short communication

The role of team and sport social contexts: Are three-level models needed in studies of sport conduct?

Mariya A. Yukhymenko-Lescroart

Department of Curriculum and Instruction, California State University, Fresno 5005 N. Maple Ave., M/S ED 2, Fresno, CA 93740, USA

ARTICLE INFO

Keywords:
Multilevel modeling
Sport conduct
Invasion team sports
Collusion
Contact

ABSTRACT

In this work, I examined how variation in sport conduct is allocated across the three different levels (i.e., athlete, team, and sport) in order to consider how the team and sport environment might contribute to athletes’ sport conduct. Specifically, three-level models were utilized for sportspersonship, technical and mental gamesmanship, and instrumental aggression to examine whether there is unique variance due to sport and team clustering with athletes (N = 1412) nested within teams (J = 92) that are in turn nested within sports (K = 17). Results showed that clustering due to sport accounted for unique variance above and beyond team clustering. However, sex and invasion team sport explained most of the sport-level variance. The implication is that it is important to account for the sport clustering. However, when accounting for the sport-level clustering is not feasible, controlling for sex is necessary at the very least.

Several recently published systematic reviews about the development of empathy and prosocial behavior (Silke et al., 2018), fair play and sporting competition (Serrano-Dura et al., 2020), and moral development and sportsmanship (Pennington, 2017) have highlighted that there are several opportunities for athletes to learn behaviors from their athletic environments, including team and sport.

A team ethical atmosphere is established through athletes’ interactions with teammates and coaches. A team’s training and competition context is linked with sport ethical conduct (Serrano-Dura et al., 2020), because it is shaped by a collective set of values, attitudes, and norms (Taylor, 2020). For example, a team environment that emphasizes winning at all cost may facilitate unethical conduct. Pennington (2017) argued that athletes’ unethical sport conduct is learned and reinforced from the leadership of coaches. Indeed, coaches’ ethical and abusive behavior have been linked with athletes’ perception of inclusion climate on the team and team members’ willingness to cheat (Yukhymenko-Lescroart, 2015). Above and beyond team influences, athletes engage with their opponents, which creates further contextual (e.g., local, regional, national) influences. While not all athletes might directly experience various levels of contextual influences, watching others play (e.g., on media) subjects athletes to sport-specific customs, norms, and culture. Consequently, there are two main common contextual levels in which sport behaviors of athletes are formed: the team level and the sport level.

Considering sport context is important because it can provide further insights into sporting norms and deviance and help identify areas for the improvement of athletes’ character development. Norms, defined as a shared expectation that are used to identify what is acceptable and unacceptable in a social world (Coakley, 2015), serve as the moral standards used to identify deviance. Each sport has formal norms, official expectations and written rules, and informal norms, a shared understanding of unwritten rules, customs, ways of thinking, and actions. Coakley (2015) pointed out that actions, which may be defined as deviance in other settings, can be accepted and defined as normal in sports. In some sports (e.g., boxing, men’s ice hockey), athletes are encouraged to engage in behaviors that are outlawed and classified as criminal in real life. As a result, sport culture has implications for the attitudes, perceptions, emotional responses, norms, and behaviors. For example, doping, or the use of banned performance-enhancing substances or methods, is more common in some sports (e.g., boxing, cycling, weightlifting) than others (WADA, 2019); likewise, the use of gamesmanship, or the use of dubious methods of winning that are within the rules, is more prevalent in some sports (e.g., diving/flopping in soccer). Yukhymenko-Lescroart (2019) posited that, because the views of what specific conduct is perceived as “normal” is not universal across all sports, it is important to identify and examine sporting cultures that are more likely to jeopardize athletes’ character.

Several studies addressing sportspersonship, aggression, willingness to cheat, and other aspects of moral functioning in sport have utilized a multilevel approach (e.g., Chow et al., 2009; Gano-Overway et al., 2005;
In these studies, researchers have used two-level models, nesting athletes within teams, which accounted for the contextual influences of a team culture on one’s behavior and generally showed the presence of interdependence of individual measures collected within a team setting. However, past studies have often failed to address the doubly nested structure of athletes who are nested within teams that are in turn nested within sports. Because parsimony is commonly sought (i.e., simpler models are preferred over complex ones, which is often referred to as “Occam’s razor,” Vandekerckhove et al., 2015), it is important to examine whether there is unique variance due to sport clustering above and beyond team clustering. Theoretically, sport clustering could be due to the sociodemographic characteristics (e.g., sex and the type of sport) that Serrano-Durá et al. (2020) highlighted as important, but less studied in relation to ethical sport conduct. Sex has been shown to be a significant team-level variable in studies on sport conduct employing two-level models (e.g., Chow et al., 2009; Yukhymenko-Lescroart et al., 2015), with some evidence pointing to sex as a relevant sport-level construct. Specifically, Yukhymenko-Lescroart et al. (2015) showed that men’s teams in general were much more willing to cheat to win a game than women’s teams, with men’s football, basketball, and baseball (i.e., conceptualized as “high profile” sports) reporting the highest willingness to cheat compared to other men’s and women’s sports. Serrano-Durá et al. (2020) noted that there is a relationship between the type of sport and anti-sporting conduct, with greater contact and impact goals, share common strategies, tactics, and patterns of play (Lamas et al., 2014), and unethical sport conduct, such as gamesmanship and instrumental aggression, can be viewed as a strategy to win (Yukhymenko-Lescroart, 2016), the sport culture of invasion team sports is likely to be more open to winning through dubious strategies. Therefore, the purpose of this work was to examine how variation in ethical and unethical sport conduct is allocated across the three different levels (i.e., athlete, team, and sport). Additionally, this work examined the extent to which sport characteristics (i.e., sex and invasion team sport) explained variance at the sport level.

1. Method

For the purposes of this brief research note, a part of existing data from a larger project was analyzed. Participants were 1412 athletes (44.2% females) participating in college varsity or club team sports. These participants came from 92 teams, representing 17 sports, including nine men’s sports: basketball (10.8%), basketball (3.0%), football (18.8%), ice hockey (1.3%), lacrosse (7.3%), roller hockey (0.3%), rugby (3.7%), soccer (9.4%), and volleyball (1.2%), and eight women’s sports: basketball (6.2%), field hockey (3.7%), ice hockey (2.6%), lacrosse (7.1%), rugby (0.4%), soccer (11.8%), softball (5.7%), and volleyball (6.5%). The following sports were classified as non-invasion: men’s baseball, men’s and women’s volleyball, and women’s softball; all other sports were classified as invasion.

After ethical approvals were obtained, participants were invited to complete a number of instruments in a paper-and-pencil format, including the Conduct in Sport toward Opponent Scale (CSOS, Yukhymenko-Lescroart, 2016). The CSOS consists of 18 items designed to measure the self-reported frequency of ethical and unethical conduct among athletes in team sports. Examples of items are “I stopped playing when an opponent got injured (not my fault),” “I created a foul situation to gain advantage over an opponent,” “I used physical taunting (e.g., making faces or showing obscene gestures) to make an opponent commit a mistake,” and “I purposely injured an opponent to win.” For each item, athletes indicated frequency of conduct during the last year on a (never) to 5 (almost daily) scale with all response options labeled. The CSOS was originally developed as a 3-factor scale, measuring sportsmanship, gamesmanship, and instrumental aggression; however, Yukhymenko-Lescroart (2019) later suggested that more work should be done to further extrapolate the spectrum of athletes’ conduct in sport and proposed that the gamesmanship dimension can further be broken into two subtypes to address the technical and mental aspects of gamesmanship.

Data were analyzed in Mplus, Version 8.3 (Muthén & Muthén, 2012–2019). First, following Yukhymenko-Lescroart’s (2019) recommendation, two CFA models were tested to examine dimensionality of the CSOS: a 3-factor model and a 4-factor model. The CFA models were estimated using maximum likelihood with robust standard errors (MLR), which is robust to non-normality and missing data, accounting for the clustered nature of the data. The CFA models were evaluated based on a chi-square test and a combination of fit indices, including cutoff values close to 0.95 for CFI and TLI (higher is better), in combination with SRMR < 0.09 and RMSEA < 0.06 for RMSEA (Hu & Bentler, 1999). Mean scores were then computed for the measures of sport conduct and subjected to multilevel modeling as main analysis.

The multilevel models were estimated using Bayesian Markov Chain Monte Carlo (MCMC), which has been shown to achieve unbiased estimates with low numbers of clusters, even when fewer than 10 clusters are present (McNeish & Stapleton, 2016). Prior distributions were not specified to not let prior beliefs influence the results (Raudenbush & Bryk, 2002). Bayesian exploration of model fit can be done based on Posterior predictive checking (PPC, Gelman et al., 1996); however, in this study the multilevel models were just-identified, that is, no restrictions on the means, variances, and covariances were imposed, thus, PPCs are not reported (Muthén et al., 2017, p. 408). First, three-level random-effect unconditional (i.e., baseline) models were specified, in which athletes were nested within teams nested within sports, to examine how much variance was in common at each level of analysis. As a result of these analyses, the intraclass correlation coefficients (ICC, Raudenbush & Bryk, 2002), which measure the proportion of the variance explained by the nested structure (Hox, 2002), were obtained. ICC, which is also known as cluster effect, indicates whether multilevel modeling is appropriate and needed, because, when assumption of independence of observation is not met, standard errors cannot be correctly estimated. In the 3-level model used in this paper, ICC for level-2 (team level) can be interpreted as the degree of similarity in sport conduct among athletes within teams after accounting for sport-level similarity. Likewise, ICC for level-3 (sport level) can be interpreted as the degree of similarity in sport conduct among athletes within sports after accounting for team-level similarity. The larger the ICC, the larger the deviation from the assumption of independence (i.e., more similar the observations are within a nested level) and the larger clustering effects are on standard errors. One motivation for the use of multilevel modeling is to correctly estimate standard errors in complex sampling designs. To better understand the effects of ICC on standard errors, design effect (DEFF), which shows the extent to which standard errors would be underestimated with positive ICCs if a conventional single-level analysis was used (Muthén & Satorra, 1995), were also computed. As a rule of thumb, design effects are considered high when DEFF is greater than 2, indicating that a multilevel approach to data analyses is warranted (Muthén & Satorra, 1995), but even DEFF as small as 1.1 can warrant multilevel modeling under certain conditions (Lai & Liu, 2012). In the final step, invasion and non-invasion team sport, which were dummy coded dichotomous variables (i.e., 0 vs. 1, with men and non-invasion sports as reference groups), were added to the models at level-3 (sport level). The convergence, which was checked based on the Potential Scale Reduction values being close to 1 (Muthén et al., 2017), was obtained for all models; all variance and covariance estimates were admissible. Full results for models, including 95% credibility intervals and effect sizes, are provided in online supplementary material.
2. Results

2.1. Preliminary analysis

Missing data on individual items ranged from 0 to 2.3% and were assumed to be missing at random. Two CFA models were tested: a 3-factor model and a 4-factor model. While both the 3-factor CFA model ($\chi^2 (132, N = 1412) = 784.93, \text{RMSEA} = 0.059, \text{CI}[0.055, 0.063], \text{SRMR} = 0.055$) and the 4-factor model ($\chi^2 (129, N = 1412) = 683.73, \text{SCF} = 1.41, p < .001, \text{CFI} = 0.949, \text{TLI} = 0.939, \text{RMSEA} = 0.055, \text{CI}[0.051, 0.059], \text{SRMR} = 0.053$) showed good fit to the data, the fit of the 4-factor model was superior to the 3-factor model based on fit indices and the chi-square test of difference, $\chi^2 (3, N = 1412) = 157.12, p < .001$. Therefore, the 4-factor solution was accepted as final, which included the following four factors: (a) a 6-item sportspersonship, (b) a 3-item technical gamesmanship, (c) a 4-item mental gamesmanship, and (d) a 5-item instrumental aggression. Scores for the four scales were computed for all cases who responded to at least one item, and no missing data were present. Reliability estimates, descriptive statistics, and correlation estimates for the four measured variables are presented in Table 1. Reliability estimates ranged from 0.73 to 0.97.

2.2. Three-level unconditional models

Table 2 shows results for the four three-level unconditional models for sportspersonship, technical aspects of gamesmanship, mental aspects of gamesmanship, and instrumental aggression, with two “between” levels: team (level-2) and sport (level-3). Results provided evidence for the nested structure for the four dimensions of sport conduct. Specifically, between 11.9% and 28.7% of the unique variance in outcomes was attributed to the nested structure of the data (team and sport levels combined); whereas DEFFs for sport level clustering ranged from 1.98 to 5.05. Both ICCs and DEFFs were the lowest for self-reported frequency of sportspersonship and the highest for mental gamesmanship. Overall, these results suggested that when examining sport conduct, it is important to account for team and sport clustering because the shared variance indicates that some of athletes’ sport conduct can be explained through contextual influences of the team and sport culture.

For all outcomes, sport clustering accounted for more variance than team clustering, though for sportspersonship this difference was negligible: 5.8% of the variance was uniquely accounted for by team clustering, and an additional 6.1% of the variance by sport clustering. Specifically, an unconditional model for sport conduct can be helpful to account for team and sport clustering because the shared variance indicates that some of athletes’ sport conduct can be explained through contextual influences of the team and sport culture.

For all outcomes, sport clustering accounted for more variance than team clustering, though for sportspersonship this difference was negligible: 5.8% of the variance was uniquely accounted for by team clustering, and an additional 6.1% of the variance by sport clustering. Specifically, an unconditional model for sport conduct can be helpful to account for team and sport clustering because the shared variance indicates that some of athletes’ sport conduct can be explained through contextual influences of the team and sport culture.

Table 1

<table>
<thead>
<tr>
<th>Measured Variable</th>
<th>Reliability ( \alpha ) (n)</th>
<th>M</th>
<th>SD</th>
<th>Correlation Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Sportspersonship</td>
<td>.80 (.79)</td>
<td>3.39</td>
<td>0.91</td>
<td>-</td>
</tr>
<tr>
<td>Gamesmanship:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical aspects</td>
<td>.73 (.79)</td>
<td>2.39</td>
<td>1.15</td>
<td>-.07*</td>
</tr>
<tr>
<td>Mental aspects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamesmanship:</td>
<td>.82 (.87)</td>
<td>2.36</td>
<td>1.16</td>
<td>-.14***</td>
</tr>
<tr>
<td>Mental aspects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrumental Aggression</td>
<td>.97 (.97)</td>
<td>1.45</td>
<td>.94</td>
<td>.02</td>
</tr>
</tbody>
</table>

Note. \( \alpha \) = Omega reliability. \( \alpha \) = Cronbach’s alpha. The 95% confidence intervals for correlation estimates are presented in brackets. ‘*’ \( p < .05 \), ‘**’ \( p < .01 \), ‘***’ \( p < .001 \).

3. Discussion

Clustered data are extremely common in sport research, with typical examples of athletes nested within teams. This study examined the variation in ethical and unethical sport conduct at the athlete, the team, and the sport levels. As expected (e.g., Yukhymenko-Lescroart, 2018; Yukhymenko-Lescroart et al., 2015), findings showed that team accounted for unique variance above and beyond sport clustering for all outcomes. Yet, unethical conduct was primarily driven by the sport norms, which was especially true for gamesmanship. These findings extend previous research (e.g., Kavussanu & Ring, 2020; Shields et al., 2016) with regards to how bracketed morality and interpretation of contesting environments is manifested in athletic contexts. Yet, the current research is the first to explicitly differentiate between the sources of bracketed morality with regards to whether it is team- or sport-driven, and the first to relate these sources to different types of unethical conduct, such as gamesmanship (i.e., using dubious methods of winning) and instrumental aggression (i.e., using illegal methods of winning).

In addition to team-level clustering, sport-level clustering was also evident for all outcomes, pointing to the importance of the influences of the sport norms above and beyond the team ethos. While previous studies accounted to the team-level clustering (e.g., Chow et al., 2009; Gano-Overway et al., 2005; Yukhymenko-Lescroart et al., 2015), findings in this study highlighted the importance of sport-specific practices and norms in athletes’ self-reported frequency of sportspersonship, technical and mental gamesmanship, and instrumental aggression.

This work also examined the extent to which sex and invasion team sport explained sport-level variance, because understanding how these sport characteristics contribute to sport-level clustering can be helpful.
for practice and research. Consistent with expectations (e.g., Yukhymenko-Lescroart et al., 2015), sex differences were found for all types of sport conduct, with athletes in women’s sports reporting engaging in unethical conduct more frequently and in unethical conduct less frequently compared to men’s sports athletes. Yet, for invasion team sports, differences were found for gamesmanship that focuses on technical aspects only. Lamas et al. (2014) proposed that invasion team sports utilize team strategies for winning that support a collective performance. Because unethical conduct can be viewed as a team strategy to win (Yukhymenko-Lescroart, 2016), the finding that invasion team sports reported more frequent engagement in technical and mental gamesmanship compared to athletes in non-invasion team sports was expected. Interestingly, among all types of sport conduct examined in this study, mental gamesmanship had the highest values of sport-level intraclass coefficients. This was also the only sport conduct outcome for which all of the sport-level variance was completely explained by sex and invasion team sport.

These findings lead to several recommendations. First, findings confirmed the nested nature of the sport conduct data, suggesting that there are unique contextual influences of the team and the sport environments; therefore, using a multilevel approach to data analysis is required. McNeish and Wentzel (2017) showed the importance of accounting for the incidental level-3 clustering to obtain trustworthy estimates, even when ICC was as low as 0.05. The magnitude of the ICC correction for variance parameters and, thus, unbiased results; but it is very large (Elff et al., 2020), the ML variance parameters in multilevel models have a small-sample bias, even when the sample size at the group level is large (Elff et al., 2020). REML, on the other hand, provides bias and small-sample bias, even when the sample size at the individual level is very large (Elff et al., 2020). REML, on the other hand, provides bias correction for variance parameters and, thus, unbiased results; but it

#### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Sportspersonship</th>
<th>Gamesmanship: Technical Aspects</th>
<th>Gamesmanship: Mental Aspects</th>
<th>Instrumental Aggression</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC: Athlete level</td>
<td>88.1%</td>
<td>81.2%</td>
<td>71.3%</td>
<td>82.7%</td>
</tr>
<tr>
<td>ICC: Team level</td>
<td>5.8%</td>
<td>1.7%</td>
<td>3.4%</td>
<td>5.4%</td>
</tr>
<tr>
<td>ICC: Sport level</td>
<td>6.1%</td>
<td>17.1%</td>
<td>25.3%</td>
<td>11.9%</td>
</tr>
<tr>
<td>ICC: Sport and team level (combined)</td>
<td>11.9%</td>
<td>18.8%</td>
<td>28.7%</td>
<td>17.3%</td>
</tr>
<tr>
<td>DEFF: Sport level</td>
<td>1.98</td>
<td>3.74</td>
<td>5.05</td>
<td>2.90</td>
</tr>
</tbody>
</table>

Note. ICC: Team Level = % of variance in outcomes that is accounted at the sport level; ICC: Sport Level = % of variance in outcomes that is accounted at the sport level; ICC: Sport and Team Level (Combined) = % of variance in outcomes that is accounted at the level of teams nested within sports. DEFF = design effect; DEFF is computed as: DEFF = 1 + (average cluster size - 1) * ICC.

#### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Sportspersonship</th>
<th>Gamesmanship: Technical Aspects</th>
<th>Gamesmanship: Mental Aspects</th>
<th>Instrumental Aggression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects: Sport Level</td>
<td>.37*** (.10)</td>
<td>-.52** (.16)</td>
<td>-.91*** (.03)</td>
<td>-.52*** (.11)</td>
</tr>
<tr>
<td>Invasion team sport</td>
<td>-.06 (.11)</td>
<td>.53** (.19)</td>
<td>.19 (.15)</td>
<td>.13 (.12)</td>
</tr>
<tr>
<td>Null Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>.05 (.04)</td>
<td>.25 (.18)</td>
<td>.38 (.24)</td>
<td>.11 (.07)</td>
</tr>
<tr>
<td>Model with sex and invasion team sport</td>
<td>.01 (.02)</td>
<td>.06 (.07)</td>
<td>.02 (.04)</td>
<td>.01 (.02)</td>
</tr>
<tr>
<td>Proportion of explained variance</td>
<td>76.5%</td>
<td>77.0%</td>
<td>94.8%</td>
<td>88.3%</td>
</tr>
</tbody>
</table>

Note. Posterior SDs are presented in parenthesis. Sex: 0 = men’s sports, 1 = women’s sports. Invasion team sport: 0 = non-invasion, 1 = invasion. *p < .05, **p < .01, ***p < .001. Because Bayesian estimation method was used, these are one-tailed p-values. Full results for models, including 95% credibility intervals and effect sizes, are presented in online supplementary material.
cannot be used to compare two nested models with a likelihood ratio test (Raudenbush & Bryk, 2002). The second consideration is that the minimum required sample size at each level depends on the complexity of the model. A general recommendation is that the number of clusters needs to be larger than the number of estimated parameters; otherwise, while the software might estimate such a model, estimated parameters might not be reliable (Silva et al., 2019). Yet, these recommendations are based solely on obtaining unbiased estimates and do not consider issues related to statistical power. Substantive researchers should use a Monte Carlo simulation study to decide on sample size and determine power (e.g., see Muthén & Muthén, 2002).

In conclusion, the self-report nature of the current data prevents from making inferences about actual athletes’ behaviors; this is especially true for unethical conduct, the self-reported frequency of which could be prone to the social desirability bias. Yet, this study added to the body of literature on the contextual influences of the team and the sport norms in athletes’ ethical and unethical sport conduct. The practical significance of this work is that coaches and athletes should be aware that the team ethos and sport norms matter in how they conduct themselves during a sport competition. More attention should be given to promoting environments conducive to athletes’ ethical conduct in future studies.

Author Statement

The preparation of this manuscript, including the conceptualization of this study, data collection, data analyses, and writing, was completed solely by Dr. Mariya Yukhymenko-Lescroart.

Author note

Data collection were sponsored by the National Collegiate Athletic Association.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.psychsport.2020.101848.

References


