organization of activities related with Sport Education could reduce the time that students had for practice the skills (during learning tasks) and consequently negatively influence their learning. Notwithstanding, in the present study it could be assumed that students became progressively familiarized with these organizational and instructional issues throughout the three seasons, spent less time on these tasks and, consequently, spent more time on learning tasks. For instance, during the first season, students spent considerable time changing from one game to another within the competition phase. However, in the following seasons they become more familiarized with the organizational issues and the teacher could dedicate more time for practice or competition within the lesson.

The combined use of Sport Education and the SGA also seemed to make positive contributions to students' improvements throughout the three years. Indeed, from the beginning of the study students were presented with a number of pedagogies that could enhance their learning. First and foremost, game forms and technical and tactical skills were introduced according to the problems disclosed by the game play within a step-by-step approach (Mesquita et al., 2005). This phased progression meant that the level of complexity remained aligned with the students' tactical understanding and skill level (Hastie and Mesquita, 2016; Mesquita et al., 2015) throughout the three years.

In addition, skill practices across the three seasons were presented mostly as game-like situations, notably through the use of structuring and adaptation tasks. This allowed students to practice within the same technical and tactical structure and the variability of practice that resembled the actual game conditions (Mesquita et al., 2015; Pereira et al., 2011). The integration of technical and tactical components promoted significant gains in overall game performance (which promotes both technical and tactical skill development) and served to minimize the technical/tactical dualism previously discussed by other researchers (French et al., 1996; Turner and Martinek, 1999). These results therefore support the need for holistic-grounded programmes rather than the instruction of isolated skills as the preferred building blocks of game competence (Mesquita et al., 2005).

Findings according to students' gender

As noted earlier, previous research on Sport Education has shown a differentiating effect of students' gender on their skill and game play improvements. In these three hybrid Sport Education–SGA seasons, however, both boys and girls showed similar game performance improvements from the beginning of the study (seventh-grade) until the end of the three years (ninth-grade). This suggests that the application of more than one Sport Education season over time might serve to promote more equitable learning opportunities. While a number of equity-promoting features were instigated during the first season, the fact that these were sustained and consolidated throughout the second and third seasons may have given the students more time to appreciate them. Students started to feel comfortable seeing all their teammates (regardless of gender) participating in all the learning tasks, assuming key positions such as student-coach or referee, and support activities such as statistician or manager.

In addition, the affiliation achieved through the notion of persisting teams embedded within Sport Education may have also contributed to the full participation of all students, a feature found in previous research on Sport Education. Specifically, by remaining on the same team across the three years and by practicing together, it is more likely they would have supported each other and held each other accountable. This "content-embedded accountability" (Hastie and Siedentop, 2006: 219) has been shown to promote students' task engagement and achievement of other common performance goals regardless of their gender (Hastie et al., 2013; Pereira et al., 2015; Pritchard et al., 2008, 2014).

Findings according to students' skill level

In this study, all students improved independently of their skill level during the first and second seasons. This can be observed in Figure 2, in which both higher and lower-skill level students improved in a similar way during these two seasons. This was accomplished by the adjustment of content and learning tasks according to the students' skill levels, mainly through the use of task modification. In particular, despite the participation of both skill levels in the same tasks, the difficulty of the task was adjusted to students' skill level. For instance, in a task in which the goal was to sustain ball circulation across the net, high skill-level students could only touch the ball one time while lower skill-level students were allowed multiple contacts. In addition, the use of the graded competition implemented during the second season also contributed to the reduction of difference between skill levels during game play.

The results also show that despite the improvements of all students, lower skill-level students achieved greater improvements across the entire intervention. In fact, by the end of this third season, the lower skill-level students were approaching the game performance scores of higher skill-level students. An alternate proposition is that those higher-skilled students were perhaps reaching a ceiling effect in terms of the possible quality of their game play. It may well have been the case that in this third season, the graded competition should not only have included the adaptation of rules, but also incorporated a completely different game form appropriate to the students' skill level. That is, the higher-skilled students may have needed a more challenging game (e.g. 3*vs*3) in order for everyone to be working according their needs and levels, and thereby challenging them to work within their "zone of proximal development" (Vygotsky, 1978). Overall, this study clearly shows the value of using more than one season over time to effectively guarantee the validity of learning measurement. Indeed, learning is a process that occurs over time, rather than simply at the time of a single measurement of performance. Therefore, repeating volleyball over three successive years may have the potential to enable learning to be measured as it could be determined whether skills, knowledge, concepts and competencies were retained from one moment to the next. Hence, longitudinal research in Sport Education is crucial in order to analyse the learning impact of the model because it might allow researchers to detect and establish the nature of students' learning over time (Ruspini, 1999), thereby generating dynamic data and clarifying the direction as well as the magnitude of any changes (Hakim, 1987).

Conclusions

This study has shown the value of implementing multiple seasons of the same sport within Sport Education, particularly when team composition is kept constant. The results show that all students achieved significant improvements in game play with each successive season, regardless of their gender and skill level, although it was noted that the higher skill-level students may have been under-challenged by the third year.

While this study provided an analysis of more explicit and measurable variables of time, gender and skill related to the goals of Sport Education, it offers but one example and the results should be treated cautiously, particularly with the interpretation of variables apart from time. Certainly, while labour intensive, replications of this study are warranted, with larger sample sizes that would allow for the examination of the effect of differentiated versions of games used for those students with lower and higher skills. That being said, the present study strongly reinforces the need to move forward from short-duration units to those of longer duration (Hastie and Mesquita, 2016).

Declaration of Conflicting Interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Portuguese Foundation for Science and Technology (FCT) / POPH / QREN / European Social Fund under Grant number SFRH/BD/72361/2010.

References

Alexander K and Luckman J (2001) Australian teachers' perceptions and uses of the sport education curriculum model. *European Physical Education Review* 7: 243–267.

Araújo R, Mesquita I and Hastie P (2014) Review of the status of learning in research on sport education: Future research and practice. *Journal of Sports Science and Medicine* 13: 846–858.

Araújo R, Mesquita I, Hastie PA, et al. (2016) Students' game performance improvements during a hybrid sport education–step-game-approach volleyball unit. *European Physical Education Review* 22: 185–200.

Atkinson G and Nevill A (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sport Medicine* 26: 217–238.

Bates D, Mächler M, Bolker B, et al. (2014) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 55: 1–9.

Baumgartner TA and Jackson AS (1995) *Measurement for Evaluation in Physical Education and Exercise Science*. Dubuque, IA: Brown & Benchmark.

Brock SJ, Rovegno I and Oliver KL (2009) The influence of student status on student interactions and experiences during a sport education unit. *Physical Education and Sport Pedagogy* 14: 355–375.

Brunton J (2003) Changing hierarchies of power in physical education using sport education. *European Physical Education Review* 9: 267–284.

Bunker D and Thorpe R (1982) A model for the teaching of games in secondary schools. *Bulletin of Physical Education* 18: 5–8.

Carlson T and Hastie PA (1997) The student social system within sport education. *Journal of Teaching in Physical Education* 16: 176–195.

Cho O, Richards K, Blankenship B, et al. (2012) Motor skill development of students enrolled in a sport education volleyball season delivered by in-service physical education teachers. *Physical Educator* 69: 375–394.

Farias C, Mesquita I and Hastie PA (2015) Game performance and understanding within a hybrid sport education season. *Journal of Teaching in Physical Education* 34: 363–383.

Fleiss J, Levin BA and Paik MC (2003) *Statistical Methods for Rates and Proportions*. Wiley, NJ: John Wiley and Sons.

French K, Werner P, Rink J, et al. (1996) The effects of a 3-week unit of tactical, skill, or combined tactical and skill instruction on Badminton performance of ninth-grade students. *Journal of Teaching in Physical Education* 15: 418–438.

Hakim C (1987) *Research Design: Strategies and Choices in the Design of Social Research*. London, UK: Routledge.

Harvey S and Jarret K (2014) A review of the game-centred approaches to teaching and coaching literature since 2006. *Physical Education and Sport Pedagogy* 19: 278–300.

Hastie P 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: 422–431.

Hastie P and Mesquita I (2016) Sport-based physical education. In: Ennis C (ed) *Routledge Handbook of Physical Education Pedagogies*. London, UK: Routledge, pp.367–379.

Hastie PA (1998a) The participation and perceptions of girls within a unit of sport education. *Journal of Teaching in Physical Education* 17: 157–171.

Hastie PA (1998b) Skill and tactical development during a sport education season. *Research Quarterly for Exercise and Sport* 69: 368–379.

Hastie PA, Calderón A, Rolim R, et al. (2013) The development of skill and knowledge during a sport education season of track and field athletics. *Research Quarterly for Exercise and Sport* 84: 336–344.

Hastie PA and Curtner-Smith MD (2006) Influence of a hybrid Sport Education - Teaching Games for Understanding unit on one teacher and his students. *Physical Education and Sport Pedagogy* 11: 1–27.

Hastie PA and Siedentop D (2006) The classroom ecology paradigm. In: Kirk D, Macdonald D and O'Sullivan M (eds) *The Handbook of Physical Education*. London: Sage Publications, pp.214–226.

Long JS (2012) *Longitudinal Data Analysis for the Behavioral Sciences using R*. Thousand Oaks, CA: Sage Publications.

Memmert D and Harvey S (2008) The Game Performance Assessment Instrument (GPAI): Some concerns and solutions for further development. *Journal of Teaching in Physical Education* 27: 220–240.

Mesquita I, Afonso J, Coutinho P, et al. (2015) Modelo de abordagem progressiva ao jogo no ensino do voleibol: conceção, metodologia, estratégias pedagógicas e avaliação. In: Tavares F (ed) *Jogos Desportivos Coletivos: ensinar a jogar*. 2 ed. Porto: Editora FADEUP, pp.73–122.

Mesquita I, Farias C and Hastie PA (2012) The impact of a hybrid Sport Education-Invasion Games Competence Model soccer unit on students' decision making, skill execution and overall game performance. *European Physical Education Review* 18: 205–219.

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: 469–492.

Metzler MW (2011) *Instructional Models for Physical Education*. Scottsdale, Arizona: Holcomb Hathaway, Publishers, Inc.

Mitchell S, Oslin J and Griffin L (2006) *Teaching Sport Concepts and Skills: A Tactical Games Approach*. Champaign, IL: Human Kinetics.

Oslin J, Mitchell S and Griffin L (1998) The Game Performance Assessment Instrument (GPAI): Development and preliminary validation. *Journal of Teaching in Physical Education* 17: 231–243.

Patton MQ (2002) *Qualitative Research & Evaluation Methods*. London: Sage Publications.

Pereira F, Graça A, Blomqvist M, et al. (2011) Instructional approaches in youth volleyball training settings according to player's age and gender. *International Journal of Sport Psychology* 42: 227–244.

Pereira J, Hastie PA, Araújo R, et al. (2015) A comparative study of students' track and field technical performance in sport education and in a direct instruction approach. *Journal of Sports Science and Medicine* 14: 118–127.

Pritchard T, Hawkins A, Wiegand R, et al. (2008) Effects of two instructional approaches on skill development, knowledge, and game performance. *Measurement in Physical Education and Exercise Science* 12: 219–236.

Pritchard T, McCollum S, Sundal J, et al. (2014) Effect of the Sport Education Tactical Model on coeducational and single gender game performance. *The Physical Educator* 71: 132–154.
R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: the R Foundation for Statistical Computing.
Rink J (1993) *Teaching Physical Education for Learning*. St. Louis: Mosby.
Ruspini E (1999) Longitudinal research and the analysis of social change. *Quality and Quantity* 33: 219–227.
Siedentop D, Hastie PA and Van der Mars H (2011) *Complete Guide to Sport Education*. Champaign, IL: Human Kinetics.
Snijders TAB (2005) Power and sample size in multilevel linear models. In: Everitt BS and Howell DC (eds) *Encyclopedia of Statistics in Behavioral Science. UK*: Wiley: Chichester, pp.1570–1573.
Tabachnick BG and Fidell LS (2007) *Using Multivariate Statistics*. Boston: Pearson.
Turner AP and Martinek TJ (1999) An investigation into teaching games for understanding: Effects on skill, knowledge, and game play. *Research Quarterly for Exercise and Sport* 70: 286–296.
Vygotsky L (1978) *Mind in Society*. Cambridge: Harvard University Press.
Wallhead T and O'Sullivan M. (2005) Sport education: Physical education for the new millennium? *Physical Education and Sport Pedagogy* 10: 181–210.
Wickham H and Francois R (2014) *dplyr: A Grammar of Data Manipulation*. New York: Springer.
Williams AM and Hodges N. (2005) Practice, instruction and skill acquisition in soccer: Challenging tradition. *Journal of Sports Sciences* 23: 637–650.

Author biographies

Rui Araújo is a Lecturer at the Faculty of Sport, University of Porto and researcher at CIFI^{2}D, Porto, Portugal.

Peter Hastie is a Professor at Auburn University, Auburn, AL, USA.

Keith R. Lohse is an Assistant Professor at Auburn University, Auburn, USA.

Cristiana Bessa is a PhD student at CIFI^{2}D, Faculty of Sport, University of Porto, Porto, Portugal.

Isabel Mesquita is a Professor of the Faculty of Sport, University of Porto and researcher at CIFI^{2}D, Porto, Portugal.