Confirmatory Factor Analysis of the Athlete Sleep Behavior Questionnaire

Author Name and Credentials: Emilie N. Miley, DAT, ATC
Affiliation: Emilie N. Miley is an Assistant Professor at Charleston Southern University; and is also a Ph.D. student in Education, Department of Movement Sciences, University of Idaho, Moscow, ID
Address: 9200 University Blvd, Charleston, South Carolina 29406
Phone: (843) 863-7414
E-mail: emiley@csuniv.edu

Author Name and Credentials: Bethany L. Hansberger, DAT, ATC
Affiliation: Bethany Hansberger is an Athletic Trainer at Towson University, Towson, MD
Address: 8000 York Road, Towson, MD 21252
Phone: 410-704-4356
E-mail: b_hansberger@yahoo.com

Author Name and Credentials: Madeline Casanova, Ph.D.
Affiliation: Madeline Casanova is Postdoctoral Fellow in the WWAMI Medical Education Program, University of Idaho, Moscow, ID
Address: 875 Perimeter Drive, Moscow, Idaho 83844-2401
Phone: (208) 885-1194
E-mail: mcasanova@uidaho.edu

Author Name and Credentials: Russell T. Baker Ph.D., DAT, ATC
Affiliation: Russell Baker is the Associate Director of Medical Research with the WWAMI Medical Education Program and a Clinical Associate Professor in the Athletic Training Programs, University of Idaho, Moscow, ID
Address: 875 Perimeter Drive, Moscow, Idaho 83844-2401
Phone: (208) 885-2065
E-mail: russellb@uidaho.edu

Author Name and Credentials: Michael A. Pickering, PhD
Affiliation: Trident University
Address: 2200 East Germann Road, Suite 150
Phone: 208.874-7154
E-mail: michael.pickering@trident.edu
Readers should keep in mind that the in-production articles posted in this section may undergo changes in the content and presentation before they appear in forthcoming issues. We recommend regular visits to the site to ensure access to the most current version of the article. Please contact the JAT office (jat@slu.edu) with any questions.
Confirmatory Factor Analysis of the Athlete Sleep Behavior Questionnaire
Abstract

Context: Sleep has long been understood as an essential component for overall well-being, significantly impacting physical health, cognitive functioning, mental health, and quality of life. Currently, the Athlete Sleep Behavior Questionnaire (ASBQ) is the only known instrument designed to measure sleep behaviors in the athletic population. However, the psychometric properties of the scale in a collegiate student-athlete and dance population have not been established.

Objective: To assess model fit of the ASBQ using a sample of collegiate student-athletes and competitive dancers.

Design: Observational study.

Setting: Twelve colleges and universities.

Patients or Other Participants: Student-athletes and dancers competing at the collegiate level.

Main Outcome Measure(s): A confirmatory factor analysis (CFA) was conducted to assess the factor structure of the ASBQ. Principal component analysis (PCA) extraction and covariance modeling analyses were performed to identify an alternate model. Multi-group invariance testing was performed on the alternate model to identify if group differences existed between sex, athletic activity, injury status, and division of competition.

Results: The CFA on the ASBQ indicated the model did not meet recommended model fit indices. An alternate three-factor, nine-item model with improved fit was identified; however, the scale structure was not consistently supported during multi-group invariance testing procedures.

Conclusions: The original three factor, 18-item ASBQ was not supported for use with collegiate athletes in our study. The alternate ASBQ was substantially improved; however, more research
should be completed to ensure the nine-item instrument accurately captures all dimensions of sleep behavior relevant for collegiate athletes.

**Key Words:** measurement, covariance modeling, athletic population

**Abstract Word Count:** 267

**Body of Manuscript Word Count:** 4125

**Key Points:**

1. Sleep is multifactorial in nature and is an important component for athletic trainers to consider in the treatment of collegiate athletes.
2. The ASBQ did not meet contemporary model fit recommendations.
3. Clinicians should use caution when using the ASBQ (original or modified) given the model fit concerns and issues with instrument design.
Sleep has long been understood as an essential component for overall well-being.\textsuperscript{1-3} Specifically, sleep significantly impacts physical health, cognitive function, mental health, and quality of life.\textsuperscript{1-3} Sleep has been proposed as a multifactorial construct that is affected by external (e.g., anxiety, noise, the need to use the bathroom, early event times) and internal (e.g., circadian rhythm) factors, which often impact each other.\textsuperscript{4,5} For example, when sleep disturbances occur due to external factors, changes to the internal factors, such as the circadian rhythm, may result. Internal and external influences affect physical and cognitive functioning, which can subsequently impact sports performance in athletes.\textsuperscript{6,7} Adequate sleep quality and quantity are important for optimal performance indices (e.g., reaction time, learning, memory tasks).\textsuperscript{8}

Researchers have indicated that athletes have reduced sleep quality compared to non-athletes. Most athletes report sleeping under the recommended target hours (i.e., eight hours of sleep) the night before competition, and approximately 70\% of athletes have reported experiencing problematic or poor sleeping patterns prior to competition compared to their normal routine.\textsuperscript{4} Poor sleeping patterns increase fatigue and tension, which are negatively correlated with precompetitive relative sleep quality and total sleep time.\textsuperscript{4} The poor sleep quality patterns identified may also have a detrimental effect on student-athlete success.\textsuperscript{9}

Researchers investigating sleep quality and overall well-being for athletes have indicated athletes who experienced acute partial (i.e., average of 2.5 hours) sleep deprivation report an increase in negative mood states (i.e., depression, tension, confusion, fatigue and anger) and decreased vigor (i.e., physical strength, good health, energy).\textsuperscript{10,11} Additionally, sleep quality and quantity are reported to impact an athlete’s ability to recover following activity;\textsuperscript{12} good sleep patterns (i.e., quality and quantity) are considered to be an important recovery method.\textsuperscript{9,13} Long-term negative changes in mood, possibly due to poor sleep patterns, may be related to increased
injury risk\textsuperscript{11} and high player workloads; continual poor sleep history may further increase injury risk due to elevated levels of chronic fatigue.\textsuperscript{13,14}

Therefore, sleep behavior may be an important factor to measure when assessing injury risk or recovery.\textsuperscript{13} Many self-report instruments have been developed to evaluate sleep in a general population; however, few self-report instruments exist specifically designed for the athletic population. Two questionnaires created to assess sleep in the athletic population are the Athlete Sleep Screening Questionnaire (ASSQ) and the Athlete Sleep Behavior Questionnaire (ASBQ).\textsuperscript{12} The ASSQ assesses clinical sleep difficulties in athletes and was designed to identify at-risk individuals who many need a referral to a sleep specialist.\textsuperscript{12} Although the ASSQ assesses six factors related to sleep difficulty (i.e., total sleep time, insomnia, sleep quality, sleep chronotype, sleep disordered breathing, and travel disturbance), it does not gather information on sleep behavior practices.\textsuperscript{12} Thus, researchers developed the ASBQ to assess sleep in elite international athletes using a combination of newly developed items and items drawn from previously validated questionnaires (i.e., the Sleep Hygiene Index and the International Classification of Sleep Disorders).\textsuperscript{12} The ASBQ has been proposed as an 18-item reflective measurement instrument designed for quick and efficient administration to identify sleep behavior practices in a competitive athlete population.\textsuperscript{12,14} Researchers used principal component analysis (PCA) and identified a 3 construct solution (i.e., routine, behavioral, and sport-related components), thus supporting the use of the scale.\textsuperscript{12}

Despite positive initial findings, several limitations or concerns support the need for further psychometric examination before using the ASBQ in clinical practice. For example, initial scale development of the ASBQ utilized a sample of elite athletes and non-athletes,\textsuperscript{14} without inclusion of other traditional competitive athletic categories (e.g., collegiate athletes) or...
a more general physically active population (i.e., recreational athletes), thus limiting the ability to ensure scale applicability and utilization across different levels within the athletic population. Other scale design concerns, such as the use of double-barreled questions (i.e., items that ask about more than one issue) or poor internal consistency (i.e., Cronbach’s α > .70), are present in the scale. Additionally, scale structure has not been confirmed in subsequent work; researchers who translated the scale into the Turkish language identified a four factor solution which did not match the originally proposed three factor model. Finally, recommended scale psychometric evaluation procedures, including confirmatory factor analysis (CFA) and invariance testing, have not been identified in the literature, and these analyses are necessary to establish the measurement properties of a scale to recommend its use in research and practice.

The lack of published CFA and invariance results, combined with the inconsistent factor structure, warrants further investigation of the psychometric properties of the ASBQ in larger, more diverse collegiate athletic samples. Thus, the primary purpose of the study was to assess the proposed reflective measurement model using CFA to assess model fit and psychometric properties of the original ASBQ scale in a sample of collegiate athletes. The secondary purposes of the study, if the ASBQ did not meet recommended model fit criteria were to (1) perform PCA and covariance modeling to identify and assess alternate models and (2) if the alternate model fit criteria were met, perform multi-group (i.e., males vs. females, collegiate dancers vs. other student-athletes, healthy vs. injured, and division of competition) invariance testing of the scale.

METHODS

Participants

The study was identified to be exempt by the University Institutional Review Board and informed consent was obtained from all participants prior to data collection. A convenience
sample of athletic trainers and dance faculty were used to recruit healthy and injured individuals from 12 different colleges and universities in the United States; participants were recruited from different collegiate levels (e.g., NCAA Division I, NAIA). Student-athletes and competitive dancers were grouped by health status (i.e., healthy, acute-, sub-acute-, persistent injury; Table 1), sex (male, female), and sport activity (e.g., dance, team sport; Table 2).

**Instrumentation**

A survey packet that included the ASBQ and a demographic questionnaire was created in identical paper and electronic formats. Demographic information collected included age, sport, division of competition (e.g., NCAA Division 1, NAIA), and self-reported injury category (i.e., healthy, acute, subacute or persistent). The electronic version of the packet was created using Qualtrics software (Qualtrics, LLC, Provo, UT). Responses from the paper packet were input into Qualtrics by the participating athletic trainer or were mailed back to the research team to be input into the system. The electronic version of the packet was completed by the participant using an electronic link to the survey.

**Athletic Sleep Behavior Questionnaire.** The ASBQ is a three factor, 18-item instrument designed to assess sleep patterns related to routine, behavior, and sport in elite athletes. Participants are asked how frequently they engage in particular sleep and sport behaviors over the past month. Participants rate each question using a 5-point Likert scale ranging from 1 (never) to 5 (always). Items are summed to create a global ASBQ score, ranging from 18 to 90, with higher scores indicative of poorer sleeping behaviors. The ASBQ is reported to have excellent reliability (ICC = 0.87, r = 0.88, CV = 6.4%); however, the reported Cronbach's α for internal consistency was 0.63.

**Data Analysis**
Survey responses were downloaded and analyzed using the Statistical Package for Social Sciences Version 25 (SPSS; IBM Corp, Armonk, NY) and the Analysis of Moment Structure software Version 25 (AMOS; SPSS Inc, Chicago, IL). Missing responses for each survey item were calculated for each respondent and individuals missing more than 10% of the ASBQ items (i.e., two or more questions) were removed from the data set. Individuals missing less than 10% of the data were retained and respective missing items were replaced with the rounded mean score for analysis. Participants missing demographic data were not excluded from the analysis. Skewness and kurtosis values, as well as histograms were assessed for normality. Data were analyzed for univariate outliers using z-scores and multivariate outliers were identified using Mahalanobis distance at \( P < .01 \).

**Confirmatory Factor Analysis.** We conducted the CFA using maximum likelihood estimation in AMOS software (IBM Corp., Armonk, NY) across the sample. We conducted two additional CFAs for sub-groups (i.e., in-season traditional collegiate athletes and dancers) of the sample because the original ASBQ was designed for elite athletes and may be more commonly used while athletes are participating in sport (i.e., in-season training). The proposed three factor structure was assessed using \textit{a priori} model fit indices. The goodness-of-fit indices computed were the Comparative Fit Index (CFI; \( \geq .95 \)), Goodness of Fit Index (GFI; \( \geq .95 \)), Tucker-Lewis Index (TLI; \( \geq .95 \)), Bollen’s Incremental Fit Index (IFI; \( \geq .95 \)), and Root Mean Square Error of Approximation (RMSEA; \( \leq .05 \)). Model fit was also assessed for localized areas of strain including statistical significance of parameter estimates.

**Alternate Model Generation.** A PCA with varimax rotation was conducted using SPSS to identify a more parsimonious model if criteria were not met for the original CFA. A PCA with varimax rotation (i.e., orthogonal rotation) was used to replicate the original analysis procedures.
utilized to establish the scale.\textsuperscript{12} Items were removed one at a time and the PCA was repeated until a solution that met recommendations was identified. Item removal was guided by the statistical (e.g., low loadings \(< .40\), high cross-loadings \(\geq .30\) with other items, etc.),\textsuperscript{16} theoretical (e.g., does the item make sense with the other items that have factored), and survey design (e.g., double-barreled, etc.) concerns identified across items at each step of the iterative PCA extraction process.\textsuperscript{12,16,23} The PCA process was repeated by two separate investigators to ensure the same final solution was reached. Kaiser-Meyer-Olkin Measure of Sampling Adequacy (\(\geq .70\)) and Bartlett’s test of sphericity \(P < .001\) were assessed for violations\textsuperscript{16} and the extraction was fixed to retain three components as specified by previous research.\textsuperscript{12}

The alternate model identified during the PCA was then assessed in a covariance model using AMOS software. The same criteria used to assess model fit for the CFA were also used to assess model fit for the covariance model. Additionally, parameter estimates, and modification indices were assessed for local strain identification. Additionally, bivariate correlation analyses using the cumulative (i.e., total) and construct scores from the 18-item ASBQ scale and the cumulative and construct scores on the newly proposed modified 9-item ASBQ were conducted; \textit{a priori} thresholds (\(<0.1, \text{trivial}; 0.1-0.3, \text{small}; 0.3-0.5, \text{moderate}; 0.5-0.7, \text{large}; 0.7-0.9, \text{very large}; \text{and 0.9-1.0, almost perfect}\) were used to assess the magnitude of correlation between scales and constructs.\textsuperscript{12}

\textbf{Multi-group Invariance Testing.} Multi-group CFA invariance testing (i.e., configural, metric, and scalar) was planned for the original model or the alternate model depending on which model met contemporary recommendations. Multi-group invariance testing was performed using AMOS software to assess model fit across different sub-groups: sex (i.e., males or females), sport type (i.e., competitive dancer or traditional student-athlete), self-reported injury status (i.e.,
healthy or injured), and level of competition (i.e., NCAA Division I or non-Division I athlete). The CFI difference test (CFI\textsubscript{DIFF}) and the chi-square difference test ($\chi^2$\textsubscript{DIFF}) were used to assess model fit, with a P-value cut-off of 0.01.\textsuperscript{16,19,25} More emphasis was placed on the CFI\textsubscript{DIFF} test due to the sensitivity of the $\chi^2$\textsubscript{DIFF} test to sample size;\textsuperscript{19,25} if a model exceeded the $\chi^2$\textsubscript{DIFF} test but met the CFI\textsubscript{DIFF} test, invariance testing would continue.

RESULTS

A total of 605 individuals completed the survey; nine individuals (1.5%) were missing responses to more than 10% of the items and were removed from the dataset. A total of 49 (8.1%) participants reported scores that were identified as univariate (z scores $\geq$ 3.4) or multivariate (Mahalanobis distance $\geq$ 33) outliers and were subsequently removed from the data set.\textsuperscript{21,25} These participants included both sexes, three injury categories (i.e., healthy, acute, persistent), and various divisions (e.g., Division I, NAIA). Removing these participants from the sample resulted in a normal data distribution for both individual items and summary indexes of the items. A total of 556 (91.9%) participants were retained for analysis (104 males, 452 females; mean age = 19.84 $\pm$ 1.62 years, age range = 16 - 32 years; Table 2). Most respondents ($n$ = 325, 58.5%) participated at the NCAA Division I athletics level and were classified as healthy ($n$ = 412, 74.1%; Table 2).

Confirmatory Factor Analysis

The CFA of the three factor, 18-item ASBQ did not meet recommended model fit indices to the full sample data ($\chi^2$ = 600.90, $P$ < .001, CFI = 0.586, TLI = 0.520, GFI = 0.888, IFI = 0.584, RMSEA = 0.080; Figure 1); the model also failed to meet recommendations for in-season athletes ($\chi^2$ =216.09, $P$ < .001, CFI = 0.562, TLI = 0.492, IFI = 0.605, RMSEA = 0.087) and dancers ($\chi^2$ =372.29, $P$ < .001, CFI = 0.602, TLI = 0.539, IFI = 0.616, RMSEA = 0.078).
Moreover, there were several potential fit concerns within the full sample: non-significant (P > 0.05) factor loadings, low item loadings (< 0.40), and standardized path coefficients >1. The latent variable correlations were moderate (Routine and Behavioral components $r = 0.45$, Routine and Sport $r = 0.49$, Behavioral and Sport $r = 1.09$; Figure 1). Additionally, modification indices indicated numerous meaningful cross-loadings and error correlation specifications were present.

**Alternate Model Generation**

The initial fixed three-component PCA solution included items with low loadings (< .40) and one item with a high cross-loading (Table 3). In total, nine items with low loadings, high cross-loadings, or high inter-item correlations were removed throughout subsequent analyses during the PCA procedures. The resulting three-component nine-item solution contained items with loadings >.49 and without substantial cross-loadings. The solution accounted for 23.4% of the variance with Cronbach’s $\alpha$ ranging from 0.47 - 0.52 (Table 4). Total scores on the 9-item modified ASBQ were strongly correlated ($r = .850; R^2 = .722$) with the scores for the 18-item ASBQ. Strong relationships were found between the Behavior ($r = .635, R^2 = .403$) and Routine ($r = .643; R^2 = .413$) constructs of the modified 9-item ASBQ and the 18-item ASBQ, while a small relationship was found between the Sport constructs ($r = .134; R^2 = .018$).

**Covariance Model Refined ASBQ**

Covariance modeling of the refined three-factor, nine-item scale had improved model fit with goodness of fit-indices meeting nearly all criteria ($\chi^2 = 43.018, P = .010, CFI = 0.951, TLI = 0.926, GFI = 0.983, IFI = 0.952, RMSEA = 0.038$; Figure 2). The latent variable correlations were low (Routine and Behavioral components $r = 0.25$ [$R^2 = 0.06$], Routine and Sport $r = 0.28$ [$R^2 = 0.08$], behavioral and sport $r = -0.03$ [$R^2 = 0.001$]; Figure 2). All factor loadings were
statistically significant, ranging from 0.21 - 0.91, and modification indices did not reveal any meaningful cross-loadings or error covariance specifications.

**Multi-group Invariance Testing Across Sub-groups**

**Sex.** All 556 individuals in the sample reported their sex (males = 104, females = 452). Individual CFAs by sex met some but not all recommended fit criteria for males (CFI= 0.944; TLI = .916; RMSEA = 0.039; Table 5) and females (CFI= 0.96; TLI = .939; RMSEA = 0.034; Table 5). The configural model (i.e., equal form) met most model fit indices (CFI = .957; $\chi^2 = 64.72$; RMSEA = .025; Table 5). The metric model (i.e., equal loadings) passed both the CFI$_{DIFF}$ test (CFI$_{DIFF}$ = .001) and the $\chi^2_{DIFF}$ test ($\chi^2 = 70.10$). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variance model passed the CFI$_{DIFF}$ test (CFI$_{DIFF}$ = .001) and the $\chi^2_{DIFF}$ test ($\chi^2 = 8.78$), indicating variances were not different between groups. The scalar model (i.e., equal intercepts) did not pass the CFI$_{DIFF}$ test (CFI$_{DIFF}$ = .014) but passed the $\chi^2_{DIFF}$ test. As such, completing the subsequent steps of the multi-group invariance testing process (i.e., testing of means) was not deemed appropriate.

**Sport type.** Of the 556 individuals in the sample, 545 (98%) reported their athletic classification (competitive dancer = 304, traditional student-athlete = 241) and were used for analysis. Individual CFAs by individual sport type indicated the model did not meet recommended fit for traditional student-athletes (CFI= 0.843; TLI = .764; RMSEA = 0.063; Table 6) but did meet recommended fit for dancers (CFI= 0.982; TLI = .972; RMSEA = 0.022; Table 6). The configural model (i.e., equal form) did not meet recommended model fit indices (CFI= 0.914; $\chi^2 = 76.74$; RMSEA = 0.066; Table 6). Therefore, completing the subsequent steps
of the multi-group invariance testing process (e.g., metric, equal latent means, etc.) was not
deemed appropriate.

**Injury status.** Of the 556 individuals in the sample, 551 (99.1%) reported their injury
status (healthy = 412, injured = 139) and were used for analysis. Individual CFAs by injury
status indicated the model met some of the recommended fit criteria for the healthy group (CFI=
0.951; TLI = .926; RMSEA = 0.039; Table 7) and all the fit criteria for the injured group (CFI=
0.991; TLI = .986; RMSEA = 0.014; Table 7). The configural model (i.e., equal form) met most
model fit indices (CFI = .958; $\chi^2 = 63.61$; RMSEA = .024; Table 8). The metric model (i.e.,
equal loadings) did not pass the CFI_DIFF test (CFI_DIFF = 0.014). Therefore, completing the
subsequent steps of the multi-group invariance (e.g., scalar, equal latent means, etc.) testing
process was not warranted.

**Level of competition.** Of the 556 individuals in the sample, 491 (88.3%) reported their
division of competition (Division 1 = 325, Lower division of competition = 166) and were used
for analysis. Individual CFAs by the level of competition indicated the models did not meet
recommended fit criteria for NCAA D1 athletes (CFI= 0.932; TLI = .897; RMSEA = 0.046;
Table 8) or the lower division (CFI= 0.867; TLI = .801; RMSEA = 0.069; Table 8) groups. The
configural model (i.e., equal form) did not meet model fit indices (CFI = .907; $\chi^2 = 83.38$;
RMSEA = .039; Table 8). As such, the multi-group invariance testing process (e.g., metric,
scalar, etc.) was not warranted.

**DISCUSSION**

The first purpose of this study was to examine the psychometric properties of the
originally proposed ASBQ in a broader athletic population (i.e., college traditional student-
athletes and competitive dancers). The secondary purpose, because model fit indices were not
met, was to use PCA and alternate model generation to identify if a modified ABSQ could be
identified from the item pool for use with the college athlete population. Contemporary
psychometric analysis methods were used to assess model fit of the ASBQ and the alternate
model to guide recommendations for use in future research and clinical practice. Our results
suggest the original ASBQ has poor psychometric properties and should not be used in college
athlete populations. The alternate model met many fit recommendations; however, due to scale
concerns and psychometric testing results that do not meet all recommended criteria, further
exploration is warranted prior to adoption in clinical practice and research given.

**Confirmatory Factor Analysis of the ASBQ**

Our CFA findings do not support the model scale structure proposed in the original
study. Model fit was poor with specific concerns related to low item loadings (< .40) and model
misspecification as evidenced by the standardized loading between the Routine and Behavioral
latent constructs being above one \( (r = 1.09) \). Further, high latent variable correlation values
(i.e., ≥ .95) indicates potential multicollinearity and a lack of unique constructs being
measured; thus, item removal or modification of the items was warranted. The instrument
may be improved by condensing the scale, rewording items, or developing new items to more
effectively measure the originally proposed dimensions. Further testing (i.e., invariance
analyses) on the original ASBQ was not supported within our sample; therefore, exploration of
alternate models was conducted. Our results failed to support the originally proposed three-
factor, 18-item ASBQ instrument using a collegiate athletic population or in sub-groups of
physically active individuals who were actively engaged in training (i.e., in-season athletes and
collegiate dancers); therefore, we do not recommend it for use in this population without
alteration. Moreover, upon further review of the items identified within the original ASBQ,
most of items appeared to be formative (i.e., change to the indicator is associated with variation of the latent construct) and not reflective (i.e., the item reflects change of the latent construct) in nature.\textsuperscript{26,27} This finding could be indicative of model-misspecification which may be leading to poor fit of the model.\textsuperscript{26,27}

**Alternate Model Generation and Multi-group Invariance Testing**

The alternate model produced a similar three factor structure to the original ASBQ scale;\textsuperscript{12} however, model fit (e.g., low factor loadings, Cronbach’s alpha levels below recommended levels.) and instrument design concerns remain.\textsuperscript{15,19,24} Further, the factor structure of the alternate model was not consistent with the original ASBQ. For example, some of the items associated with the latent construct (i.e., Routine, Behavioral, Sport) did not factor into the originally proposed construct. For the Routine construct, four of the original items (5, 16, 17, 18) were retained in the nine-item model; only two of these items (5, 16) loaded onto the Routine construct while the other two (items 17 and 18) loaded onto the Sport construct. For the Behavioral construct, three of the original items (2, 4, 13) were retained in the nine-item model; two of these items (2, 4) loaded on the Behavioral construct while one item loaded on the Sport construct. Only two items (3, 6) from the original Sport construct were retained in the final nine-item model, however, neither loaded on the Sport construct; one item loaded on the Behavioral and the other on the Routine variable.

Despite item removal, the strong correlation ($r = .850$) between total scores on the modified ASBQ and the original ASBQ suggests the 9-item version accounted for most of the variance in participants responses found in the 18-item version and indicates the modified scale is capturing a similar theoretical measurement of sleep behaviors as the original scale. Further, the total score correlation findings suggest a similar phenomenon is being measured across the
two scales and serves as evidence of item redundancy in the 18-item model. The strong construct
correlations between the Routine ($r = .643$) and Behavioral ($r = .635$) constructs across scales
also indicates a similar phenomenon is being measured in each construct across the two scales.
The correlation between the Sport constructs was small, however, and indicates a different
phenomenon is primarily being assessed across these two constructs. The strong but not perfect
correlations for the total score between scale versions, the Routine constructs, and the Behavior
constructs are expected because of the similarity between items (i.e., the modified ASBQ
contains 9 of the original 18 items, 2 of the 3 items in the modified Routine and Behavior
constructs were retained from the original constructs). Additionally, a strong correlation between
scale versions was expected because the redundant and poor fitting items that result in
measurement error and variation in the original scale were removed. The small correlation for
the Sport constructs is also expected because none of the items are shared between the two
constructs. The correlational findings should be interpreted with caution as construct scoring has
not been recommended for the original or modified ASBQ; the necessary testing has not been
performed to establish the criterion validity of the constructs and psychometric assessment has
not been fully completed to support construct scoring.

The scale modifications (i.e., item removal) were necessary to address the previously
discussed fit (e.g., multicollinearity between the latent variables, substantial item cross-loading)
and design concerns. Item removal resulted in a substantial decrease in the correlations between
the Behavioral and Sport, Routine and Behavioral, and Routine and Sport constructs. The
improved correlational values may have reduced the likelihood of multicollinearity between the
constructs, resulting in a more parsimonious model.\textsuperscript{16,19} Our findings were confirmed in the
covariance model, which had substantially improved model fit and reduced latent variable
correlations (Figure 2). Of note, eight of the nine items (89%) removed, presented as formative indicators. Because most of the items removed were formative in nature, the removal of these items may have reduced model misspecification, which may also explain the increase in model fit statistics.

Concerns with the alternate model, however, were identified when multi-group invariance testing between sub-groups was performed. Baseline models were assessed between sex sub-groups: for males and females, model fit criteria met some but not all contemporary model fit recommendations. Group differences in variances were not noted for poor sleep behaviors found between sex. Because the scalar model did not meet contemporary model fit recommendations, further invariance testing for group means differences was not supported. The scalar model results indicated males and females do not conceptualize sleep behavior similarly using the modified ASBQ. Therefore, group mean differences between males and females on the modified scale should not be interpreted as true group differences without further testing with a larger sample is conducted to refute the multi-group invariance test findings in our study.

Multi-group invariance testing was also performed on the modified ASBQ across college-aged traditional student-athletes and competitive dancer sub-groups. For competitive dancers, model fit criteria were satisfied; however, for traditional student-athletes, model fit did not meet contemporary model fit recommendations. The failure of the configural model suggests that traditional student-athletes and competitive dancers do not conceptualize sleep behavior similarly across the modified ASBQ items. Further testing should be done with a larger sample from both groups to refute our findings. Additionally, researchers may want to rewrite or modify items in the scale that may be better suited for the different populations.
The modified ASBQ was then subjected to invariance testing across injury status (i.e., healthy vs. injured). For those who self-reported to be healthy, model fit criteria were satisfied; however, for those who were injured, model fit did not meet contemporary model fit recommendations. The configural model indicated model fit criteria were satisfied when both groups were included in the model; however, model fit indices for the metric invariance model were not met, indicating an inconsistent factor structure. Thus, the modified ASBQ may not be a psychometrically sound scale for tracking sleep in the injured population or for examining group differences between respondents who are healthy and those who are injured. Utilization of this scale in these groups (i.e., injured or healthy) is not recommended without testing in another sample of collegiate athletes. Lastly, multi-group invariance testing was also performed across two levels of competition: NCAA Division I athletes vs. lower-division athletes. The model fit did not meet contemporary model fit recommendations when each group was tested individually. The configural model was then assessed and model fit criteria were not satisfied. Potential fit and design concerns should be considered beyond the multi-group invariance results despite the alternate model having improved fit for most fit indices (i.e., CFI, IFI, GFI, and RMSEA). First, our Cronbach’s $\alpha$ values ranged from 0.47 - 0.52, which are similar to previous literature (i.e., 0.41 - 0.56); however, these values were lower than recommended values of $\geq 0.70$ and $\leq 0.89$. The low internal consistency may indicate that the content of the items are too heterogeneous or items are not relevant to the sample of individuals who responded. Second, the TLI value was below recommended values (< .95), which may be related to model misspecification from the combination of omitting cross loadings and low factor loadings (e.g., item ‘21’ with a loading of .21). Model misspecification may introduce bias by not accounting for all parameters, correlations, or other pertinent values. Best practice
recommendations for survey item development (e.g., avoiding double-barreled items and redundancy between items)\textsuperscript{15} and analysis procedures (e.g., exploratory factor analysis to allow for an oblique solution, parallel analysis to determine factor retention)\textsuperscript{16,32} may be employed to produce a more parsimonious and valid scale. This can allow for the scale to be more easily understood and consistently answered, which may result in improved model fit and more precise assessment of sleep behaviors in college athletes.

The mentioned survey design concerns and psychometric findings raise concerns for using the modified ASBQ in practice and serve as possible explanations for why the modified ASBQ did not meet multi-group invariance testing recommendations. However, other plausible explanations may help us understand the multi-group invariance testing results which indicated model fit for the modified ASBQ exceeded contemporary recommendations when tested in female respondents and dancers (who were also primarily female). The model failed to meet recommendations when tested in sub-groups more heavily dominated by males or other sport student-athletes. Group differences on other variables (e.g., GPA, academic preparation, etc.) may have impacted how each group interpreted or responded to items, which might indicate item design was not effective for a specific sub-group and resulted in the subsequent failure of the model to meet multi-group invariance testing requirements conducted by other demographic variable (e.g., sport type). For example, researchers have previously reported females tend to maintain higher grades than their male counterparts, and that non-athletes enter college with higher test scores than traditional athletes.\textsuperscript{30,31} In our study, males reported having a lower GPA (3.28) compared to females (3.57), and those who danced had a higher GPA (3.62) than traditional student-athletes (3.32; Table 2). Thus, it is possible that prior academic preparation may have influenced how respondents in our sample understood or responded to items, which
may mean that certain item characteristics (e.g., double-barreled questions, item reading level, item bias) influenced certain groups more than others resulting in increased response variation among certain sub-groups. Therefore, the modified ASBQ might be sufficient to use in certain populations, but not others without further refinement to address survey design concerns.

**Implementation in Clinical Practice and Research**

Measuring sleep patterns in the collegiate athletic population is an important component in athletic training; however, we do not recommend the use of the original version of the ASBQ in the physically active collegiate population. The modified ASBQ may be a viable alternative with certain populations because of the improved model fit for several fit indices (i.e., CFI, GFI, IFI, RMSEA); however, caution is warranted without further research is needed to confirm which subgroups and, in what contexts (e.g., repeated testing), those subgroups’ scores may be interpreted. For example, low-factor loadings, poor internal consistency, a lower than recommended TLI (< .95), and poor multi-group invariance testing raise concerns. The invariance testing results, along with the design of the items not following many recommended best practice standards, provide evidence that the ASBQ items may be biased or ineffective for measuring sleep behaviors in certain subgroups of the collegiate athlete population without further item refinement.

Additionally, other necessary steps in scale development, such as longitudinal invariance testing, or the assessment of scale responsiveness were also not conducted in this study and should be completed to inform users on how to use and interpret scale results prior to widespread adoption. Another important step for guiding use in clinical practice would be to establish the criterion validity of the modified ASBQ by correlating the scale with other established measures to better understand what is being assessed with construct and total scores.
Further, due to the multifactorial nature of sleep,\textsuperscript{4,5} it is unlikely that this modified version successfully captures sleep behavior patterns in collegiate student-athletes. Thus, it would be prudent to develop an instrument that adequately assesses the multifactorial nature of sleep behaviors in athletes, regardless of sex, collegiate division of competition, type of sport, or injury status.

**Limitations**

There are several limitations present in the current study. First, while the fit indices for the alternate model met most of the recommended standards, we were unable to assess responsiveness, test-retest reliability, or perform longitudinal invariance testing due to data collection occurring at a single time point. Secondly, although our large and diverse sample contained a similar number of participants per sub-group (i.e., competitive dancers and traditional student-athletes), there could have been more diversity in terms of sex, ethnicity, division of competition for the student-athletes, and injury types. The analysis approaches utilized in this study also perform better with large sample sizes; thus, certain multi-group invariance testing procedures should be tested again in larger and more equally represented samples. For example, the majority (74.1\%) of the participants in our study were self-reported to be healthy; thus, conducting multigroup invariance testing with larger samples of healthy and injured respondents with more equal group representation may be valuable for assessing the scale. In addition, differences in socioeconomic status, work-life balance, reading level, and other factors that were not assessed in this study may have had an impact on the way the scale was interpreted between groups and should be considered in future research. The measurement properties and theoretical design of the original ASBQ identified concerns with whether the scale adequately measures the proposed constructs; while removing items from the scale produced a
modified version with improved model fit, it is unclear if the 9-item ASBQ adequately captures
the intended measures (i.e., sleep behaviors, sleep routine, and sport). and further criterion
validity (e.g., correlating scores on the constructs of the modified scale to other
instruments/items thought to test the same) assessment is needed.

Lastly, we conducted our analysis of the ASBQ using a reflective measurement model, to
remain consistent with prior researchers; however, our evaluation of item content suggests
the ASBQ may be better served utilizing formative or mixed-model analysis. Therefore,
future researchers should consider both the individual items in the ASBQ and the best course of
analysis (e.g., reflective, formative, mixed-model) to create a scale that accurately assesses sleep
behaviors in athletes. Our purpose was to replicate the original ASBQ testing with the addition
of CFA and multi-group invariance testing procedures; however, future researchers should
consider other recommended procedures (e.g., EFA with oblique solutions, parallel analysis,
formative models) to test these items or develop new items to establish a psychometrically sound
instrument to measure the multi-factorial nature of sleep behaviors in collegiate athletes.

CONCLUSION

The CFA of the original three factor, 18-item ASBQ did not meet contemporary fit
recommendations. Alternate model generation was performed which led to the creation of a three
factor nine-item model. Although model fit was substantially improved, more research should be
done to ensure a valid reflective measurement model can accurately capture all dimensions of
sleep behavior relevant for collegiate athletes, as sleep remains a vital component of overall
well-being and performance.
REFERENCES

1. Luyster FS, Strollo PJ, Zee PC, Walsh JK. Sleep: a health imperative. *Sleep*. 2012;35(6):727-734.
2. Williamson A, Lombardi DA, Folkard S, Stutts J, Courtney TK, Connor JL. The link between fatigue and safety. *Accident Analysis & Prevention*. 2011;43(2):498-515.
3. Association NATA. Sleep Handout. [https://www.nata.org/blog/beth-sitzler/sleep-handout-now-available](https://www.nata.org/blog/beth-sitzler/sleep-handout-now-available). Published March 8, 2020.
4. Lastella M, Roach GD, Halson SL, Sargent C. Sleep/wake behaviours of elite athletes from individual and team sports. *Eur J Sport Sci*. 2015;15(2):94-100.
5. Samuels C. Sleep, recovery, and performance: the new frontier in high-performance athletics. *Neurol Clin*. 2008;26(1):169-180.
6. Drust B, Ahmed Q, Roky R. Circadian variation and soccer performance: Implications for training and match-play during Ramadan. *J Sport Sci*. 2012;30(sup1):S43-S52.
7. Lee A, Galvez JC. Jet lag in athletes. *Sports Health*. 2012;4(3):211-216.
8. O’Toole ML. Overreaching and overtraining in endurance athletes. *Overtraining in Sport*. 1998;3:18.
9. Leeder J, Glaister M, Pizzoferro K, Dawson J, Pedlar C. Sleep duration and quality in elite athletes measured using wristwatch actigraphy. *J Sport Sci*. 2012;30(6):541-545.
10. Sinnerton S, Reilly T. Effects of sleep loss and time of day in swimmers. *Biomechanics and Medicine in Swimming*. 1992:399-404.
11. Reilly T, Edwards B. Altered sleep–wake cycles and physical performance in athletes. *Physiology & Behavior*. 2007;90(2-3):274-284.
12. Driller MW, Mah CD, Halson SL. Development of the athlete sleep behavior questionnaire: A tool for identifying maladaptive sleep practices in elite athletes. *Sleep Science*. 2018;11(1):37.
13. Dennis J, Dawson B, Heasman J, Rogalski B, Robey E. Sleep patterns and injury occurrence in elite Australian footballers. *J Sci Med Sport*. 2016;19(2):113-116.
14. Johnson U. Athletes’ experiences of psychosocial risk factors preceding injury. *Qualitative Research in Sport, Exercise and Health*. 2011;3(1):99-115.
15. Dillman DA, Smyth JD, Christian LM. *Internet, phone, mail, and mixed-mode surveys: the tailored design method.* John Wiley & Sons; 2014.
16. Leech NL, Barrett KC, Morgan GA. *IBM SPSS for intermediate statistics: Use and interpretation.* Routledge; 2014.
17. Darendeli A, Diker G, Çınar Z. Athlete Sleep Behavior Questionnaire-Turkish Version: Study of Validity and Reliability. *Journal of Turkish Sleep Medicine*. 2019;6(2):43.
18. Bryant FB, Yarnold PR. Principal-components analysis and exploratory and confirmatory factor analysis. 1995.
19. Kline RB. *Principles and practice of structural equation modeling.* Guilford publications; 2015.
20. Baker RT, Burton D, Pickering MA, Start A. Confirmatory factor analysis of the disablement in the physically active scale and preliminary testing of short-form versions: A calibration and validation study. *J Athl Train*. 2019;54(3):302-318.
21. Ullman J, Tabachnick B, Fidell L. Using multivariate statistics. *Struct Equ Modeling*. 2001:653-771.
22. Tabachnick BG, Fidell LS. Using multivariate statistics. (4th edn.) Needham Heights, MA: Allyn & Bacon. 2001.
23. Grimm LG, Yarnold PR. Reading and understanding multivariate statistics. American Psychological Association; 1995.
24. Brown TA. Confirmatory factor analysis for applied research. Guilford publications; 2015.
25. Byrne BM. Structural equation modeling with AMOS: Basic concepts, applications, and programming. Routledge; 2010.
26. Coltman, Tim, Timothy M. Devinney, David F. Midgley, and Sunil Venaik. Formative versus reflective measurement models: Two applications of formative measurement. Journal of Business Research. 2008, 61(12): 1250-1262.
27. Kim, Sangmook. Testing a revised measure of public service motivation: Reflective versus formative specification. Journal of Public Administration Research and Theory. 2011;21(3): 521-546.
28. Kenny DA, McCoach DB. Effect of the number of variables on measures of fit in structural equation modeling. Struct Equ Modeling. 2003;10(3):333-351.
29. Shi D, Lee T, Maydeu-Olivares A. Understanding the model size effect on SEM fit indices. Educ Psychol Meas. 2019;79(2):310-334.
30. Turner RW, Vissa K, Hall C, Poling K, Athey A, Alfonso-Miller P, Gehrels J, Grandner MA. Sleep problems are associated with academic performance in a national sample of collegiate athletes. J AM Coll Health. 2019:1-8.
31. Aries E, McCarthy E, Salovey P, Banaji MR. A comparison of athletes and non-athletes at highly selective colleges: Academic performance and personal development. Res High Educ. 2004; 45(6):577-602.
32. Matsunaga M. How to factor-analyze your data right: do’s, don’ts, and how-to’s. International Journal of Psychological Research. 2010;3(1):97-110.
LEGENDS TO FIGURES

Figure 1: ASBQ: Athlete Sleep Behavior Questionnaire; e: error; chisq: Chi-square; df: degrees of freedom; IFI: Bollen incremental fit index; RMSEA: root mean square error of approximation; TLI: Tucker-Lewis index; GFI: goodness-of-fit index; CFI: comparative fit index

Figure 2. ASBQ: Athlete Sleep Behavior Questionnaire; e: error; chisq: Chi-square; df: degrees of freedom; IFI: Bollen incremental fit index; RMSEA: root mean square error of approximation; TLI: Tucker-Lewis index; GFI: goodness-of-fit index; CFI: comparative fit index
chi^2 = 600.900  pvalue = .000  df = 132
IFI = .593  RMSEA = .080
TLI = .520  GFI = .888
CFI = .586
chisq = 43.018  pvalue = .010 df = 24
IFI = .952 RMSEA = .038 TLI = .926
GFI = .983 CFI = .951
| Criterion          | Definition                                                                 |
|--------------------|-----------------------------------------------------------------------------|
| Healthy            | Free from musculoskeletal injury and fully able to participate in sport or activity |
| Acute Injury       | A musculoskeletal injury that precludes full participation in sport or activity for at least 2 consecutive days (0–72 hours post-injury) |
| Sub-acute Injury   | A musculoskeletal injury that precludes full participation in sport or activity for at least 2 consecutive days (3 days to 1 month post-injury) |
| Persistent Injury  | A musculoskeletal injury that has been symptomatic for at least 1 month |
### Table 2. Demographic Information

|                          | Frequency (%) | Grade Point Average (GPA) |
|--------------------------|---------------|----------------------------|
| **Sex**                  |               |                            |
| Males                    | 104 (18.7)    | 3.28 (.45)                 |
| Females                  | 452 (81.3)    | 3.57 (.46)                 |
| **Ethnicity**            |               |                            |
| Caucasian                | 424 (76.3)    | 3.56 (.48)                 |
| African American         | 65 (11.7)     | 3.25 (.45)                 |
| Hispanic/Latino          | 55 (9.9)      | 3.44 (.39)                 |
| Asian/Pacific Islander   | 22 (4.0)      | 3.58 (.34)                 |
| Other                    | 7 (1.3)       | 3.34 (.57)                 |
| **Injury Status**        |               |                            |
| Healthy                  | 412 (74.8)    | 3.52 (.5)                  |
| Acute                    | 19 (3.4)      | 3.43 (.38)                 |
| Subacute                 | 27 (4.9)      | 3.39 (.52)                 |
| Persistent               | 93 (16.9)     | 3.51 (.37)                 |
| **Level of Competition** |               |                            |
| NCAA Division I          | 325 (58.5)    | 3.43 (.51)                 |
| NCAA Division II         | 19 (3.4)      | 3.36 (.34)                 |
| NCAA Division III        | 129 (23.2)    | 3.62 (.36)                 |
| NAIA                     | 6 (1.1)       | 3.68 (.31)                 |
| Junior College           | 12 (2.2)      | 3.67 (.21)                 |
| **Sport Type**           |               |                            |
| Dance                    | 304 (55.8)    | 3.62 (.42)                 |
| Traditional Student-Athlete | 241 (44.2)  | 3.35 (.5)                  |

*The sum does not equal 100% because percentages were rounded.*

*GPA is represented as mean (standard deviation).*
| Item                                                                 | Component 1 | Component 2 | Component 3 |
|----------------------------------------------------------------------|-------------|-------------|-------------|
| I think, plan and worry about issues not related to my sport when I am in bed. | .752        | -.054       | .014        |
| I think, plan and worry about my sporting performance when I am in bed. | .727        | .090        | .065        |
| I go to bed feeling thirsty.                                         | .524        | .228        | .007        |
| I go to bed at different times each night (more than ±1-hour variation). | .432        | .283        | .108        |
| At home, I sleep in a less than ideal environment (e.g. too light, too noisy, uncomfortable bed/pillow, too hot/cold). | .409        | .174        | .109        |
| I use light-emitting technology in the hour leading up to bedtime (e.g. laptop, phone, television, video games). | .380        | -.008       | -.003       |
| Travel gets in the way of building a consistent sleep-wake routine.  | .032        | .744        | -.113       |
| I sleep in foreign environments (e.g. hotel rooms).                 | -.011       | .671        | -.044       |
| I get up at different times each morning (more than ±1-hour variation). | .278        | .494        | -.075       |
| I wake myself and/or my bed partner with my snoring.                | -.221       | .436        | .236        |
| I take afternoon naps lasting two or more hours.                    | .158        | .389        | -.014       |
| I wake myself and/or my bed partner with my muscle twitching.       | .096        | .342        | .218        |
| I wake to go to the bathroom more than once per night.              | .262        | .332        | .111        |
| I use sleeping pills/tables to help me sleep.                       | .159        | .280        | .255        |
| I exercise (train or compete) late at night (after 7pm).            | .059        | -.137       | .711        |
| I use stimulants when I train/compete (e.g. caffeine).              | .097        | -.006       | .673        |
| I consume alcohol within 4 hours of going to bed.                   | -.069       | .148        | .580        |
| I go to bed with sore muscles.                                       | .413        | .002        | .430        |
| Eigenvalue                                                          | 2.96        | 1.79        | 1.47        |

*Bolded values show the loading for each component*
| Component | Item                                                                 | 1   | 2   | 3   |
|-----------|---------------------------------------------------------------------|-----|-----|-----|
| Routine   | 6. I go to bed feeling thirsty                                       | .768|     |     |
|           | 5. I go to bed at different times each night (more than ±1-hour variation) |     | .697|     |
|           | 16. At home, I sleep in a less than ideal environment (e.g. too light, too noisy, uncomfortable bed/pillow, too hot/cold) |     |     | .611|
| Behavior  | 2. I use stimulants when I train/compete (e.g. caffeine)            |     | .713|     |
|           | 3. I exercise (train or compete) late at night (after 7pm)          |     | .683|     |
|           | 4. I consume alcohol within 4 hours of going to bed                 |     |     | .671|
| Sport     | 18. Travel gets in the way of building a consistent sleep-wake routine |     |     | .816|
|           | 17. I sleep in foreign environments (e.g. hotel rooms)              |     |     | .776|
|           | 13. I wake myself and/or my bed partner with my snoring             |     |     | .492|
| Eigenvalue|                                                                      | 1.83| 1.52| 1.26|
| Cronbach’s alpha |                                                      | .48 | .47 | .52 |
Table 5. Goodness-of-Fit Indices for Measurement Invariance Analyses Across Sex

|                               | $\chi^2$ | df | $\chi^2_{diff}$ (df$_{diff}$) | CFI | CFI$_{diff}$ | TLI | RMSEA |
|-------------------------------|----------|----|-------------------------------|-----|-------------|-----|-------|
| Males (n = 104)               | 27.79    | 24 | ----                          | .944| ----        | .916|.039  |
| Females (n = 452)             | 36.81    | 24 | ----                          | .96 | ----        | .939|.034  |
| Configural (equal form)       | 64.72    | 48 | ----                          | .957| ----        | .935|.025  |
| Metric (equal loadings)       | 70.1     | 54 | 5.38 (6)                      | .958| .001        | .944|.023  |
| Equal factor variances*       | 73.58    | 57 | 8.78 (9)                      | .957| .001        | .940|.023  |
| Scalar (equal indicator intercepts) | 81.76 | 60 | 17.04 (12)                    | .943| **.014**    | .932|.026  |
| Equal latent means*           | ----     | ---- | ----                          | ----| ----        | ----| ----   |

* = Substantive questions; ---- = Value not calculated; **Bolded** = did not meet cutoff criteria
Table 6. Goodness-of-Fit Indices for Measurement Invariance Analyses Across Sport Type

|                              | $\chi^2$ | df | $\chi^2_{diff}$ (df$_{diff}$) | CFI  | CFI$_{diff}$ | TLI  | RMSEA |
|------------------------------|----------|----|-------------------------------|------|-------------|------|-------|
| Non-dancers (n = 240)        | 46.79    | 24 | ----                          | .843 | ----        | .764 | .063  |
| Dancers (n = 303)            | 27.46    | 24 | .843                          | .982 | .972        | .022 |       |
| Configural (equal form)      | 74.26    | 48 | ----                          | .921 | ----        | .881 | .032  |
| Metric (equal loadings)      | ----     | ----| ----                          | ---- | ----        | ---- | ----   |
| Equal factor variances*      | ----     | ----| ----                          | ---- | ----        | ---- | ----   |
| Scalar (equal indicator intercepts) | ----     | ----| ----                          | ---- | ----        | ---- | ----   |
| Equal latent means*          | ----     | ----| ----                          | ---- | ----        | ---- | ----   |

* = Substantive questions; ---- = Value not calculated; **Bolded** = did not meet cuff off criteria
|               | $\chi^2$ | df  | $\chi^2_{\text{diff}}$ (df<sub>diff</sub>) | CFI   | CFI<sub>diff</sub> | TLI  | RMSEA |
|---------------|----------|-----|--------------------------------|-------|------------------|------|-------|
| Healthy (n = 412) | 38.91    | 24  | ----                          | .951  | ----              | .926 | .039  |
| Injured (n = 139)     | 24.66    | 24  | ----                          | .991  | ----              | .986 | .014  |
| Configural (equal form) | 63.61    | 48  | ----                          | .958  | ----              | .938 | .024  |
| Metric (equal loadings) | 64.61    | 54  | 1.00 (6)                      | .972  | .014             | .962 | .019  |
| Equal factor variances* | ----     | --- | -                             | ----  | ----              | ---- | ----   |
| Scalar (equal indicator intercepts) | ---- | --- | -                             | ----  | ----              | ---- | ----   |
| Equal latent means*   | ----     | --- | -                             | ----  | ----              | ---- | ----   |

* = Substantive questions; ---- = Value not calculated; **Bolded** = did not meet cutoff criteria
| Level of Competition                  | $\chi^2$ | df  | $\chi^2_{\text{diff}}$ (df_{\text{diff}}) | CFI   | CFI_{\text{diff}} | TLI   | RMSEA |
|--------------------------------------|---------|-----|----------------------------------------|-------|------------------|-------|-------|
| NCAA D1 (n = 325)                    | 40.32   | 24  | ----                                   | .932  | ----             | .897  | .046  |
| Lower level (n = 166)                | 43.02   | 24  | ----                                   | .867  | ----             | .801  | .069  |
| Configural (equal form)              | 83.38   | 48  | ----                                   | .907  | ----             | .861  | .039  |
| Metric (equal loadings)              | 93.95   | 54  | 10.57 (6)                               | .896  | .011             | .861  | .039  |
| Equal factor variances*              | ----    | ----| ----                                    | ----  | ----             | ----  | ----  |
| Scalar (equal indicator intercepts)  | ----    | ----| ----                                    | ----  | ----             | ----  | ----  |
| Equal latent means*                  | ----    | ----| ----                                    | ----  | ----             | ----  | ----  |

* = Substantive questions; ---- = Value not calculated; **Bolded** = did not meet cutoff criteria