The Physical Educator Vol. 76 • pp. 568–587 • 2019
https://doi.org/10.18666/TPE-2019-V76-I2-8667

SPORT OFFICIATING

Referee Engagement and Officiating Accuracy in a College-Level Volleyball Sport Education Season

Zachary Wahl-Alexander, Oleg A. Sinelnikov, K. Andrew R. Richards

Abstract

This study investigated university students' levels of engagement and accuracy of officiating volleyball during their first Sport Education season. The participants in this study included 20 students (12 male, 8 female) enrolled in a beginner volleyball activity course. The volleyball Sport Education season included 21 lessons that lasted 75 min each. All season gameplay was video recorded and data were collected so that we could determine student engagement and officiating accuracy for every referee throughout the season. A series of paired-samples t tests examined changes over time, and bivariate correlations examined the relationships among engagement and officiating decision variables at preseason and during the formal competition. The results of this study indicated a significant increase in active engagement of officials from preseason to the formal competition phase with the corresponding decrease in passive engagement. Furthermore, while there were no significant differences in the accuracy of some officiating decisions

Zachary Wahl-Alexander is an assistant professor, Department of Kinesiology and Physical Education, Northern Illinois University. Oleg A. Sinelnikov is an associate professor, Department of Kinesiology, The University of Alabama. K. Andrew R. Richards is an assistant professor, Department of Kinesiology and Community Health, The University of Illinois at Urbana-Champaign. Please send author correspondence to zachwahl18@gmail.com

over time, the overall officiating accuracy significantly increased from preseason to formal competition. The levels of engagement were related to the complexity of officiating calls and to an overall accuracy. Officiating competency needs to be intentionally developed during Sport Education, and teachers should allocate ample time for students to practice and improve their officiating abilities, because officiating serves as an important integral part of this pedagogical model.

In the majority of sporting contexts, officials or referees are tasked with enforcing rules, while attempting to be dependable and accurate, as to not become a significant contributing factor to the outcome of the game. In today's sporting culture, all officials, regardless of experience, are expected to perform their obligations error free. While lofty, this appears an unobtainable goal, as examinations of Major League Baseball (MLB) umpires' accuracy (88% accurate on ball/strike calls) and England's best rugby officials (6/11 tackles adjudicated incorrectly) posit that even for professional referees, calling the "perfect game" is improbable (Downes, 2016; Mascarenhas, Collins, & Mortimer, 2005). Downes (2016) also indicated that when assessing MLB pitches thrown within 2 in. of all corners of the plate, umpires got the call incorrect 31.7% of the time. Although most referees are not professional, results suggest amateur officials are held to a similar standard in regard to officiating accuracy (Burke, Joyner, Pim, & Czech, 2000).

Over the past 30 years, the levels of stress and burnout among intramural, interscholastic, and other certified youth officials have increased. Several early studies revealed low levels of stress and negligible reports of burnout (Goldsmith & Williams, 1992; Rainey, 1995; Rainey & Hardy, 1999); however, more recently, officials have reported higher levels of anxiety due to fear of making incorrect calls, being out of position, and receiving verbal abuse from players or coaches (Dorsch & Paskevich, 2007; Voight, 2009). These findings are due in part to the unrealistic expectations placed upon all officials (Kellett & Warner, 2011).

More recently, scholars identified a substantial decrease in the amount of qualified sporting referees (Ridinger, 2015). The American Sport Education (2011) program went so far to call it a national crisis in U.S sport, with the survival of sport dependent upon approaches of

successfully training and retaining referees while motivating youth to become officials. Although referees need experience and practice officiating (Catteeuw, Helsen, Gilis, & Wagemans, 2009), there are few opportunities for teaching these skills.

A potential benefit of one of the pedagogical models, namely, Sport Education, being implemented in contemporary physical education is its stated objective of developing "literate sportspeople," which includes teaching students how to appropriately officiate, judge, and umpire sports and activities within the context of regular physical education (Siedentop, Hastie, & van der Mars, 2011). Sport Education's primary objective is to create an authentic sporting experience (Siedentop, 1994) while ensuring students become competent, literate, and enthusiastic sportspeople. Each "season" within Sport Education is longer in duration than typical physical education units, with students competing on fixed teams throughout the season. The key features of the Sport Education model include team affiliation, seasons, formal competition, record keeping, festivity, and a culminating event (Siedentop et al., 2011). These six key tenets, along with small-sided developmentally appropriate gameplay, foster higher levels of learning (Browne, Carlson, & Hastie, 2004; Hastie, Sinelnikov, & Guarino, 2009) and provide participants with an authentic sporting experience unique to other physical education classes.

One distinctive characteristic of this pedagogical model is the nonnegotiable requirement for students to fulfill various roles outside that of a team player. For example, students might be tasked with coaching a team, leading a daily warm-up, collecting statistics, writing regular newspaper reports, or scouting future opponents (Siedentop et al., 2011), in addition to their participation in gameplay. These additional roles increase student excitement (Kinchin, Wardle, Roderick, & Sprosen, 2004), aid in building team affiliation (MacPhail & Kinchin, 2004), and help students grow in their ability to conduct these responsibilities as a season progresses (Kim, Penney, Cho, & Choi, 2006).

In addition to being a player, and fulfilling one of the aforementioned team roles, each student acts as an official at various stages throughout every season. A primary purpose of this duty-team role is for students to take active part in managing their season and to

become more literate sportspeople. Additionally, as a result of being an engaged observer such as an official in game, students tend to not only develop deeper cognitive understanding of the rules of the game, but also become more tactically astute as a player (Sinelnikov & Hastie, 2010). However, despite officiating being a required duty-team role in any Sport Education season (Siedentop et al., 2011), many impediments need to be overcome for students to be effective officials. Specifically, at a minimum, students must have a deep understanding of gameplay and a significant understanding of the rules to referee a sporting contest successfully (Layne & Hastie, 2014).

Although the Sport Education model has been comprehensively examined over the past 25 years, there has been limited contradicting evidence on the role of student officials within a season. Initially, there were concerns that students would not pay attention to the game or activity if they were not playing or that they would not be able to officiate accurately. Hastie (1996) was the first to design a study evaluating decision-making accuracy and involvement of referees within a Sport Education season. The findings from Hastie's study demonstrated student officials' ability to stay fully attentive to gameplay (96% of the time) and a gradual increase in their ability to make accurate decisions. Although student officials were not initially accurate without teacher interference, refereeing success rate reached higher levels (94%) at the conclusion of the season. While these findings appear promising, it should be noted that officiating calls in that study occurred every 26 s and were simple in nature. Since that time, other studies confirmed the ability of students to stay actively involved when officiating during Sport Education. For example, in a study of Russian students' participation in a Sport Education season, referees were actively involved 90.46% of the time (Hastie & Sinelnikov, 2006).

Following this line of research, Layne and Hastie (2014) examined officiating attentiveness and accuracy of calls made by first-time student officials in the fourth grade. The results of this study indicated that officials were more engaged and accurate with their officiating decisions as the season progressed, with students making successful calls (84% of the time) in the postseason.

In addition to using systematic observation methods to evaluate officiating performance, several studies focused specifically on student perception and memories to better understand the student perspective. Students from several iterations displayed a strong affinity toward officiating, characterizing it as fun, enjoyable, and important (Hastie, 1996; Hastie & Sinelnikov, 2006); however, other students prioritized winning and deemphasized the importance of peer officiating (Mowling, Brock, & Hastie, 2006). Sinelnikov and Hastie (2010) examined students' autobiographical memories relative to their involvement in multiple Sport Education seasons and reported that opportunities to officiate were very memorable, challenging, and rewarding to students, and as such, the act of officiating resulted in a deeper understanding of rules and strategies that lasted well beyond the Sport Education season.

Performing the obligations of a referee at any level is difficult and stressful. With the lack of officials in sport today, it appears prudent for youth to be afforded additional opportunities to gain refereeing experience. While the literature suggests that most students enjoy the role of officiating within the Sport Education model, limited evidence has depicted the accuracy and quality of officiating calls, with most research conducted with younger students (upper elementary and middle school). While Sport Education has been used as a pedagogical model in college-level activity courses (e.g., Bennett & Hastie, 1997; Hastie & Sinelnikov, 2007), to date no studies have investigated officiating in Sport Education with university students. Furthermore, researchers have raised concerns over the quality of student officiating competency and accuracy during initial seasons (Wahl-Alexander, Sinelnikov, & Curtner-Smith, 2017). Therefore, this study investigated university students' levels of engagement and accuracy of officiating volleyball during their first Sport Education season. Specifically, these research questions guided data collection and analysis: (1) What was the level of student engagement while they officiated? (2) How accurate were the students when making officiating decisions differentiated by specific rules? (3) How, if at all, did engagement and success rate change throughout the season?

Method

Participants

The participants in this study were 20 students (12 male, 8 female) enrolled in one beginner volleyball activity course. Three students were first-year students (freshmen), six were in their second year of study (sophomores), five were completing their third year of study (juniors), and six were enrolled in their fourth or fifth year (seniors) at the university. The average age of participants was 19.2. Seventeen participants had little to no experience playing volleyball, while three (2 females, 1 male) had played in high school. Additionally, none of the participants in this study had experience with the Sport Education model. Informed consent was obtained for all participants in this study prior to data collection, and the first author's university review board for research on human subjects approved the research protocol.

The instructor of the course was a 25-year-old female Sport Pedagogy graduate student who had ample experience with the volleyball content area. In addition to playing for over 6 years at a club level, she was a certified volleyball official who officiated junior high and high school matches. The instructor also had experience using the Sport Education model, having taught over 15 Sport Education seasons across multiple age groups (middle school, high school, and college) and across multiple content areas (volleyball, soccer, physical conditioning, pickle ball, and others).

Setting

The study took place in a research university in the Midwest of the United States. Over 20,000 students were enrolled in this public institution at the time of the study with 58.1% represented as Caucasian, 16.2% as African American, 14.4% as Hispanic, 4.9% as Asian, and 6.4% as other. The university offered a variety of basic activity courses for which students were allowed to sign up based upon their specific interests. Each volleyball class met in a large gymnasium on campus. The gymnasium had a large area with two regular-sized volleyball courts and sufficient space for gameplay and team practices.

Volleyball Sport Education Season Plan

The volleyball Sport Education season consisted of 21 lessons, with classes meeting twice each week. Each lesson lasted 75 min. The instructor designed the season according to Siedentop's (1994) key tenets of Sport Education: seasons, formal competition, record keeping, team affiliation, festivity, and a culminating event. The first two classes consisted of a series of skill assessments for evaluating the students' ability levels. Team selection process included preseason allocation of students, with consideration of skill assessment data for creation of five equally skilled and mixed-gender teams (Siedentop et al., 2011). Within each team, students decided who would assume the role of coach, statistician, warm-up leader, and cheer captain throughout the season. In addition to these roles, at various points of the season all participants performed the duty-team roles of official and statistician. This season followed the typical structure of Sport Education: an initial skill development and team practice phase, followed by inconsequential preseason games and consequential formal competition regular season games, and concluding with a festive culminating event.

The skill development phase included 2 days of teacher-led skill development and skill assessment activities, followed by 8 days of a team training camp during which teams were selected and students chose team roles. Consistent with Sport Education's tenet of gradually shifting responsibility for learning from teacher to learner (Siedentop et al., 2011), the following five lessons were mostly led by students and entailed a team warm-up and team skill practices for improving skills, tactics, and game content knowledge.

The preseason phase lasted three lessons during which inconsequential games between teams took place. A typical lesson in this phase began with each team's warm-up leader initiating a warm-up, followed by a 10- to 15-min student-initiated (coach-led) team practice focused on various skills. This was followed by inconsequential games managed by the duty team tasked with officiating, scorekeeping, and organizing responsibilities.

The formal competition phase lasted eight lessons and included a round-robin tournament and playoffs during which each team had an opportunity to play against other teams and to perform the roles of a duty team. Statistics and records were kept during the season, and

an awards ceremony culminated the last day of the Sport Education experience.

Officiating Training During Sport Education

Officiating a volleyball match was specifically taught by the instructor. While the instructor taught some of the rules of the game during training camp, explicit instructions on officiating a match took place during the preseason phase of the volleyball Sport Education season. While there are 25 referee's official hand signals (Federation Internationale De Volleyball, 2016), for the purposes of this study we have classified officiating signals into the following categories: (1) service faults (the server touches the court); (2) inbounds and out-of-bounds calls (ball "in"; ball "out"; ball hits "ceiling"); (3) net calls (net touched by player or player crossing under the net); and (4) illegal hits (catch; four hits).

To teach students how to officiate, the instructor initially explained the basic rules of officiating, taught officiating signals, and provided demonstrations during a sample game. Second, during each lesson the instructor allocated time for each team to review rules and officiating signals. This review was facilitated by each team's student statistician, who provided leadership during officiating. Third, students viewed video segments that described officiating calls and hand signals, and video clips distinguishing the differences between calls. Finally, all students practiced officiating during nonconsequential preseason games. During these games, the instructor focused on teaching officiating positioning and accuracy of signals. During these practice games, gameplay pace was slowed and the instructor provided feedback and helped ensure that all players received ample time in the role of the official. All officials had a whistle, which they would blow to stop play and then would indicate the appropriate call with hand signals.

Data Collection

Lessons during the preseason and formal competition phases were video recorded with two GoPro Hero 4 cameras. Each GoPro camera was situated in a corner of the gym on a tripod stand as to not interfere with the lesson. Following each lesson, the footage was transferred to a desktop computer, which provided a clearer view of the recording from the lessons. We utilized event recording for

two variables (active and passive engagement) in this study to determine the official's involvement during each rally. Table 1 details the descriptions for each student involvement category, and following previous studies protocol (Hastie, 1996; Hastie & Sinelnikov, 2006), this study followed the subsequent sequence:

- At the beginning of every play, one official was selected for observation.
- That official was observed for the entirety of that play.
- The researcher would make a determination based upon the already established definitions if that official was actively or passively involved, distracted, or off-task.

This process was repeated with every official for every rally.

Table 1
Active and Passive Involvement for Officials (Hastie, 1996)

Category	Definition
Actively involved	Keeps up with the ball, follows play, consistently enforces the rules, uses whistle definitively.
Passively involved	Watching play, but not moving to keep up with the ball; makes occasional rulings or uses whistle passively.
Distracted	In the field of play, but attending to outside factors (i.e., the other match); misses a call due to inattention.
Off-Task	Not watching or following play; not making rule decisions; engaged in activity detrimental to officiating performance.

To determine each official's accuracy during every point, the observer focused on one official during each rally. Once the rally concluded, the results were documented, and the observer re-watched the rally to observe the other official. To determine the accuracy, frequency data were collected for each rally, and the observer identified if the official made a correct or incorrect decision for each of the following:

- Was the server's foot behind the service line during the serve? (Service fault)

- Did any player on either team contact any part of the net? (Net fault)
- Did each team contact the ball more than a maximum of three times before returning the ball? (Illegal hit – four-hit rule)
- Did a player have prolonged contact with the ball (lift)? (Illegal hit – catch)
- Was the ball in or out when making contact with the floor? (In/out rule)
- Did the ball make contact with the ceiling at any point during the rally? (Illegal hit – ceiling rule)

During each rally, the official being observed was coded as either correct or incorrect for every opportunity for making a call. For example, if the referee succeeded in calling a server's foot behind the line and the serve out, the observer would tally one successful call for service fault and in/out. Further, if during a rally, an official made an incorrect call followed by several correct calls, a tally would be documented for the incorrect and correct calls. If the official failed to make a call, an incorrect tally was recorded. During each game, one official was located on the side of the court close to the net, whereas another official was on the opposite side in the corner. The video camera was positioned to record actions of the referees, along with the corresponding gameplay.

Interobserver Reliability

Interobserver agreement was determined via the formula [Agreements/(Agreements + Disagreements)] × 100 (van der Mars, 1989). Fifty percent of games were randomly analyzed by a trained second observer, which is more than the 20% threshold recommended by Cooper, Heron, and Heward (2014). The interobserver reliability was 98.5%, which exceeds the recommended standard for systematic observations (van der Mars, 1989).

Data Analysis

The data analysis process began with standard procedures for data cleaning and screening recommended in the literature (Tabachnick & Fidell, 2013). Through this process, we determined that the data met the basic assumptions of inferential statistical analysis and that

we could proceed. However, no instances of off-task behavior were recorded, and referees had few opportunities to make calls using the ceiling rule. As a result, both of these variables were omitted from further analysis. Next, all data related to engagement was converted into percentile scores using Equation 1. This provided the percentage of time in each game that an official was actively engaged, passively engaged, and distracted.

Equation 1

$$\text{Percent}_x = \frac{X}{(\text{Active Engaged} + \text{Passive Engaged} + \text{Distracted})}$$, where X is type of engagement.

Similarly, the accuracy of officiating decisions (i.e., in/out calls, serve contact, net balls, four-contact rule, lifts) were transformed into percentage scores using Equation 2. This provided the percentage of calls that were either accurate or inaccurate for each rule. An additional variable that represented overall officiating accuracy was created by dividing the total number correct officiating decisions by the total number of decisions made.

Equation 2

$$\text{Percent}_x = \frac{\text{Correct Calls}_x}{(\text{Correct Calls}_x + \text{Incorrect Calls}_x)}$$, where X is the specific type of call.

Once the data had been converted to percentile score, a series of paired-samples *t* tests examined changes over time (i.e., preseason to formal competition). One test was run for each study variable, which resulted in five tests. Cohen's *d* was calculated as a measure of effect size for all *t* tests. A Cohen's *d* between .20 and .50 is associated with a small effect, between .50 and .80 with a medium effect, and greater than .80 with a large effect (Cohen, 1992). All statistical analyses were conducted using SPSS 23.0.

Next, bivariate correlations examined the relationships among engagement and officiating decision variables at preseason and during the formal competition. Effect sizes for the correlational analyses were interpreted using Cohen's (1988) guidelines. A correlation coefficient ≥ |.10| is associated with a small effect, ≥ |.30| with a medium effect, and ≥ |.50| with a large effect.

Results

Table 2 overviews descriptive statistics and the results of the paired-samples *t* tests for all engaged and officiating decision data. Relative to engagement, the percentage of time that referees were actively engaged increased significantly from preseason to formal competition, $t(19) = 3.09$, $p = .006$, $d = 1.01$, and passive engagement decreased significantly, $t(19) = -3.35$, $p = .003$, $d = 1.09$. While the percentage of time referees were distracted decreased slightly from preseason to formal competition, the change was not significant.

In regard to the officiating decisions, the percentage of correct calls related to the line serve rule increased significantly from preseason to formal competition, $t(19) = 2.73$, $p = .013$, $d = .89$. The accuracy of calls related to netballs, $t(19) = 3.20$, $p = .005$, $d = 1.40$, and lifts, $t(19) = 4.63$, $p < .001$, $d = 1.50$, also increased over time. There were no significant differences in the accuracy of in/out or four-hit-rule calls over time, but overall officiating accuracy significantly increased from preseason to formal competition, $t(19) = 4.42$, $p < .001$, $d = 1.47$.

Following the paired-samples *t* tests, bivariate correlations examined the association between engagment and officiating decision variables at preseason and during formal competition. In the preseason analyses, there was no significant association between any of the engagement and officiating decision variables (see Table 3). However, during the formal competition, there were several significant correlations (see Table 4). Specifically, active involvement was positively associated with the accuracy of calls related to net balls, $r(20) = .58$, $p = .008$; lifts, $r(20) = .68$, $p = .001$; and overall accuracy, $r(20) = .60$, $p = .005$. Passive engagement was negatively related to the accuracy of netballs, $r(20) = -.48$, $p = .032$; lifts, $r(20) = -.72$, $p < .001$; and overall accuracy, $r(20) = -.54$, $p = .013$. Finally, being distracted was negatively related to in/out calls, $r(20) = -.47$, $p = .038$; net balls, $r(20) = -.60$, $p = .005$; and overall accuracy, $r(20) = -.53$, $p = .016$.

Table 2

Descriptive Statistics and Independent t-Test Results

Enagement and officiating decision	Pre		Post		Paired-samples *t* test		
	M	*SD*	*M*	*SD*	*t*	*p*	Cohen's *d*
Engagement							
Actively involved**	.75	.13	.85	.09	3.10	.006	1.01
Passively involved**	.20	.10	.11	.07	-3.35	.003	1.09
Distracted^{NS}	.05	.04	.04	.03	-1.26	.224	.41
Officiating Decisions							
In/out rule^{NS}	.98	.02	.99	.01	1.26	.224	.41
Service fault*	.97	.03	.99	.01	2.73	.013	.89
Net fault**	.44	.33	.66	.28	3.20	.005	1.04
Four-hit rule^{NS}	.58	.37	.71	.32	.98	.341	.32
Catch rule**	.29	.19	.56	.31	4.63	< .001	1.50
Overall accuracy**	.90	.04	.94	.04	4.42	< .001	1.47

Note. $N = 20$ participants; mean values expressed as percentile scores; ^{NS}not significant.

*$p < .05$. **$p < .01$.

Table 3

Correlation Between Referee Engagement and Officiating Decision During Preseason

Variable	Actively involved	Passively involved	Distracted
In/out rule	-.19^{NS}	.10^{NS}	.37^{NS}
Service fault	-.13^{NS}	.14^{NS}	.06^{NS}
Net fault	.05^{NS}	-.03^{NS}	-.11^{NS}
Four-hit rule	-.17^{NS}	.09^{NS}	.35^{NS}
Catch rule	-.03^{NS}	-.01^{NS}	.12^{NS}
Overall accuracy	-.06^{NS}	-.02^{NS}	.28^{NS}

Note. $N = 20$ participants; ^{NS}not significant.

*$p < .05$. **$p < .01$.

Table 4
Correlation Between Referee Engagement and Officiating Decision During Formal Competition

Variable	Actively involved	Passively involved	Distracted
In/out rule	.26^{NS}	-.13^{NS}	-.66*
Service fault	.06^{NS}	-.18^{NS}	.24^{NS}
Net fault	.58**	-.48*	-.60**
Four-hit rule	.02^{NS}	-.04	-.04^{NS}
Catch rule	.68**	-.71**	-.38^{NS}
Overall accuracy	.60**	-.54*	-.53*

Note. $N = 20$ participants; ^{NS}not significant.

*$p < .05$. **$p < .01$.

Discussion

This study examined university students' levels of engagement and accuracy of officiating volleyball during their first Sport Education season. This is the only study to date that provides data-driven evidence regarding officials' engagement and accuracy of officiating calls during a college-level activity course. The results of the study indicated that the students reported high levels of active engagement and improvements in officiating accuracy. The results also showed a significant connection between engagement and officiating accuracy. The results of this study contribute to and extend the literature related to officiating within pedagogical models, specifically within Sport Education. Furthermore, this is the first study to scrutinize each referee's decision in each game of the complete Sport Education season.

In relation to student engagement in officiating, the results of this study indicate a significant increase in active engagement as the season progressed from preseason (75%) to formal competition (85%). These findings align with those in studies demonstrating high levels of student enthusiasm about officiating (Hastie, 1996;

Hastie & Sinelnikov, 2006; Layne & Hastie, 2014) and its importance (Sinelnikov & Hastie, 2010). While active engagement of officials in this iteration of Sport Education was slightly lower than that of officials in previous studies, the university students were observed being highly and actively engaged in their role as official. This finding suggests that university students take duty-team roles of an official seriously. Because the literature identified a link between enjoyment and engagement (Hastie & Sinelnikov, 2006), the finding of this study of high levels of active engagement underscores a potential for some students to engage in officiating outside of a class setting, a possibility that was anecdotally confirmed when some participants elected to become volleyball officials for intramural volleyball at the university. With a lack of referees at every level (Ridinger, 2015), future research that examines university-aged students' perceptions of officiating and explores their intent to further pursue officiating following a Sport Education season may prove fruitful.

The results also revealed a significant increase in overall officiating accuracy as students improved from the preseason (90% accurate) to formal competition (94% accurate). These were similar findings to those in research that indicated significant improvement between the formal competition and postseason phase (Layne & Hastie, 2014). Although the accuracy level of making officiating calls in this study was relatively high, the simplicity and quantity of the in/out (98%, 99%) and service line calls (97%, 99%) in the preseason and formal competition, respectively, contributed to high percentages of correct calls.

While overall accuracy was favorable, students exhibited an inability to correctly identify net calls (44% accurate) and illegal hits (29% accurate) during the preseason. During completion phase, the students significantly raised their accuracy of both judgment calls; however, their accuracy was still low (net calls, 66% accurate; illegal hits, 56% accurate). These findings confirm that student learning takes time and students need deliberate practice to improve. Similar to the development of skill competence and tactical knowledge (Hastie et al., 2009), the development of officiating skills in Sport Education results from the structure of the model in which students are afforded significant practice in an authentic setting that is meaningful to them. Nonetheless, 56% and 66% accuracy of officiating

calls during formal competition still seems poor, especially with the high standards customary in place for all officials (Burke et al., 2000); however, because of the difficulty of determining these infractions, the improvement rate is encouraging.

Furthermore, these results also suggest that level of accuracy depends on the level of ambiguity and complexity of an officiating call. In this study, students exhibited high levels of accuracy of officiating calls that have low levels of ambiguity. For example, officials determined with a high rate of success whether the ball landed in or out of bounds or whether the server stepped on the end line during the serve. As the level of ambiguity of an officiating call increased (e.g., Was that a prolonged contact with a ball that constitutes a lift or was it not? Was it a double hit or not?), the level of accuracy of officiating calls diminished. However, one of the most encouraging aspects of these results is that students can significantly improve even when dealing with highly complex and ambiguous officiating calls during a Sport Education season.

Each Sport Education season requires all students to perform the role of official and enforce the rules of the game. The findings of this study suggest that although students can obtain a high standard of accuracy, especially with nonambiguous officiating calls, this does not happen automatically. To be successful, students need a deep understanding of gameplay and a significant knowledge of the rules (Layne & Hastie, 2014). Considering the results of this study, when employing Sport Education teachers need to realize that officiating competency does not occur simply because students have a chance to officiate. Devoting adequate time to training officials (Siedentop et al., 2011), teaching and assessing game performance in-season (Farias, Mesquita, & Hastie, 2015), and providing ongoing training to officials is vital to officiating success. Further research needs to identify best practices for adequately providing instruction and training to student officials in a variety of sporting contexts, not just team sports.

The findings also suggest that active engagement was associated with higher accuracy for the net faults and illegal hit calls. Similarly, passive engagement was associated with lower percentages of overall correct calls. While some studies have indicated the connection between engagement and accuracy (Layne & Hastie, 2014),

this is the first study that confirmed this relationship. The findings demonstrate the necessity for officials to be actively involved while officiating. With this in mind, instructors may find it prudent to incorporate the officiating duty into the record-keeping component of the model. It is unlikely that student officials will call the perfect game (Downes, 2016); however, maintaining active engagement is realistic, tangible, and, as this study demonstrated, associated with officiating accuracy.

This study has several limitations. First, we measured each student's engagement and accuracy data at two points in the season, preseason and formal competition. While this is fairly typical for a round-robin format, other Sport Education competition formats may provide more opportunities for students to officiate. Second, with the absence of a control group, we found it difficult to determine if the increases in accuracy and engagement were due to students' participation in Sport Education or another extraneous variable. Finally, the sample size for this study was limited to 20 participants and the content was limited to a net/wall game of volleyball. Moving forward, future research could utilize several classes, increase the sample, and examine different context to add to the generalizability of the results.

In conclusion, this study provided empirical evidence that suggests high levels of student engagement and improvement of officiating accuracy in a season of Sport Education. While participants in this study had no experience with Sport Education, future research could examine the effects on officiating for participants over multiple Sport Education seasons. Additionally, researchers could investigate potential transfer relative to officiating within the same game category (e.g., racquetball and tennis) and between games or activities from different game categories (e.g., between net/wall and invasion). Finally, research efforts investigating the link between content knowledge and officiating success and engagement could prove fruitful.

References

American Sport Education Program. (2011). *Successful sports officiating* (2nd ed.). Champaign, IL: Human Kinetics.

Bennett, G., & Hastie, P. (1997). A Sport Education curriculum model for a collegiate physical activity course. *Journal of Physical Education, Recreation, & Dance*, *68*(1), 39–44. https://doi.org/10.1080/07303084.1997.10604876

Browne, T. B. J., Carlson, T. B., & Hastie, P. A. (2004). A comparison of rugby seasons presented in traditional and Sport Education formats. *European Physical Education Review, 10*, 199–214. https://doi.org/10.1177/1356336X04044071

Burke, K. L., Joyner, A. B., Pim, A., & Czech, D. R. (2000). An exploratory investigation of the perceptions of anxiety among basketball officials before, during, and after the contest. *Journal of Sport Behavior, 23*, 11–19.

Catteeuw, P., Helsen, W., Gilis, B., & Wagemans, J. (2009). Decision-making skills, role specificity, and deliberate practice in association football refereeing. *Journal of Sport Sciences, 27*, 1125–1136. https://doi.org/10.1080/02640410903079179

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Pyschological Bulletin, 122*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cooper, J. O., Heron, T. E., & Heward, W. L. (2014). *Applied behavior analysis*. London, England: Pearson Education.

Dorsch, K. D., & Paskevich, D. M. (2007). Stressful experiences among six certification levels of ice hockey officials. *Psychology of Sport & Exercise, 8,* 585–593. https://doi.org/10.1016/j.psychsport.2006.06.003

Downes, C. (Producer). (2016, September 27). *Real sports with Bryant Gumbel: Man vs. machine* [Television broadcast]. New York, NY: Home Box Office.

Farias, C. F., Mesquita, I. R., & Hastie, P. A. (2015). Game performance and understanding within a hybrid Sport Education season. *Journal of Teaching in Physical Education*, *34*, 363–383. https://doi.org/10.1123/jtpe.2013-0149

Federation Internationale De Volleyball. (2016). *Official volleyball rules 2017–2020.* Retrieved from http://www.fivb.org/EN/Refereeing-Rules/documents/FIVB-Volleyball_Rules_2017-2020-EN-v06.pdf

Goldsmith, P. A., & Williams, M. J. (1992) Perceived stressors for football and volleyball officials from three rating levels. *Journal of Sport Behavior, 15*, 106–118.

Hastie, P. (1996). Student role involvement during a unit of Sport Education. *Journal of Teaching in Physical Education, 16,* 88–103. https://doi.org/10.1123/jtpe.16.1.88

Hastie, P., & Sinelnikov, O. (2006). Russian students' participation in and perceptions of a season of Sport Education. *European Physical Education Review, 12,* 131–150. https://doi.org/10.1177/1356336X06065166

Hastie, P. A., & Sinelnikov, O. A. (2007). The use of Web-based portfolios in college physical education activity courses. *Physical Educator*, *64*, 21–28.

Hastie, P., Sinelnikov, O., & Guarino, A. (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

Kellett, P., & Warner, S. (2011). Creating communities that lead to retention: The social worlds and communities of umpires. *European Sport Management Quarterly, 11*, 471–494. https://doi.org/10.1080/16184742.2011.624109

Kim, J., Penney, D., Cho, M., & Choi, H. (2006). 'Not business as usual': Sport Education pedagogy in practice. *European Physical Education Review, 12*, 361–379.

Kinchin, G. D., Wardle, C., Roderick, S., & Sprosen, A. (2004). A survey of year 9 boys' perceptions of Sport Education in one English secondary school. *Bulletin of Physical Education, 40*, 27–40.

Layne, T., & Hastie, P. (2014). Active involvement and accuracy of calls of novice referees during a season of Sport Education. *Physical Educator*, *71*, 459–471.

MacPhail, A., & Kinchin, G. (2004). The use of drawings as an evaluative tool: Students' experiences of Sport Education. *Physical Education and Sport Pedagogy, 9*, 87–108. https://doi.org/10.1080/1740898042000208142

Mascarenhas, D. R. D., Collins, D., & Mortimer, P. (2005). The accuracy, agreement, and coherence of decision-making in rugby union officials. *Journal of Sport Behavior, 28*, 253–271. https://doi.org/10.1123/jtpe.25.1.9

Mowling, C. M., Brock, S. J., & Hastie, P. A. (2006). Fourth grade students' drawing interpretations of a Sport Education soccer unit. *Journal of Teaching in Physical Education, 25*, 9–35.

Rainey, D. (1995). Sources of stress among baseball and softball umpires. *Journal of Applied Sport Psychology, 7*, 1–10. https://doi.org/10.1080/10413209508406296

Rainey, D. W., & Hardy, L. (1999). Sources of stress, burnout, and intention to terminate among rugby union referees. *Journal of Sports Sciences, 17*, 797–806. https://doi.org/10.1080/026404199365515

Ridinger, L. L. (2015). Contributors and constraints to involvement with youth sports officiating. *Journal of Amateur Sport*, *1*, 103–127. https://doi.org/10.17161/jas.v1i2.4946

Siedentop, D. (1994). *Sport Education: Quality PE through positive sport experiences*. Champaign, IL: Human Kinetics.

Siedentop, D., Hastie, P., & van Der Mars, H. (2011). *Complete guide to Sport Education* (2nd ed.). Champaign, IL: Human Kinetics.

Sinelnikov, O. A., & Hastie, P. A. (2010). Students' autobiographical memory of participation in multiple sport education seasons. *Journal of Teaching in Physical Education, 29*, 167–183.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

van der Mars, H. (1989). Observer reliability: Issues and procedures. In P. W. Darst, D. Zakrajsek, & V. H. Mancini (Eds.), *Analyzing physical education and sport instruction* (2nd ed., pp. 55–80). Champaign, IL: Human Kinetics.

Voight, M. (2009). Sources of stress and coping strategies of US soccer officials. *Stress & Health: Journal of the International Society for the Investigation of Stress*, *25*, 91–101. https://doi.org/10.1002/smi.1231

Wahl-Alexander, Z., Sinelnikov, O., & Curtner-Smith, M. (2017). A longitudinal analysis of students' autobiographical memories of participation in multiple Sport Education seasons. *European Physical Education Review, 23*, 25–40. https://doi.org/10.1177/1356336X15624246