Sports Participation and Moral Development Outcomes: Examination of Validity and Reliability of the Prosocial and Antisocial Behavior in Sport Scale

by Chunxiao Li, Koon Teck Koh, John Wang and Lit Khoon Chian

Reprinted from International Journal of Sports Science & Coaching Volume 10 · Number 2+3 · 2015
Sports Participation and Moral Development Outcomes: Examination of Validity and Reliability of the Prosocial and Antisocial Behavior in Sport Scale

Chunxiao Li1, Koon Teck Koh2, Chee Keng John Wang2, and Lit Khoon Chian2
1Department of Health and Physical Education, The Hong Kong Institute of Education, 10 Lo Ping Road, New Territories, Hong Kong
E-mail: cxlilee@gmail.com
2National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore, 637616

ABSTRACT
The purpose of the study was to examine the psychometric properties of the Prosocial and Antisocial Behavior in Sport Scale (PABSS) developed by Kavussanu and Boardley [13] to measure moral behaviors in Singapore’s context. A total of 574 (boys = 296, girls = 278) school team athletes were recruited for the study. Results showed that internal reliability, convergent validity and discriminant validity of the PABSS were supported. Evidence of configural, metric and scalar invariance of the PABSS across school and gender groups were also found. The findings suggest that the PABSS can be used to measure prosocial and antisocial behaviors among school athletes in Singapore.

Key words: Moral Behaviour, Scholastic Athletics, Social Cognitive Theory, Sport Participation

INTRODUCTION
Participation in sports can lead to positive experiences and beneficial outcomes such as increased self-esteem, confidence, citizenship, character building, identity development, meaningful adult and peer relationships, academic achievement, and decreased delinquency [1-4]. Participation in organized sports thus provides an appealing environment for acquisition of sports skills, as well as life skills that enhance personal growth [5]. This is especially so for Singapore schools as Co-Curricular Activities, which include a range of sports and games, are mandatory components in the curriculum [6]. Besides, there is an increase in the concern for personal development in the younger generation through sports, as noted in the significant rise in incorporating social-emotional learning approaches and development in many present-day organizations [5], including the Ministry of Education in Singapore [6]. At the recent National School Games opening ceremony, the Minister for...
Education, Mr. Heng Swee Keat reiterated the important link between sporting excellence and character development, and highlighted the use of sports as a vehicle to facilitate students to learn life lessons and develop valuable character traits, which will be useful to them at the later part of their lives [7].

However, sports participation is associated with negative outcomes as well, particularly due to its competitive nature and the excessive pressure to win. Aggression has long been considered as a major problem, observable in sporting events from youth sports to games played at the highest levels of competition [1]. Furthermore, the professionalization of sports has led to increasing pressures on competition outcomes, inducing the ‘winning at all costs’ mentality highly associated with substance abuse, along with reducing competition fairness and demeaning positive societal values [3,4]. Therefore, understanding the moral behaviors of youth participation in organized sports is an important area of research.

The term moral behavior refers to “a broad range of intentional acts that could result in positive or negative consequences for others’ psychological and physical welfare” [8, p. 366]. Sport behaviors are morally relevant as they can positively or negatively affect other people who are involved in sports [8]. While good sport behaviors (e.g., encouraging teammates and helping fallen opponents get off the floor) are expected in youth sport contexts; negative sport behaviors especially towards opponents (e.g., shoving and taunting opponents) are not unusual and often observed [9]. In addition, empirical evidence has shown that negative sport behaviors can result in both physical and emotional injury [10]. Thus, investigation of sport behaviors merits attention to advance knowledge within this field.

According to social cognitive theory, there are dual aspects of sport behaviors: proactive and inhibitive [11, 12]. Proactive sport behaviors are expressed in the power to behave humanely, whereas inhibitive sport behaviors are manifested in the power to refrain from behaving inhumanely [11]. In that view, high levels of morality are evident when athletes manifest positive sport behaviors as well as refrain from expressing negative sport behaviors. However, the majority of studies that examined morally relevant sport behaviors have focused on negative aspects including aggressive behavior, hitting opponents and poor sportsmanship [13]; only recently have both negative and positive moral behaviors been investigated in the sport settings. The terms ‘prosocial behavior’ and ‘antisocial behavior’ are therefore used for referring to good and bad sport behaviors respectively [9, 14]. According to the social cognitive theory [11], conducting appropriate or inappropriate sport behaviors are determined by multidimensional rules (e.g., consequences of the conduct and whether it is directed at other athletes). Athletes may learn good or bad behaviors through observing or receiving reinforcement from significant others such as coaches, opponents and teammates [11].

Building upon the social cognitive theory and the definitions of prosocial and antisocial behaviors, the Prosocial and Antisocial Behavior in Sport Scale (PABSS) was developed and validated using British samples [13]. It was an important step forward as researchers have advocated a need to establish a sound instrument for measuring sport behaviors of morality [11]. Some researchers have suggested the potential of using the PABSS to monitor youth athletes’ moral development [15], and understanding the social and moral behaviors occurring in sports [16].

The PABSS consists of four factors: a) prosocial behavior toward teammates, b) prosocial behavior toward opponents, c) antisocial behavior toward teammates, and d) antisocial behavior toward opponents. The internal reliability (i.e., all cronbach’s α values > .70), factorial validity (i.e., first-order four-factor model), concurrent validity (e.g., its subscales
including prosocial opponent behavior, antisocial opponent behavior, and antisocial teammate behavior were correlated with empathy), and discriminant validity (i.e., strength of factor correlations ranged from small to strong) of the PABSS were supported using the British team-sport athletes (Mage = 21.64 years) [13]. The measurement invariance of the scale was also evidenced regarding gender and sport type such as soccer, rugby, hockey, basketball and netball [13].

Testing measurement invariance concerns the extent to which instrument items convey the same meaning to different group of participants (e.g., gender, school, sport and country) [17,18]. There are four primary forms of measurement invariance: configural, metric, scalar, and residual invariance [19]. The sequence for evaluating measurement invariance are as follows: a) test the measurement model separately in each group from the perspectives of both parsimony and substantive meaningfulness; b) examine configural or “equal form” invariance (i.e., a least-constrained model with the equal parameters across groups); c) evaluate metric invariance or “weak factorial invariance” (i.e., equal factors loadings across groups); d) assess scalar, intercept, or “strong factorial invariance” (i.e., equal observed item or indicator scores across groups); and e) test item residual, error, or “strict factorial invariance” (i.e., equal item or indicator residuals across groups) [17,20]. Examining error invariance is generally optional because it overly restrictive and unimportant for evaluation of measurement invariance [20]. In the context of substantively meaningful invariance, tests of population heterogeneity (i.e., factor variance, factor covariance, and latent means) can be conducted [20]. Evaluation of equality of factor variance (i.e., equal amount of within-group variability or dispersion of the factor across groups) and factor covariance (i.e., equal latent factor correlations across groups) usually has unclear or little implications in applied research [20]. It is useful, however, to test latent mean differences (i.e., whether groups differ in their levels of a factor) in the context of metric and scalar invariance [20]. More information about how to evaluate measurement invariance should refer to the great work done by experts within this field such as Byrne [17, 18], Meredith [19], and Brown [20].

Although the PABSS is a promising instrument for measuring morally relevant sport behaviors, its measurement invariance across different school groups has not been tested. Furthermore, the PABSS is developed using British athletes, its psychometric properties in other cultures, contexts, and independent samples have not been examined to ensure that the instrument is robust, reliable and valid [13]. Therefore, the aims of the current study were to: a) test the multigroup invariance of the PABSS across different groups (i.e., gender and school) of athletes; and b) examine other psychometric properties of the PABSS, including internal reliability, convergent validity and discriminant validity using Singapore schools athletes.

**METHOD**

**PARTICIPANTS**

A sample of 574 school athletes (secondary school, n = 354; junior college, n = 220) between 13 to 17 years old (secondary school, M = 13.62, SD = 0.51; junior college, M = 15.20, SD = 0.44) were recruited in this study. The participants consisted of 296 boys and 278 girls. The participants were from five team sports: field hockey (n = 87), soccer (n = 78), netball (n = 166), rugby (n = 109) and basketball (n = 134).

**MEASURE**

Participants’ prosocial and antisocial sport behaviors were assessed using the PABSS [13]. The PABSS consisted of 20 items describing four subscales: prosocial behavior toward
teammates (four items; e.g., “Gave positive feedback to a teammate”), prosocial behavior toward opponents (three items; e.g., “Helped an opponent off the floor”), antisocial behavior towards teammate (five items; e.g., “Argued with a teammate”), and antisocial behavior towards opponent (eight items; e.g., “Criticised an opponent”). The internal reliability of the four subscales ranged from .74 to .86 in the British sample [13]. A five-point Likert scale was used for responses (1 = never; 5 = very often).

PROCEDURE
Data were collected at different time points of the season for different teams, which lasted for about three months. Before the data collection, the ethical approval from the principal investigator’s institution, and permission from school principals were granted. Participants’ informed consents and parents’/guardians’ written consents were also collected before the survey. Subsequently, the respective heads of school sports department and research assistants distributed the paper-based questionnaire. The official language in Singapore is English, hence, the PABSS can be directly administrated to the participants. The participants were informed that participation in the study was voluntary and their responses were kept confidential. The participants were asked to refer to the extra-curricular sport they participated in the current season when reporting their sport behaviors. They were also encouraged to respond honestly. It took about 10 minutes to complete the questionnaire.

DATA ANALYSES
As discussed in the Introduction section, testing for the factorial invariance of an instrument allows researchers to examine whether components of the measurement model are invariant or equivalent across particular groups [17]. In other words, evidence of measurement invariance provides a valid basis for making group comparisons. To examine the three aspects of measurement invariance (i.e., configural, metric and scalar variance) of the PABSS across groups (male vs. female; secondary school vs. junior college), we followed the procedure proposed by Byrne using EQS 6.1 [17].

Specifically, before invariance testing, we determined the measurement model (first-order four-factor structure) for each group from the perspectives of both parsimony and substantive meaningfulness [18]. Building upon the previous stage, configural invariance was examined by estimating the measurement model (i.e., a least-constrained model with the same parameters across groups) across the two groups simultaneously. The fit of the simultaneously estimated model (i.e., configural invariance model) can be used as the baseline value to compare with subsequent specified models [18]. Next, we tested the metric invariance by comparing the fit of metric invariance model (i.e., the model with all of its factor loadings being constrained) with the fit of the configural invariance model. Finally, we tested the scalar invariance by comparing the fit of the scalar invariance model (i.e., the model with all of its factor loadings and intercepts being constrained) with the configural invariance model. The three aspects of measurement invariance provide increasing evidence of group invariance (each subsequent step provides higher evidence) [17].

Maximum Likelihood estimation was chosen as most of the items had skewness and kurtosis values between +1 and −1 (one item had a skewness value of 1.20 and another one with a skewness value of 1.14), and the distribution of the data showed multivariate normality (Mardia’s coefficient was 42.75 and the Normalized estimate was 17.26) [21, 22]. Model fit was evaluated using comparative fit index (CFI), root mean square errors of approximation (RMSEA), and standardized root mean squared residual (SRMR). For a CFI value, the higher the better model fit. A smaller value for RMSEA/SRMR indicates a better
model fit. There were ranges of cut-off criteria associated with these CFA fit indices and some disparities on setting cut-off values across researchers. For example, against the use of traditional cut-off values (CFI ≥ .90, RMSEA ≤ .08, SRMR ≤ .08), Hu and Bentler [23] recommended a more rigorous cut-off criteria (CFI ≥ .95, RMSEA ≤ .06, SRMR ≤ .05). Marsh et al. [24] warned that these more stringent cut-off values should not be used as the golden rule to assess model fit. As such, the traditional criteria were applied as indicators of acceptable fit and Hu and Bentler’s [23] criteria were adopted as evidence of good fit. The changes of CFI and χ² values can be used for a model comparison. However, we did not use the Δχ² test as researchers have argued that the Δχ² value is sensitive to sample size [17]. Consequently, the ΔCFI test was used in the current study and the difference in value greater than .01 (ΔCFI ≥ .01) indicate a significant difference between two models [17].

Cronbach’s α was used to determine internal reliability of factors and α values greater than .70 indicate adequate internal consistency [25]. Convergent validity refers to the common proportion of variance shared by the items within a specific factor and correlations with other purported measures of a similar factor [26, 27]. The average variance extracted (AVE), the mean variance extracted for the items loading on a factor, was used to examine convergent validity of the scale. An AVE value higher than .50 within a factor indicates adequate convergent validity [26].

Discriminant validity concerns the extent to which a factor is truly distinct from other factors and a factor with high discriminant validity should be unique and captures some phenomena other factors do not [25]. Discriminant validity was examined by assessing the factor correlations among the four factors and a correlation lower than .85 indicates adequate discriminant validity [20, 28].

RESULTS

The results of confirmatory factor analysis showed that the data (entire sample) fit the four-factor first-order model adequately, χ² (164) = 439.56, CFI = .943, SRMR = .052, RMSEA = .054. Item factor loadings ranged from .55 to .84 (M = .71). Results of the descriptive statistics, internal reliability, convergent validity, and factor correlations of the scale are presented in Table 1. Athletes reported “rarely” to “sometimes” in antisocial behaviors and “sometimes” in prosocial behaviors. The α values for the four factors ranged from .79 to .88, indicating acceptable to good internal reliability. All the AVE values were close to or higher than .50, indicating adequate convergent validity of the PABSS. Adequate discriminant validity of the PABSS was also found as all the factor correlations ranged from .04 to .71.

Table 1. Descriptive statistics, internal reliability, convergent validity, and factor correlations

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>AVE</th>
<th>Mean</th>
<th>SD</th>
<th>AT</th>
<th>PT</th>
<th>AO</th>
<th>PO</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>.82</td>
<td>.49</td>
<td>1.83</td>
<td>.63</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PT</td>
<td>.85</td>
<td>.58</td>
<td>3.76</td>
<td>.70</td>
<td>.04</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AO</td>
<td>.88</td>
<td>.48</td>
<td>1.76</td>
<td>.62</td>
<td>.71**</td>
<td>.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PO</td>
<td>.79</td>
<td>.57</td>
<td>2.69</td>
<td>.87</td>
<td>.23**</td>
<td>.51**</td>
<td>.25**</td>
<td>–</td>
</tr>
</tbody>
</table>

Note. AVE = Average Variance Extracted; SD = Standard Deviation; AT = Antisocial Behaviour toward Teammates; PT = Prosocial Behaviour toward Teammates; AO = Antisocial Behaviour toward Opponents; PO = Prosocial Behaviour toward Opponent.

**p < .01.
The results of the invariance analyses are summarized in Table 2. Regarding the gender invariance, adequate to good fit was achieved for all models and there was no substantial difference in the CFI across all the models (i.e., models 3-5; all ΔCFIs < .01). To examine whether males and females differed on the four factor means, factor intercepts of males were fixed to zero for each factor (i.e., male acted as the reference group). The results showed that females have higher scores on antisocial behavior towards teammates (.06, p < .05), prosocial behavior towards teammates (.13, p < .05), antisocial behavior towards opponents (.05, p < .05), and prosocial behavior towards opponents (.07, p < .05). In terms of the findings of school level differences, adequate to good fit was achieved for all models and there was no group invariance across all the models (i.e., models 8-10, all ΔCFIs < .01). In an effort to evaluate whether junior college students and secondary school students differed on the four factor means, junior college level was used as the reference group. Significant mean differences were not observed on all the four factor means between groups (all ps > .05).

Table 2. Fit indices for multisample gender and school analyses

<table>
<thead>
<tr>
<th>Model</th>
<th>χ2 (df)</th>
<th>CFI</th>
<th>SRMR</th>
<th>RMSEA (90% CI)</th>
<th>Model comparison</th>
<th>ΔCFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1(Male)</td>
<td>274.36 (164)</td>
<td>.958</td>
<td>.052</td>
<td>.048 (.038, .057)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Model 2 (Female)</td>
<td>307.31 (164)</td>
<td>.918</td>
<td>.059</td>
<td>.058 (.048, .068)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Model 3 (Configural Invariance)</td>
<td>581.67 (328)</td>
<td>.942</td>
<td>.055</td>
<td>.053 (.046, .060)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Model 4 (Metric Invariance)</td>
<td>602.63 (344)</td>
<td>.941</td>
<td>.060</td>
<td>.052 (.045, .059)</td>
<td>3 vs. 4</td>
<td>-.001</td>
</tr>
<tr>
<td>Model 5 (Scalar Invariance)</td>
<td>667.87 (360)</td>
<td>.945</td>
<td>.061</td>
<td>.053 (.046, .059)</td>
<td>3 vs. 5</td>
<td>.003</td>
</tr>
<tr>
<td>School</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 6 (Secondary School)</td>
<td>314.45 (164)</td>
<td>.952</td>
<td>.052</td>
<td>.049 (.040, .057)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Model 7 (Junior College)</td>
<td>286.76 (164)</td>
<td>.930</td>
<td>.065</td>
<td>.064 (.051, .076)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Model 8 (Configural Invariance)</td>
<td>605.96 (328)</td>
<td>.943</td>
<td>.060</td>
<td>.054 (.048, .061)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Model 9 (Metric Invariance)</td>
<td>616.12 (344)</td>
<td>.944</td>
<td>.062</td>
<td>.053 (.046, .059)</td>
<td>8 vs. 9</td>
<td>.001</td>
</tr>
<tr>
<td>Model 10 (Scalar Invariance)</td>
<td>650.27 (360)</td>
<td>.944</td>
<td>.062</td>
<td>.053 (.046, .059)</td>
<td>8 vs. 10</td>
<td>.001</td>
</tr>
</tbody>
</table>

Note. CFI = Comparative Fit Index; SRMR = Standardized Root Mean Squared Residual; RMSEA = Root Mean Square Error of Approximation; CI = Confidence Interval.

**DISCUSSION**

There has been a growing interest on sport morality research since the last decade [13], as sports have been suggested as a conducive platform to teach life skills, given their prominence as a social activity and the great interest from youth for sports participation [5]. An important progress in this area is the development of the PABSS which enables researchers to examine a wide range of morally relevant behaviors in sport settings [13]. Given that PABSS was preliminarily developed in the United Kingdom, its reliability and validity to be used in other cultures and contexts were not investigated. Schutz and Park [29] also stated that validating a measure is an ongoing process to determine what the instrument is designed to measure and accurate conclusion can be drawn about the meaning of the scores from the scale. Therefore, the present study examined the measurement invariance and other psychometric properties of the PABSS in Singapore using school team athletes. Results showed that psychometric properties of the scale were adequately supported.

One significant contribution of the present study was that the factorial validity, internal
reliability, convergent validity and discriminant validity of the PABSS were established in the Singaporean’s context. The internal reliability of the four factors (αs = .79 to .88) were comparable or even higher than the study conducted by Kavussanu and Boardley (αs = .73 to .86) [13]. Magnitudes of factor correlations found in the current study (|0.4 to 0.71|) were also similar to the previous validation study (|0.4 to 0.74|) [13]. As such, researchers should be confident to use the instrument to examine prosocial and antisocial behaviors towards teammates and opponents for Singapore school athletes.

Another important contribution of the present study was the examination of the invariance of the PABSS across gender and school groups. Our results showed that configural, metric and scalar invariance of the scale was established, indicating both males and females athletes in different school levels perceived the observed items and their underlying constructs in the same way. Thus, the PABSS appears to be appropriate for use in research involving school athletes of either gender and from any of the two school levels tested.

Although the present study examined the psychometric properties of the PABSS in the Singapore context and was the first to evaluate its measurement invariance across school level, there are a few limitations of our study which should be mentioned. Firstly, data were collected at different time points of the season (e.g., pre-season, in-season) for different teams which may affect our findings. The reason is because the focus and content of the training for pre- and in-season periods is different (e.g., fundamental skills vs. competitions), which might affect how athletes report moral behaviors based on disparate periods of time. Nonetheless, it is difficult to administer the survey within the same period of the season given there were a lot of teams and schools involved in the present study. Longitudinal designs might be used in future to monitor athletes’ moral behavioral to examine whether there is a large variation of the data within participants across the season. Secondly, participants of the study were mainly school athletes competing in five team sports (i.e., hockey, soccer, netball, rugby and basketball), thus the instrument might not be appropriate for other team sports until further evidence of reliability and validity is established. Lastly, although our participants were strongly encouraged to respond honestly to the PABSS, a social desirability scale was not used in the current study to assess if the participants have a tendency to over-report positive sport behaviors or under-report negative sport behaviors.

IMPLICATIONS AND FUTURE RESEARCH DIRECTIONS
Our results suggest that the PABSS can be used to monitor Singapore school athletes’ (13 to 17 years old) prosocial and antisocial tendencies in team sport. This is an important implication for school leaders, policies makers and practitioners that the PABSS has the potential to be used for evaluating the desired outcomes of sport programs in achieving moral behaviors and holistic development of the students. We acknowledged the limitations of our study and would like to offer some suggestions for future studies. Firstly, future research on equivalence of the scale across different cultures and contexts is warranted to further strengthen the instrument, as well as advance our knowledge in this line of research.

Secondly, future work needs to address the sensitivity of the scale. For example, to examine whether participants change their responses over time either in a natural or experimental intervention environment that aims to enhance positive behaviors and reduce negative behaviors in sport participation.

CONCLUSION
The current study contributed to the literature by providing statistical evidence to support the psychometric properties of the PABSS using school team athletes in the Asian context. Our
findings suggest that scholars or practitioners can use the PABSS to measure prosocial and antisocial sporting behaviors in Singapore school athletes for different age groups.

ACKNOWLEDGEMENTS

Authors would like to thank reviewers for their formative comments on the early draft of this manuscript.

REFERENCES