Mahedero, M. P., Calderón, A., Hastie, P. A., & Arias-Estero, J. L. Grouping students by skill level in mini-volleyball: Effect on game performance and knowledge in sport education.
https://doi.org/10.1177/00315125211021812

Perceptual and Motor Skills. ISSN: 0031-5125 / 1558-688X.
Vol. 128(4). Páginas, inicial: 1851. final: 1871. Fecha: Julio de 2021.

Grouping Students by Skill Level in Mini-Volleyball: Effect on Game Performance and Knowledge in Sport Education

Abstract

The purpose of this study was to explore any differences in game performance variables and knowledge among a cohort of high school students who participated in either homogeneous or heterogeneous skill level groups (N = 126) across a 12-lesson mini-volleyball sport education unit of study. This study followed a mixed-methods approach using a quasi-experimental pre-test / post-test design. The quantitative variables analyzed were decision making, skill execution, game performance, game involvement, and game knowledge. We also evaluated students' performance qualitatively, employing two methods: (a) experts' analysis of students' game performance, and (b) students' and teachers' perceptions of students' performance. We analyzed quantitative data through a series of paired samples t-tests comparing pre- and post-test scores according to the grouping strategy. Students became more competent in their game play and more knowledgeable in their technique, the sport's rules, tactical awareness, and general game knowledge. However, grouping students by skill level had no impact on gains in game performance variables and knowledge. Although sport education literature shows a preference for heterogeneity in ability-based grouping, within our data both heterogeneous and homogenous groups of higher and lower skilled students achieved improvements in game performance and knowledge, leading us to suggest that teachers who are interested in grouping students to create a meaningful learning experience should consider criteria other than student ability.

Keywords: physical education, game involvement, ability grouping, decision-making, sport pedagogy, teaching-learning contexts, sport education

Introduction

Ability grouping is defined as any classroom organization plan that is intended to reduce heterogeneity among instructional groups or teams (Slavin, 1990). Although ability grouping has been generally accepted as a useful tool, it must be used precisely if it is to have a positive effect on student achievement (Slavin & Karweit, 1985). Previous work has defined two main forms of ability grouping – between-class and within-class. Within-class ability grouping makes it possible to adapt instructional techniques to the needs of the group or makes possible individual instruction to students with different skill levels (Wilkinson et al., 2016). As reported by Tieso (2003), flexible ability grouping, combined with appropriate curricular revision or differentiation, may result in substantial achievement gains for both average and high ability learners.

In physical education, several authors have indicated that students learn most effectively when they practice at an appropriate difficulty level in homogeneous groups (e.g., Silverman et al., 1993, 1998). Nevertheless, authors of a recent study suggested that teachers should be mindful of the affects that group composition can have on learning, and they should use different strategies according to unique situational demands (Barker et al., 2015). Overall, ability-based groups have seemed to make sense for individualized learning and allow for the creation of a more intimate learning environment, especially for lower skilled students (Haynes et al., 2008; Ward et al., 2019).

In contrast, a central tenet of sport education within physical education is the value of creating a number of small, heterogeneous learning groups or teams (Siedentop et al., 2011). Within heterogenous teams, students' varied roles and responsibilities are expected to contribute to their team's successes, and students are expected to help and learn from each other. Furthermore, the philosophy of sport education incorporates the idea that all students get an equal opportunity to play, as expedited through the use of small-sided contests. With that in mind, one element included within the features of sport education is that of 'graded competition' (Siedentop et al., 2011, p. 104). In graded competition, leagues are arranged that match students of similar skill level against one another, and in some cases, even the game forms are different across the competition levels.

Of those investigations that have studied differentiating students' skill levels, the results tend to have been equivocal. Mesquita et al. (2005) found that a hybrid combination of sport education and Teaching Games for Understanding approaches to a volleyball unit was more effective in promoting performance improvements among seventh grade lower skilled than higher skilled students. Furthermore, Mesquita et al. (2012) highlighted improvement among fifth grade students with different skill levels during a hybrid soccer unit using a sport education-Invasion Games Competence Model. Again, in that case, lower (versus higher) skill level students achieved the greatest gains. Of note, in both of these studies, the authors recognized that some students' lower entry performance provided more room to show progress than was the case for students with an initially higher technical background, for whom improvement may have been bounded by a ceiling effect. However, Mesquita et al. (2005, 2012) followed no grouping strategies. In contrast to outcome studies favoring lower skilled students, Mahedero et al. (2015) reported that the improvements of both the highest and lowest skilled eighth grade students were less significant than those of students with more moderate skill levels when participants were divided into mixed-ability teams based upon their pretest scores.

Recently, Hastie et al. (2017) tested the hypothesis that graded competition would increase opportunities for game involvement and success of both higher and lower skilled fourth grade students within two classes-groups. In their study, while the season format for both classes-groups was the same, the team composition varied between the two. In the first class-group, all teams were heterogeneous throughout the entire unit, while, in the second class-group, two homogeneous leagues were formed (one consisting of teams with all higher skilled students and the other with teams of all lower skilled students). Hastie et al.'s (2017) key finding was that lower skilled students were at a disadvantage in terms of their engagement rates and playing efficiencies when they participated alongside higher skilled classmates. While less severe, there were also decrements in higher skilled students' success rates and efficiencies in these mixed-skill conditions. Consequently, they concluded that while higher skilled students could thrive in either homogenous

or mixed level competitions, lower skilled students benefitted from playing against students of similar skill levels.

An analysis of the method sections of all competence-based papers within sport education research would suggest that all students participated in heterogeneous teams. Furthermore, Hastie et al. (2017) grouped students by skill level but only for the formal competition phase at the end of the season but not during the early phases (skill practice, pre-season, etc.). In this context, our purpose in this study was to explore whether there were differences in knowledge and game performance variables among a cohort of high school students participating in either homogeneous or heterogeneous skill level groups across an entire mini-volleyball sport education unit. Due to some contradictory findings and conclusions reported in previous research, we made no *a priori* hypotheses.

Method

Research Design

This study followed a mixed-methods approach, using a quasi-experimental pre-test – post-test design. This design is appropriate to achieve a deeper and wider perspective of the data as well as to search for convergence and corroboration of the results on a similar study (Morse & Niehaus, 2009). Pre-testing on decision making (DM), skill execution (SE), game performance (GP), game involvement (GI) and game knowledge (GK) took place before instruction began, while post-testing followed completion of instruction. In addition, student performance was evaluated qualitatively by expert descriptive analyses of students' game play; and we interviewed both students and teachers for their perceptions of students' possible game improvements.

Participants and Context

Participants in this study were 126 students from five coeducational public, urban, by-lingual and non-religious high school classes within a Spanish school (66 boys, 60 girls; *M* age = 16.6. years). Of a potential population of 148 students, all of these 126 students attended at least 10 of the 12 sessions of the teaching unit and completed all assessment instruments (see below for

those details). Two female teachers (aged 30 and 34 years) were responsible for all instruction. Both teachers had been teaching physical education in high schools for at least five years, and each one had had previous experience with at least one sport education season and with innovative teaching through student-centerd approaches. Students, parents and teachers all completed informed consent forms (giving participants the right to withdraw and assuring them of confidentiality). The authors' University's Research Ethics Committee approved the study protocol, conducted according to the Helsinki Declaration.

Instruction and Treatment

We followed Hastie and Casey's (2014) guidelines on fidelity in models-based practice research, including: (a) a rich description of the curricular elements of the unit, (b) a detailed validation of model implementation, and (c) a detailed description of the program context, including the teachers' and students' previous experiences with the model or with models-based practice. An external review of the unit plan confirmed all key features of sport education considering those listed in the validity check developed by Sinelnikov (2009). This checklist was completed in-vivo by an independent researcher, observing lessons three, five, seven, and nine. There was more than 85% agreement between these two researchers on the different aspects analyzed, both in the planning and the implementation of the unit.

Curricular Elements of the Sport Education Unit

The mini-volleyball unit took place twice a week over six weeks for a total of 12 lessons (see Table 1), with each lesson scheduled for 55 minutes. The sport education season began with three lessons that were initially teacher-directed and that focused on the skills and tactics of mini-volleyball and were aimed at developing a game of 4-vs-4 mini-volleyball with a degree of ball continuity. In these early lessons, students were introduced to the rules and officiating procedures of the game. The next five lessons consisted of small-sided games within peer-teaching instructional tasks. The unit concluded with a formal competition spanning three lessons that took the form of a

no elimination round-robin challenge. After the final games, a closing ceremony provided a formal end to the unit and in which various awards were presented to students.

[Insert Table 1 about here.]

Description of the Program Context

Teacher Expertise. Sport education requires the teacher to have mastered a significant number of managerial protocols and routines to successfully implement a season (Hastie & Casey 2014). Consequently, both teachers received formal instruction and training in sport education following directions from Dyson et al. (2004). The comprehensive experience conducted included: (a) training by different sport education experts (academics and practitioners with a practical and theoretical knowledge on sport education and several publications on the topic in peer review journals) on the basic principles of sport education; (b) training in the development of a specific module of sport education; (c) observations of the actual implementation of a teaching unit (badminton and/or mini-volleyball) with sport education in high school students; and (d) experience in implementing a sport education unit with classes other than those used in this study.

Students' Previous Experience. None of these students had prior experiences with sport education. They experienced only direct instruction methodologies during their high school physical education class. While the students did have previous experience with volleyball (between one and four years), game play during these lessons followed the full 6-vs-6 game format. Only 21 students indicated that they played volleyball outside of school, and among these, 19 reported that the format of this play was informal and not associated with a sports club.

Allocation of Students to Teams. Two weeks before instruction began, we collected quantitative data on the students' game performance variables during 4-vs-4 mini-volleyball games. All games were videotaped, and student performance was subject to analysis using the Game Performance Assessment Instrument (GPAI, Oslin et al., 1998). Based upon these scores, all students were classified as being either higher or lower skilled (we used a median split of scores to the GPAI, with 50% of students in each category). Using this classification, all students were then

randomly allocated to one of the following team classifications: (a) homogeneous higher-level team (with all players scoring high on the GPAI), (b) homogeneous lower-level team (with all players scoring low on the GPAI) or (c) heterogeneous skill level team (with half the players scoring high and half the players scoring low on the GPAI). These teams of 4-5 players each remained constant throughout the season, and students were never informed of their ability level of ability or of the homogeneity or heterogeneity of their team's composition (the note in Table 1 further describes the distribution of these teams across the five classes.)

Data Collection

We collected quantitative data through the GPAI and questionnaire, and qualitative data through the experts' analysis on students' game performance, and students' and teachers' focus group.

Game Performance andAassessment Reliability

We made digital video records of student game performance variables at both pre-test and post-test, based on 8-minute recordings of each student's play during these observations. We then assessed game performance variables using the most common and fundamental indices of the (Table 2). To ensure the reliability of the GPAI data, three observers were trained for more than 30 hours throughout 12 meetings in which they viewed and analyzed video clips of mini-volleyball games that were not part of the current study. The observers were considered adequately reliable when they were able to achieve a 90% (intraclass correlation coefficient > 0.90) accuracy standard with respect to both their intra- and inter-observer reliability. This was calculated by means of the intraclass correlation coefficient (ICC) and percent agreement ((agreement/disagreement) * 100) achieved following a 12-minute assessment of three single players at two intervals separated by two weeks. We measured reliability of the pre-test and post-test data through inter-observer evaluation among the three coders. In these evaluations, the observers analyzed more than 15% of studio players (18 students). Reliability of the observation reached an ICC between .92 and .93, and percent agreement values between 84.16% and 91.66%.

[Insert Table 2 about here.]

Knowledge

We measured GK with a written test adapted from a validated volleyball questionnaire (see Appendix I; Mahedero et al., 2015). We administered this questionnaire before the unit began and again on completion with an average time for completion being 25 minutes. This test consisted of six questions concerning technique, five concerning rules, three concerning tactics, and a further three assessing the students' general knowledge about the game of mini-volleyball. An example of a question concerning technique was, "Could you describe the technique of overhand pass?" of one concerning rules was, "What is the mini-volley field size? of one concerning tactics was, "The forearm pass is appropriate for..." and of one concerning general knowledge was, "How many referees are needed to referee the game play?"

Experts' Analysis of Game Performance

As noted, to provide a more qualitative dimension to the GPAI data, ten experts in volleyball agreed to analyze sample video sequences of the participating students during their pre-test and post-test game play. Each of these experts had more than five years of experience in teaching, research and game analysis in volleyball. First, using the methodology outlined by Hastie et al. (2013), two video clips (initial and final games) of every team (60 clips in total) were analyzed from video playbacks with these evaluators during three sessions of game analysis. This process took, in all, approximately seven hours. Second, during these analysis sessions, the experts were together in a room where video projections were made. They were never informed about the homogeneity or heterogeneity of the teams in the video clips. Third, at the beginning of the session they were provided with an analysis form and given an explanation of the tasks. The first task was to judge, using a score sheet template, whether they saw distinctive qualitative differences in the overall quality of each team's performance across their team games (technical quality of the players and tactical decisions and GI), and if so, to note which facets were superior. Finally, the second task

for the experts was to debate their answers as a group. These responses were audio-recorded for analysis purposes alongside with the main researcher field notes.

Focus Group

At the completion of the season, the second researcher conducted interviews with 16 student teams (~80 students) as well as with both teachers. The selected teams included seven who were classified as at a higher homogeneous skill level, five teams at a lower homogeneous skill level, and four teams at heterogeneous skill levels. By consequence, each group interview consisted of five or six students from the same team. Each interview followed a specific script, which lasted approximately 10 minutes. The students were first asked to comment individually on their DM, SE, GI, GK and level of game play as a result of their participation in the season. Following this, the students were given the opportunity to debate their responses as a group. These interviews were recorded and later transcribed. Both teachers were also asked to respond to these items in one-on-one interviews.

Data Analysis

Game Performance and Knowledge

We used the IBM Statistical Package for Social Sciences (SPSS) version 20.0 to analyze the GP level (DM, SE, and GI) and GK. We first used a series of paired samples t-tests to compare pre-test and post-test scores for (a) the entire sample, (b) those in the homogeneous higher level, (c) those in the homogeneous lower level, and (d) those in heterogeneous teams. The performances at both time points were assessed for higher and lower skilled students within the heterogeneous teams. We used analyses of variance (ANOVA) to determine any significant improvement differences between the three groups, and we used Cohen's d to reflect the effect size (ES) after calculating partial eta squared values. The data were determined to be normally distributed by analysis using the Kolmogorov-Smirnov test.

Experts' Commentary of Game Performance

First, we calculated the level of agreement between reviewers with respect to the extent of improvement between the two videos (i.e., pre-test and post-test videos). Across 95% (n=57) of the videos, experts observed differences between these clips. Second, for clips indicating expert-perceived improvement, we calculated the 'percent correct' (i.e., those placing the videos in the appropriate temporal order). In other words, all the experts (n=10) placed the videos in appropriate temporal order (i.e., pre-test and post-test order). Third, the descriptive responses collected concerning game play quality were analyzed inductively with the intent to develop themes. Table 3 shows the consistency of the experts' comments to justify the game quality improvement.

[Insert Table 3 about here.]

Focus Group

Regarding students' and teachers' perceptions as determined from focus group interviews, two experienced researchers coded a total of 40,194 words using QSR NVivo 11 software. Data were analyzed deductively based on the pre-existing categories (DM, SE, GI, student skill level and GK). As described above, researchers coded the data and engaged in a reflective dialogue with an independent researcher to interpret the transcripts, followed by a critical examination of transcripts that were generated in the first analysis (Braun & Clarke, 2006). Trustworthiness of expert opinions was supported through continuous feedback and participative analysis on the part of the two researchers as they reviewed the codes and descriptors, resulting in findings that were dependable and credible (Lincoln & Guba, 1985). In addition, reviewers engaged in confirmatory member-checking for credibility. That is, all participants received a verbatim transcription of their interview to verify the correctness of the data, clarify confusing quotes, and add/modify information (some ideas were re-worded).

Results

The study's main findings are presented in two sections, reflecting (a) a general exploration of the whole sample, and (b) values grouped according to the students' skill levels. Globally, there were improvements for both GP and GK across all students.

Game Performance (GP) and Knowledge (GK): General **Findings**

Participant changes in means scores for GP (DM, SE, and GI) and GK across the two data collection points can be seen in Table 4. Overall, all students (regardless of their grouping or initial skill level) became more competent and knowledgeable, irrespective of any group difference in their pre-test scores. However, considering the small ES and despite the students' and teachers' perceptions, the improvement was minor in SE and GI. Based on these results, if the goal is for each student to improve at mini volleyball, grouping homogenously or heterogeneously on ability level would seem equally appropriate.

[Insert Table 4 about here.]

During the interviews with students conducted at the beginning of the unit (see Table 5), many reported that they were mostly concerned with simply sending the ball to the other side of the net. However, by the end of the season, students commented that they were more aware of the need for team collaboration, and of being more conscious of the value of using three passes for successful attacking plays. Teachers focused more on the technical aspects of the game, noting that many students improved their serve and overhead pass skills. The teachers also suggested that there was not only a higher student engagement during game play, but also a higher level of player movement during rallies (Table 5). To support these claims, the experts suggested that the quality of students' play was superior in later games compared to at the beginning of the season. Each of the experts noted that the players showed substantial improvements in hand and forehand passes, overhead and block skills (Table 3). In addition, students showed a tactical awareness and GI improvement was noted from the high frequency of verbal exchanges between students.

[Insert Table 5 about here.]

In terms of GK, students learned about technique, tactical, game rules and general knowledge about the mini-volleyball game. Students answered in the interviews that refereeing, score keeping and teacher support and feedback helped them to learn the rules much better, and they understood the importance of their initial court position in performing these skills (Table 5). In

addition, consistent with experts' comments (Table 5), the most common arguments given by students to justify their perceived improvements in the different categories analyzed (DM, SE, GI or GK) were assuming individual roles, the supportive verbal exchanges between higher skilled and the lower skilled students, and team affiliation.

Game Performance and Knowledge: Ability Level Groups

There were no significant differences between the ability level groups for either GP improvement or GK. Nevertheless, the results show some differential improvements between higher and lower skilled students within heterogeneous teams. Although higher skilled students within these teams showed statistically significant improvements across all components of the GPAI, lower skilled students showed improvements only in DM and GK. In addition, there were no significant differences among lower skilled students in GI in either group (Table 4).

Qualitative data from the experts' analysis were consistent with the GPAI data. In three teams (two homogeneous higher skilled teams and one heterogeneous team) experts expressed doubts about the level of improvement. They acknowledged that the players improved with respect to their positioning on the court but without a concurrent increase in the quality of individual and group skills. On the other hand, they observed a better collaboration among students, and a high level of participation and commitment (Table 3). Accordingly, lower skilled students mentioned during the interview the benefits of having the chance to learn from their higher skilled teammates. They also suggested that they were more active and enthusiastic participants within their teams compared with previous units. The teachers supported these comments and mentioned that the higher skilled students did indeed collaborate with the lower skilled students to improve the team performance (Table 5).

Discussion

In this study, we explored whether there were differences in GP variables and GK among a cohort of high school students who participated in sport education over an entire mini-volleyball unit grouped into either homogeneous or heterogeneous skill level groups. As a general summary,

316 we found that all students (regardless their skill level groupings or initial skill levels) improved in
317 GP and GK over this period. Grouping students homogeneously or heterogeneously had no
318 differential effect on these individual achievements. The key explanation for the development of
319 competence was that the structure and the features of the season allowed for significant practice
320 opportunities through teamwork, the students' roles and teacher support and feedback. As one
321 student commented: 'Overall, I do think that now all of us within the team have a more active
322 participation than we used to have, the teacher was very supportive', and the teacher highlighted:
323 'To succeed, you need to create a positive atmosphere, and we had that as a consequence of the
324 sport education general experience'.

325 These findings of improvement following a sport education experience are consistent with
326 previous research (Hastie, 1998; Mahedero et al., 2015; Pritchard et al., 2008). In fact, systematic
327 reviews of sport education research (see Araújo et al., 2014; Hastie et al., 2011) have concluded that
328 this pedagogical model can achieve the five common physical education content standards (i.e.,
329 motor skill development, tactical knowledge and performance, fitness, social development, and
330 student attitudes and values) regardless of the grouping strategy. With respect to teachers, there
331 were suggestions of observed improvements in students' GP, a finding consistent with those of
332 Hastie et al. (2009) who also made significant improvements in both the selection and execution
333 dimensions of game play in badminton.

334 Furthermore, while student skill level has been considered as a relevant variable in a number
335 of past studies (Hastie, 1998; Mahedero et al., 2015; Mesquita et al., 2005, 2012), the grouping
336 strategy used in those studies (homogeneous versus heterogeneous) did not allow an analysis of
337 skill level. The current study showed that students improved regardless the grouping strategy or
338 students' initial skill level. Students expressed in the interviews that they perceived themselves to
339 be better players at the end of the unit, both in technique and tactically. Based on individual and
340 group effort they assumed they had obtained slight improvements and achieved team success. Some
341 students emphasized 'I think I am more competent in the forearm pass' or 'Despite it is very

challenging, we tried to send the ball to a free zone, because now I can use it as the second pass'. These results are consistent with previous studies in which students also improved their GP (MacPhail et al., 2004; Siedentop, 2002).

With respect to lower skilled students, our results support those of Pereira et al. (2015) who found that lower skilled students improved in track and field tests and in their total scores at the end of the unit with sport education, which those authors attributed to higher levels of student autonomy and motivation. From our interview data, our findings are also consistent with Perlman's (2010) discovery that lower skilled students reported a greater sense of enjoyment with sport education. Nevertheless, in the present study, lower skilled students improved their GP, based on the DM but not on the SE, especially when they practiced in heterogeneous teams. In addition, the GI of lower skilled homogeneous teams and heterogeneous teams was not large.

Our results differed from previous studies that explored the effect of ability grouping and that reported that lower skilled students were at disadvantage when they participated alongside higher skilled classmates in terms of success rates, engagement rates and playing efficiency (Hastie et al., 2017; Ward et al., 2019). Reasons for inconsistencies in findings may include the complexities of teaching and learning in different contexts (Light, 2008), but also our different student grouping strategy. Our grouping strategy (e.g., homogeneous low and high, and heterogeneous low and high) persisted in all the phases of the season (not just for the graded competition). In previous studies (Hastie et al., 2017), students within each team of 6-7 players were able to divide into their three-player sub-teams, and the students played in three different leagues without interleague play (homogeneous higher skill, homogeneous lower skill and heterogeneous). This 'sustained' grouping affiliation in either homogeneous or heterogeneous teams, allowed for the refinement of gameplay and interpersonal dynamics, as confirmed by the study participants' and experts' comments, and their improved GP (Farias et al., 2018; Table 3 and 5).

Also, according to the teachers' comments, higher skilled students in the homogeneous and heterogeneous teams were given the autonomy to introduce new learning tasks through their roles as student-coaches. This facilitated and encouraged equal participation and engagement of all players, regardless their skill level and grouping strategy, which is aligned with recommendations from Harvey et al. (2014). It is necessary to ensure that teams are equitably challenged; especially, when the initial differences between groups were considerable, given the diversity of the Spanish high school system.

The quantitative results in this study were also supported by experts' qualitative assessments after they analyzed the students' initial and final video sequences. Experts mainly reported improvements in DM, SE and GI. Furthermore, as heterogeneous higher skill level teams increased the level of GI and their final success, teacher perceptions from interviews with students were that students' attitudes improved considerably throughout the unit. The experts also mentioned that they appreciated an observing an increasing number of verbal exchanges between the students, consistent with findings from Brock and Hastie (2017). Students' GI improvement could be understood as a measure of more extensive verbal exchanges, observed particularly to have been higher among skilled than lower skilled students. Nevertheless, this finding should be interpreted with care. The higher skilled students in the heterogeneous teams emphasized during their interviews that they took ownership of providing assistance (peer-assisted learning) and support for their teammates, allowing all members of their team to improve (Mahedero et al., 2015). This behaviour was also reported by Lafont et al. (2017) during a cooperative learning unit. Lafont et al. (2017) highlighted how social variables and interactions within the teams might have a collective benefit and may provide relevant modes of inclusion under certain conditions. Overall, it is important to highlight that despite our finding of student improvement independent of the grouping strategy, group dynamics and student behavior always carry an element of unpredictability (Barker & Quennerstedt, 2017).

Limitations and Directions for Further Research

These findings should be interpreted with caution due to the quasi-experimental nature of our research design and the lack of a control group. In addition, grouping was only determined by students' previous ability and not by other variables such us the students' psycho-social competencies. Psycho-social variables as for example, sportsmanship, motivational regulation, and self-confidence could also be relevant as a grouping criterion than ability (Barker & Quennerstedt 2017; Brock & Hastie 2017). Further research might be done exploring game performance related variables with different grouping strategies, but also non-game performance related variables like students' meaningful experiences (Kretchmar, 2000).

Conclusion

While sport education literature has tended to prefer heterogeneous ability-based student grouping, the present study demonstrated that heterogeneous and homogenous groups of both higher and lower skilled students were undifferentiated in their effects on student achievement and improvements in GP and GK after a 12-week mini-volleyball unit. While, this improvement was less significant for lower skilled students, ability type grouping was not a major influence on improvements in GP and GK through the sport education experience. Given these unpredictable findings among lower skilled students, we suggest that teachers consider a variety of grouping strategies different than skill ability, and also listen to students' voices to inform those decisions.

It is necessary for grouping strategies to go beyond mixing people up, keeping friends together, or even ensuring groups have at least one expert. This is not to suggest that one grouping strategy might be superior to others in terms of facilitating learning. Rather, teachers should be mindful of the relationships that group composition might have on teaching and learning and use different strategies according to the uniqueness of their local context with the hope to optimize group dynamics, hence learning and engagement.

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

References

Araújo R., Mesquita, I., & Hastie, P. A. (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.

Barker, D., & Quennerstedt, M. (2017). Power and group work in physical education: A foucauldian perspective. *European Physical Education Review, 23*, 339-353. https://doi.org/10.1177/1356336X15620716

Barker, D., Quennerstedt, M., & Annerstedt, C. (2015). Learning through group work in physical education: A symbolic interactionist approach. *Sport, Education and Society, 20*, 604-623. https://doi.org/10.1080/13573322.2014.962493

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77-101. https://doi.org/10.1191/1478088706qp063oa

Brock, S. J., & Hastie, P. A. (2017). Students' verbal exchanges and dynamics during sport education. *European Physical Education Review, 23*, 354-365. https://doi.org/10.1177/1356336x16629582

Dyson, B., Griffin, L., & Hastie, P. A. (2004). Theoretical and pedagogical considerations for implementing sport education, tactical games, and cooperative learning instructional models. *Quest, 56*, 225-239.

Farias, C., Mesquita, I., & Hastie, P. A. (2019). Student game-play performance in invasion games following three consecutive hybrid Sport Education seasons. European Physical Education Review, 25(3), 691-712. https://doi.org/10.1177/1356336X18769220

Harvey, S., Kirk, D., & O'Donovan, T. M. (2014). Sport education as a pedagogical application for ethical development in physical education and youth sport. *Sport, Education and Society, 19*, 41-62. https://doi.org/10.1080/13573322.2011.624594

Hastie, P. A. (1998). Skill and tactical development during a sport education season. *Research Quarterly for Exercise and Sport, 69*, 368-379. https://doi.org/10.1080/02701367.1998.10607711

Hastie, P. A., & 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. https://doi.org/10.1123/jtpe.2013-0141

Hastie, P. A., Layne, T. E., & Mesquita, I. (2013). Evaluating game performance from external experts in sport education. *Revista Portuguesa de Ciências do Desporto, 13*, 23-37.

Hastie, P.A., Martínez de Ojeda, D., & Calderón, A. (2011). A review of research on sport education: 2004 to the present. *Physical Education and Sport Pedagogy*, *16*, 103-132. https://doi.org/10.1080/17408989.2010.535202

Hastie, P. A., Sinelnikov, O. A., & Guarino, A. J. (2009). The development of skill and tactical competencies during a season of badminton. *European Journal of Sport Science, 9*, 133-140. https://doi.org/10.1080/17461390802542564

Hastie, P. A., Ward, J. K., & Brock, S. J. (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*, 316-327. https://doi.org/10.1080/17408989.2016.1203888

Haynes, J., Fletcher, T., & Miller, J. (2008). Does grouping by perceived ability sustain student attitude towards physical education? Paper presented at the *Australian Association for Research in Education (AARE)*, Brisbane. Retrieved from http://www.aare.edu.au/data/publications/2008/hay08226.pdf

Kretchmar, R. S. (2000). Movement subcultures: Sites for meaning. *Journal of Physical Education, Recreation and Dance, 71*, 19-25. https://doi.org/10.1080/07303084.2000.10605140

Lafont, L., Rivière, C., Darnis, F., & Legrain, P. (2017). How to structure group work? Conditions of efficacy and methodological considerations in physical education. *European Physical Education Review, 23*, 327-338. https://doi.org/10.1177/1356336x15626639

Lentillon-Kaestner, V., & Patelli, G. (2016). Effects of grouping forms, student gender and ability level on the pleasure experienced in physical education. *Journal of Teaching in Physical Education, 35*, 251-262. https://doi.org/10.1123/jtpe.2014-0216

Light, R. (2008). Complex learning theory - its epistemology and its assumptions about learning: Implications for Physical Education. Journal of Teaching in Physical Education, 27, 21-37. https://doi.org/10.1123/jtpe.27.1.21

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

MacPhail, A., Kirk, D., & Kinchin, G. (2004). Sport education: Promoting team affiliation through physical education. *Journal of Teaching in Physical Education, 23*, 106-122. https://doi.org/10.1123/jtpe.23.2.106

Mahedero, M. P., Calderón, A., Arias-Estero, J. L., Hastie, P- A., & Guarino, A. J. (2015). Effects of student skill level on knowledge, decision making, skill execution and game performance in a mini-volleyball sport education season. *Journal of Teaching in Physical Education, 34*, 626-641. https://doi.org/10.1123/jtpe.2014-0061

Mesquita, I., Farias, C., & Hastie, P. (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. https://doi.org/10.1177/1356336x12440027

Mesquita, I., Graça, A., Gomes, R., & Cruz, C. (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.

Morse, J. M., & Niehaus, L. (2009). *Mixed method design. Principles and procedures.* Routledge.

Oslin, J., Mitchell, S., & Griffin, L. (1998). The game performance assessment instrument (GPAI): Development and preliminary validation. Journal of Teaching in Physical Education, 17, 231-243. https://doi.org/10.1123/jtpe.17.2.231

Pereira, J., Hastie, P. A., Araújo, R., Farias, C., Rolim, R., & Mesquita, I. (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.

Perlman, D. J. (2010). Change in affect and needs satisfaction for amotivated students within the sport education model. *Journal of Teaching in Physical Education, 29*, 433-445. https://doi.org/10.1123/jtpe.29.4.433

Pritchard, T., Hawkins, A., Wiegand, R., & Metzler, J. N. (2008). Effects of two instructional approaches on skill development, knowledge, and game performance. *Measurement in Physical Education and Exercise Science, 12*, 219-236. https://doi.org/10.1080/10913670802349774

Siedentop, D. (2002). Sport education: A retrospective. *Journal of Teaching in Physical Education*, *21*, 409-418. https://doi.org/10.1123/jtpe.21.4.409

Siedentop, D., Hastie, P. A., & van der Mars, H. 2011. *Complete guide to sport education.* Human Kinetics.

Silverman, S., Tyson, L., & Krampitz, J. (1993). Teacher feedback and achievement in physical education: mediating effects of initial skill level and sex. *Journal of Human Movement Studies, 24*, 97-118.

Silverman, S., Woods, A. M., & Subramaniam, P. R. (1998). Task structures, individual student feedback, and student skill level in physical education. *Research Quarterly for Exercise and Sport, 69*, 420-424. https://doi.org/10.1080/02701367.1998.10607718

Sinelnikov, O. A. (2009). Sport education for teachers: Professional development when introducing a novel curriculum model. *European Physical Education Review, 15*, 91-114. https://doi.org/10.1177/1356336x09105213

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research, 60*, 471-499. https://doi.org/10.2307/1170761

Slavin, R. E., & Karweit, N. L. (1985). Effects of whole-class, ability grouped, and individualized instruction on mathematics achievement. *American Educational Research Journal, 22*, 351-367. https://doi.org/10.3102/00028312022003351

Tieso, C. L. (2003). Ability Grouping is not just tracking anymore. *Roeper Review, 26*, 29-36. https://doi.org/10.1080/02783190309554236

Ward, J. K., Hastie, P. A., & Strunk, K. (2019). Effects of ability grouping on students' game performance and physical activity. *Journal of Teaching in Physical Education, 38*, 187-198. https://doi.org/10.1123/jtpe.2018-0140

Wilkinson, S., Penney, D., & Allin, L. (2016). Setting and within-class ability grouping: A survey of practices in physical education. *European Physical Education Review, 22*, 336-354. https://doi.org/10.1177/1356336X15610784

Table 1. Program of Activities for the Mini-volleyball Season.

Lesson		Content	Teacher's role	Students' roles
1	Teacher directed: Skill development phase	Knowledge test Introduction to teaching approach Description of daily roles	Class leader	Participant
2		Assigned team / Pre-test 4 vs. 4	Class leader	Participant
3		Overall skills and tactics of mini-volleyball	Class leader	Participant
4	Pre-season: students to work in their teams within peer-teaching. Scrimmages	Overhead pass 1 vs. 1; 2 vs. 2	Head coach Referee advisor	Coaches, players, learn duty role, practise duty roles
5		Forearm pass 1 vs. 1	Head coach Referee advisor	Coaches, players, learn duty role, practise duty roles
6		Forearm pass 2 vs. 2	Head coach Referee advisor	Coaches, players, learn duty role, practise duty roles
7		Underhand serve 3 vs. 3	Head coach Referee advisor	Coaches, players, learn duty role, practise duty roles
8		Block 3 vs. 3, 4 vs. 4	Head coach Referee advisor	Coaches, players, duty team roles
9	Formal competition: Round robin tournament	Tournament 4 vs. 4	Programme manager	Duty team roles
10		Tournament 4 vs. 4	Programme manager	Duty team roles
11		Tournament Post-test 4 vs. 4	Programme manager	Duty team roles
12	Culminating event	Final Competition Festivity Award ceremony	Master of ceremonies	Participant

Note. Team composition by classes: Class 1 teams = 3 high, 2 low, 1 mixed (3 high, 2 low); Class 2 teams = 3 high, 1 low, 2 mixed (a: 3 high, 2 low; b: 4 high, 1 low); Class 3 teams = 2 high, 3 low, 1 mixed (3 high, 1 low); Class 4 teams = 1 high, 1 low, 4 mixed (a: 2 high, 1 low; b: 3 high, 2 low; c: 3 high, 1 low; d: 4 high, 1 low); Class 5 teams = 2 high, 2 low, 2 mixed (a: 3 high, 2 low; b: 4 high, 1 low).

Table 2. Indexes of GPAI Assessed.

Decision making	Appropriate decision making / inappropriate decision making	An appropriate DM was defined as those actions in which the player (i) passed the ball to their teammate (on the first and second contact); (ii) sent the ball across the net in a way that placed a stress on the opposing team (on the third contact); or (iii) blocked an opponent's attack.
Skill execution	Correct technical execution / incorrect technical execution	SE was considered correct in the following cases: (i) when the player made a overhead pass or forearm pass which raised above the net being able to be controlled or attacked by a teammate; (ii) the player successfully sent the ball into the opponent's court; or (iii), the player blocked the ball so that it was returned to the opponent's side of the court.
Game performance	(Decision making + Skill execution) / 2	
Game involvement	Appropriate decision making + inappropriate decision making + appropriate technical execution + inappropriate technical execution	

Table 3. Sample Experts' Comments on Game Performance Variables during Data Collection.

Pre-test Comments	Post-test Comments
Decision making	*Decision making*
E1: : 'They do not send the ball to a free-zone' E2: 'It is difficult to see the a consistent decision making' E3: 'They play more individually than as a team'	E2: 'There are a better collaboration between students and they have a better position in the practice area' E3: 'Most of then try to make sequences of three passes in each attack' E2: 'Some of them try winning points attacking the free-zone'
Skill execution	*Skill execution*
E1: 'The ball in overhead-pass is not higher enough to let for a second (or third) pass' E2: 'There is no blocking at all during the game play, and most of them are block-out' E3: 'Usually, serves go successfully into the other side, however hand and forearm pass don't'	E2: 'Technical improvement are more visible on the hand and forearm passes' E3: 'In doing the overhead pass the students move under the ball' E1: 'Sometimes the blocking appears and the play ends with winner points'
Game involvement	*Game involvement*
E1: 'Most of the students are not actively engaged in the game' E1: 'There are two students that seems to be very out of the play' E2: 'Overall, there is not a great support between the teammates'	E2: 'There are more verbal exchanges between the students in their teams' E2: 'On the whole, I see a higher involvement of all the teams. That improvement does not seem quite evidence in two or three teams! E3: 'Almost all of the students try to participate in the game; they show a great commitment'

Note. E1 = Comment made by one and four experts; E2 = Comment made by five and eight experts; E3 = Comment made by nine and ten experts.

Table 4. Pre- to Post-test Changes in Game Performance Assessment Variables.

		Pre-test	Post-test			
Grouping	Variable	*M (SD)*	*M (SD)*	*t*	*p*	*ES*
Entire sample	GP	2.64 (2.20)	4.58 (2.13)	-7.14	< .01	.84
	DM	1.80 (2.10)	3.88 (3.36)	-8.44	< .01	.88
	SE	3.49 (3.32)	5.18 (4.11)	-4.33	< .01	.53
	GI	12.30 (7.41)	16.47 (9.89)	-5.03	< .01	.56
	K	5.44 (1.87)	11.25 (2.33)	-25.72	.01	.19
Homogeneous-high	GP	3.41 (2.70)	5.54 (3.05)	-4.35	< .01	.86
	DM	2.54 (2.82)	5.20 (3.50)	-6.05	< .01	.97
	SE	4.27 (3.52)	5.88 (3.68)	-2.35	.02	.51
	GI	12.52 (7.77)	18.43 (8.60)	-4.60	< .01	.83
	K	5.39 (1.99)	11.65 (2.31)	-17.64	.01	.36
Homogeneous-low	GP	1.55 (0.99)	3.22 (2.90)	-3.89	< .01	0.99
	DM	1.23 (1.04)	2.54 (2.89)	-2.82	< .01	0.76
	SE	1.87 (1.80)	3.90 (3.38)	-3.94	< .01	0.91
	GI	9.71 (2.20)	11.60 (2.61)	-1.43	.16	0.95
	K	5.08 (2.00)	10.94 (2.61)	-11.98	.01	.93
Heterogeneous-high	GP	2.64 (1.75)	4.90 (3.21)	-3.63	< .01	1.05
	DM	1.71 (1.81)	3.88 (.320)	-4.49	< .01	1.00
	SE	3.57 (.311)	5.82 (4.99)	-2.28	.03	0.64
	GI	13.31 (7.80)	19.10 (12.45)	-3.00	< .01	0.66
	K	5.58 (1.63)	11.34 (2.05)	-15.70	.01	.60
Heterogeneous-low	GP	2.86 (2.43)	4.22 (2.79)	-1.97	.06	.60
	DM	1.08 (1.20)	3.06 (2.98)	-3.43	< .01	1.10
	SE	4.65 (4.49)	4.86 (4.65)	-.19	.85	.05
	GI	15.50 (7.28)	16.75 (11.00)	-.46	.65	.16
	K	6.12 (1.50)	10.62 (2.25)	-6.59	.01	.77

Note. GP = game performance; DM = decision making; SE = skill execution; GI = game involvement; K = Knowledge.

Table 5. Students' and Teachers' Sample Focus Group Responses Coded by Established Themes.

Students' Interview	Teachers' Interview
Decision making	
• 'At the beginning of the unit, I was just worried about sending the ball to the other side. Now, I am more aware about team collaboration' • 'We started practicing the second pass (overhead pass) to send the ball to the other side, and in the final lessons we were more conscious of the three passes' • 'Despite it is very challenging, we tried to send the ball to a free-zone'	• 'I can see that some of them try to attack the free-zone' • 'Students improved the spatial awareness, trying to keep and defence their playing area' • 'There were longer plays because now they tried a three passes sequence' • 'There was not only a higher engagement of the students during the plays but also better player movement during rallies in response to the demands of the game'
Skill execution	
• 'My overhead pass is now more effective because I improved my initial position (feet, hands, elbows)'. • 'I think I am more competent in the forearm pass because now I can use it as the second pass' • 'Overall, most of us are able to send the ball to the other side'	• 'The students improved mainly the serve and the overhead pass because the amount of practice and the autonomy to design and practice some skills according their perceived weaknesses' • 'Basically, there is a better flow in the game because technically, most of the students are now better, and especially because of the higher ball trajectory in the passes'
Game involvement	
• 'We work better now as a team, there are more supportive verbal exchanges' • 'The individual roles made you feel more important within the team and push you to work harder every lesson' • 'Overall, I do think that now all of us within the team have a more active participation than we used to have, the teacher was very supportive'	• 'On the whole, most of students were very positive toward practice; and I do think it was because of some of the key features of Sport Education, but mainly the final competition' • 'To assume different individuals and team roles made the students who usually did not have an active participation to be more engaged'
Student skill level	
• 'What I liked more is the chance to learn from my colleagues within the team, especially from the more skilful teammates' • 'I spent loads of time supporting the lower skilled students in my team, and I think it made me feel useful and it kind of made me improve' • 'I liked to belong to a team in which there was teammates with different skill level'	• 'To be honest, I do think that most of the teams improved in the social and physical domains regardless the grouping; it was hard to see any difference' • 'I have been always using mixed ability grouping, and it was not always successful. To succeed, you need to create a positive atmosphere and we had that as a consequence of the Sport Education general experience'.
Knowledge	
• 'Since we had to act as referees and score-keepers, we learned the rules much better' • 'Now, I understand a little bit better how important is the initial position to perform the skills' • 'It is easier for us to cover the whole team side; the written piece (questionnaire) helped us in doing so'	• 'Overall, knowledge improved in most of the students thanks to the refereeing and score keeping actions that the students assumed throughout the teaching unit' • 'Regardless the type of grouping, students now have a better understanding and awareness of different game play situations'