Validating the Scale of Perception of Sex Abuse in Youth Sports (SPSAYS): Confirmatory Factor Analysis

Thomas Baker III, University of Georgia
Kevin Byon, University of Georgia

Legal aspects

To date, research on child sex abuse in sport remains very limited (Hartill, 2009). This is particularly true in regards to studies aimed at determining the prevalence of sex abuse in youth sports. Prevalence studies have been conducted in Australia (Leahy, Pretty & Tenenbaum, 2002), Canada (Kirby and Graves, 1996; Kirby, Graves & Hankivsky, 2000), and Norway (Fasting, Brackenridge & Sundgot-Borgen, 2000). However, very limited attention has been paid to analyzing the perceived risk of sex abuse in youth sports (Baker & Byon, 2011). Youth sport organizations need to be equipped with the data necessary to develop policies directed at protecting athletes from sexual abuse. One probable reason for the lack of empirical studies related to this area may be due to the absence of a valid instrument that specifically deals with the risk perception toward pedophilic coaches in youth sports. In response, Baker and Byon (2011) developed a Scale of Perception of Sex Abuse in Youth Sports (SPSAYS) measuring risk perception toward unfit coaches in youth sports. The results of an exploratory factor analysis (EFA) yielded 15 items under four factors labeled as Management (5 items), Prevalence (4 items), Influence (3 items), and Likelihood (3 items). In addition, the scale showed a preliminary reliability scores. Further examination via confirmatory factor analysis (CFA) to assess the measurement properties of the four factors is warranted. Therefore, the purpose of this study was to examine psychometric properties (i.e., model fit, convergent validity, reliability, and nomological validity) of the SPSAYS using a CFA. Results from this study could enhance the factorial validity of the scale, enabling youth sport organization managers to confidently use the SPSYAS as a tool to assess the perceived prevalence of sex abuse committed by pedophilic youth sport coaches, and how that risk perception influences future behavior (participation).

Using the purposive sampling method, participants with experience in youth sports were reached from classes at a large university in the southeastern part of the United States. As a result, we collected a total of 230 participants. Due to incomplete questionnaires, we eliminated 21 surveys, leaving the final sample of 209. The sample size (N = 209) of the study 2 also met the Watson and Gore’s (2006) suggestion of minimum sample size of 200. The SPSYAS was adopted that included four factors: Management (5 items), Prevalence (4 items), Influence (3 items), and Likelihood (3 items). The items were assessed using a 5-point scale, ranging from 1 Items were assessed using a 5-point scale. Two items measuring behavioral intentions, namely Participation Consideration and Drop-Out were used to test nomological validity of the model. Procedures in the AMOS version 18 (SPSS, 2010) were utilized to conduct a CFA for the proposed four factor model. To measure model fit, several fit indexes were adopted, including \( \chi^2/2 \), \( \chi^2/df \), RMSEA, SRMR, and CFI (Hu & Bentler, 1999). To measure the reliability of the model, construct reliability (CR) and average variance extracted (AVE) were calculated using the formulae suggested by Hair et al. (2010). Construct validity was ensured via convergent validity and nomological validity (Anderson & Gerbing, 1988; Hair et al., 2010).

The four-factor model was subject to a CFA via ML estimation method (Hair et al., 2010). Goodness of fit indexes revealed that the four-factor measurement model fit the data reasonably well (\( \chi^2(84) = 187.98; \chi^2/df = 2.24; CFI = .91; \) RMSEA = .077; and SRMR = .079). Researchers argued that a model that has a good model fit only mean that the model is plausible (Kline, 2010). Thus, three alternative models were developed for a model comparison, along with the initially hypothesized correlated four-factor model (model 1). The three alternative models were: (a) a unidimensional model for which all 15 items were forced to load on a single factor (model 2), (b) a four-factor uncorrelated model for which all four factors were forced not to be correlated (model 3), and (c) a second-order model for which a global construct representing all four first-order factor was hypothesized (model 4). A chi-square difference test was employed to statistically compare the models (Hair et al., 2010). As a result, the models 2 and 3 were found to be inferior to the models 1 and 4. Both the model 1 (\( \chi^2(84) = 187.98 \)) and the model 4 (\( \chi^2(86) = 191.64 \); \( \chi^2/df = 2.23; CFI = .91; \) RMSEA = .077; and SRMR = .082) fit the data well. A chi-square difference test was not significant (\( \chi^2 \) difference (2) = 3.66, \( p > .05 \)), suggesting that the second-order model be retained because the model was more parsimonious.

The second-order measurement model showed sound psychometric properties as evidenced by all statistically significant loadings in the expected direction, supporting for convergent validity (Anderson & Gerbing, 1982). The
reliability of the factors was evaluated by Cronbach’s alpha, CR, and AVE. Cronbach’s alpha values for the model indicated that all factors were above the acceptable threshold (i.e., greater than .60) suggested by Kline (2010), ranging from .65 (Likelihood) to .87 (Management). Except for Likelihood (.65), the CR values for the model were all above the recommended threshold (Hair et al., 2010), ranging from .73 (Influence) to .88 (Management). AVE values for the model indicated that all factors except for Prevalence (.46) and Likelihood (.39), were above the acceptable threshold (i.e., greater than .50) suggested by Fornell and Larcker (1982), ranging from .50 (Influence) to .59 (Management). Based on the overall information of reliability tests, the second-order model was deemed reliable, pending further validation.

A structural equation modeling was conducted to examine the predictability of the SPSAYS to behavioral intentions. The two-step model (Anderson & Gerbing, 1988) was employed to test the predictive validity. The overall model fit was reasonably well ($\chi^2 = 309.62, p < .001, \chi^2/df = 2.38, \text{CFI} = .88, \text{RMSEA} = .082, \text{and SRMR} = .085$). Having a satisfied model fit, it was appropriate to proceed with a SEM analysis. The SEM test revealed that SPSAYS was found to be positively related ($\beta = .82, p < .001$) to Participation Consideration and Drop-Out ($\beta = .45, p < .001$). The model predicted 67% and 20% of the variances in Participation Consideration and Drop-Out, respectively, indicating that the measurement model showed high predictive validity.

The revalidation of the SPSAYS revealed that the SPSAYS is of good measurement properties, pending further validation, to measure the perceived risk associated with unfit coaches in youth sports. Practical implications as to how this scale could be utilized by youth sport organizations are further discussed in this presentation.