

Confirmatory Factor Analysis of the Athlete Sleep Behavior Questionnaire

Emilie N. Miley, DAT, ATC*†; Bethany L. Hansberger, DAT, ATC||;
Madeline Casanova, PhD‡; Russell T. Baker, PhD, DAT, ATC‡§||;
Michael A. Pickering, PhD¶

*University of Florida, Gainesville; †Department of Movement Sciences, ‡WWAMI Medical Education Program, and §Athletic Training Programs, University of Idaho, Moscow; ||Towson University, MD; ¶Trident University, Cypress, CA

Context: Sleep has long been understood as an essential component for overall well-being, substantially affecting 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 in a sample of collegiate traditional student-athletes and dancers.

Design: Observational study.

Setting: Twelve colleges and universities.

Patients or Other Participants: A total of 556 (104 men, 452 women; age = 19.84 ± 1.62 years) traditional student-athletes and dancers competing at the collegiate level.

Main Outcome Measure(s): A confirmatory factor analysis (CFA) was computed to assess the factor structure of the ASBQ. We performed principal component analysis extraction and

covariance modeling analyses to identify an alternate model. Multigroup invariance testing was conducted on the alternate model to identify if group differences existed for sex, sport type, injury status, and level of competition.

Results: The CFA on the ASBQ indicated that the model did not meet recommended model fit indices. An alternate 3-factor, 9-item model with improved fit was identified; however, the scale structure was not consistently supported during multigroup invariance testing procedures.

Conclusions: The original 3-factor, 18-item ASBQ was not supported for use with collegiate athletes in our study. The alternate ASBQ was substantially improved, although more research should be completed to ensure that the 9-item instrument accurately captures all dimensions of sleep behavior relevant for collegiate athletes.

Key Words: measurement, covariance modeling, athletic population

Key Points

- Sleep is multifactorial and an important component for athletic trainers to consider in the treatment of collegiate athletes.
- The Athlete Sleep Behavior Questionnaire did not meet contemporary model fit recommendations.
- Clinicians should be cautious when using the Athlete Sleep Behavior Questionnaire (original or modified) given the concerns regarding model fit and instrument design.

Sleep has long been understood as an essential component for overall well-being.^{1–3} Specifically, sleep substantially affects physical health, cognitive function, mental health, and quality of life.^{1–3} Sleep has been proposed as a multifactorial construct that is affected by external (eg, anxiety, noise, the need to use the bathroom, early event times) and internal (eg, circadian rhythm) factors, which often affect each other.^{4,5} For example, when sleep disturbances occur because of 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 affect sport performance in athletes.^{6,7} Adequate sleep quality and quantity are important for optimal performance indices (eg, reaction time, learning, memory tasks).⁸

Compared with nonathletes, athletes have reduced sleep quality. Most athletes described sleeping fewer than the recommended target hours (ie, 8 hours of sleep) the night

before competition, and approximately 70% of athletes reported problematic or poor sleeping patterns before competition compared with their normal routine.⁴ Poor sleeping patterns increase fatigue and tension, which are negatively correlated with precompetitive relative sleep quality and total sleep time.⁴ The poor sleep-quality patterns identified may also detrimentally affect student-athlete success.⁹

Researchers^{10,11} investigating sleep quality and overall well-being in athletes indicated that athletes who experienced acute partial (ie, average of 2.5 hours) sleep deprivation exhibited an increase in negative mood states (ie, depression, tension, confusion, fatigue, and anger) and decreased vigor (ie, physical strength, good health, and energy).^{10,11} Sleep quality and quantity also affected an athlete's ability to recover after activity¹²; good sleep patterns (ie, quality and quantity) are considered an important recovery method.^{9,13} Long-term negative changes

Table 1. The Athlete Sleep Behavior Questionnaire (ASBQ)

No.	In recent times (over the last month) . . .	Never	Rarely	Sometimes	Frequently	Always
1.	I take afternoon naps lasting two or more hours					
2.	I use stimulants when I train/compete (eg, caffeine)					
3.	I exercise (train or compete) late at night (after 7pm)					
4.	I consume alcohol within 4 hours of going to bed					
5.	I go to bed at different times each night (more than ± 1 hour variation)					
6.	I go to bed feeling thirsty					
7.	I go to bed with sore muscles					
8.	I use light-emitting technology in the hour leading up to bedtime (eg, laptop, phone, television, video games)					
9.	I think, plan and worry about my sporting performance when I am in bed					
10.	I think, plan and worry about issues not related to my sport when I am in bed					
11.	I use sleeping pills/tablets to help me sleep					
12.	I wake to go to the bathroom more than once per night					
13.	I wake myself and/or my bed partner with my snoring					
14.	I wake myself and/or my bed partner with my muscle twitching					
15.	I get up at different times each morning (more than ± 1 hour variation)					
16.	At home, I sleep in a less than ideal environment (eg, too light, too noisy, uncomfortable bed/pillow, too hot/cold)					
17.	I sleep in foreign environments (eg, hotel rooms)					
18.	Travel gets in the way of building a consistent sleep-wake routine					

Scoring:

Never = 1, Rarely = 2, Sometimes = 3, Frequently = 4, Always = 5

Total Global Score: _____

This table is reproduced in its original form according to the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>). Reprinted with permission from Driller MW, Mah CD, and Halson SL.¹²

in mood, possibly due to poor sleep patterns, may be related to increased injury risk¹¹ and high player workloads; continual poor sleep history may further increase the injury risk because of elevated levels of chronic fatigue.^{13,14}

Therefore, sleep behavior may be an important factor for researchers to measure when assessing injury risk or recovery.¹³ Many self-report instruments have been developed to evaluate sleep in a general population; however, few have been 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).¹² The ASSQ is used to assess clinical sleep difficulties in athletes and was designed to identify at-risk individuals who may need a referral to a sleep specialist.¹² Although the ASSQ evaluates 6 factors related to sleep difficulty (ie, total sleep time, insomnia, sleep quality, sleep chronotype, sleep disordered breathing, and travel disturbance), it does not gather information on sleep behavior practices.¹² Thus, investigators¹² developed the ASBQ to assess sleep in elite international athletes by using a combination of newly developed items and items drawn from previously validated questionnaires (ie, the Sleep Hygiene Index and the International Classification of Sleep Disorders; Table 1). 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.^{12,14} Researchers¹² used principal component analysis (PCA) and identified a 3-construct solution (ie, Routine, Behavioral, and Sport components), thus endorsing the use of the scale.

Yet despite positive initial findings, several limitations or concerns support the need for further psychometric examination before the ASBQ is used in clinical practice. For example, during initial development of the ASBQ, the authors included a sample of elite athletes and nonathletes¹⁴

but not other traditional competitive athletic categories (eg, collegiate athletes) or a more general physically active population (ie, recreational athletes), thereby limiting the ability to ensure scale applicability and use across different levels in the athletic population. Other scale design concerns, such as the use of double-barreled questions (ie, items that ask about >1 topic)¹⁵ or poor internal consistency (ie, Cronbach $\alpha < 0.70$)¹⁶, are present in the scale. Additionally, the scale structure has not been confirmed in subsequent work; investigators^{12,17} who translated the scale into the Turkish language identified a 4-factor solution that did not match the originally proposed 3-factor model. Finally, recommended scale psychometric evaluation procedures, including confirmatory factor analysis (CFA) and invariance testing, have not been published, and these analyses are necessary to establish the measurement properties of a scale in order to endorse its use in research and practice.^{18,19}

The lack of published CFA and invariance results, combined with the inconsistent factor structure, warrants further evaluation of the psychometric properties of the ASBQ in larger, more diverse collegiate athlete samples. Thus, the primary purpose of our study was to assess the proposed reflective measurement model by using CFA to assess model fit and the psychometric properties of the original ASBQ scale in a sample of collegiate athletes. If the ASBQ did not meet the recommended model fit criteria, the secondary purposes of the study were (1) to perform PCA and covariance modeling to identify and assess alternate models and (2), if the alternate model fit criteria were met, to perform multigroup (ie, men versus women, nondancers versus dancers, healthy versus injured, and National Collegiate Athletic Association [NCAA] Division I versus lower division of competition) invariance testing of the scale.

Table 2. Definitions of Injury Classifications

Criterion	Definition ^{20,21}
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 ≥ 2 consecutive d (0–72 h postinjury)
Subacute injury	A musculoskeletal injury that precludes full participation in sport or activity for ≥ 2 consecutive d (3 d to 1 mo postinjury)
Persistent injury	A musculoskeletal injury that has been symptomatic for ≥ 1 mo

Adapted with permission.^{20,21}

METHODS

Participants

The study was identified as exempt by the university institutional review board, and all participants provided written informed consent. A convenience sample of athletic trainers and dance faculty recruited healthy and injured individuals from 12 colleges and universities in the United States; participants were recruited from different collegiate competition levels (eg, NCAA Division I or National Association of Intercollegiate Athletics [NAIA]). Participants were grouped by injury status (ie, healthy or acute, subacute, or persistent injury; Table 2),^{20,21} sex (male or female), and sport type (eg, collegiate dance or traditional student-athlete; Table 3).

Instrumentation

A survey packet that consisted of the ASBQ and a demographic questionnaire was created in identical paper

Table 3. Descriptive Data

Variable	Frequency (%) ^a	Grade Point Average, Mean \pm SD
Sex		
Male	104 (18.7)	3.28 \pm 0.45
Female	452 (81.3)	3.57 \pm 0.46
Ethnicity ^b		
Caucasian	424 (74.0)	3.56 \pm 0.48
African American	65 (11.3)	3.25 \pm 0.45
Hispanic or Latino	55 (9.6)	3.44 \pm 0.39
Asian or Pacific Islander	22 (3.8)	3.58 \pm 0.34
Other	7 (1.3)	3.34 \pm 0.57
Injury status ^c		
Healthy	412 (74.8)	3.52 \pm 0.5
Acute	19 (3.4)	3.43 \pm 0.38
Subacute	27 (4.9)	3.39 \pm 0.52
Persistent	93 (16.9)	3.51 \pm 0.37
Level of competition ^d		
National Collegiate Athletic Association		
Division I	325 (66.2)	3.43 \pm 0.51
Division II	19 (3.9)	3.36 \pm 0.34
Division III	129 (26.3)	3.62 \pm 0.36
National Association of Intercollegiate Athletics	6 (1.2)	3.68 \pm 0.31
Junior college	12 (2.4)	3.67 \pm 0.21
Sport type ^e		
Collegiate dance	303 (55.8)	3.62 \pm 0.42
Traditional student-athlete	240 (44.2)	3.35 \pm 0.5

^a Percentages for each variable were calculated from the total number of responses for that variable.

^b Participants could select >1 ethnicity and provided 573 responses.

^c A total of 551 participants reported injury status.

^d A total of 491 participants reported level of competition.

^e A total of 543 participants reported sport type.

and electronic formats. Descriptive data collected were age, sport, self-reported injury category (ie, healthy, acute, subacute, or persistent), and level of competition (eg, NCAA Division I or NAIA). The electronic version of the packet was created using Qualtrics software (Qualtrics, LLC). Responses from the paper packet were input into Qualtrics by the participating athletic trainer or were mailed 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.

The ASBQ is a 3-factor, 18-item instrument designed to assess sleep patterns related to routine, behavior, and sport in elite athletes. Participants were asked how frequently they had engaged in specific sleep and sport behaviors over the past month.¹² They provided their answers using a 5-point Likert scale ranging from 1 (*never*) to 5 (*always*).¹² Item scores were summed to create a global ASBQ score, ranging from 18 to 90, with higher scores indicating poorer sleeping behaviors.¹² The ASBQ has been reported to have excellent reliability (intraclass correlation coefficient = 0.87, $r = 0.88$, coefficient of variation = 6.4%); however, the reported Cronbach α for internal consistency was 0.63.¹²

Data Analysis

Survey responses were downloaded and analyzed using SPSS (version 25; IBM Corp) and AMOS (version 25; IBM Corp) software. Missing responses for each survey item were calculated for each respondent, and individuals missing $>10\%$ of the ASBQ items (ie, ≥ 2 questions) were removed from the data set. Individuals missing $<10\%$ of the data were retained, and respective missing items were replaced with the rounded mean score for analysis.^{19,22} Participants missing descriptive data were not excluded from the analysis. Skewness and kurtosis values, as well as histograms, were assessed for normality. The data were analyzed for univariate outliers using z scores, and multivariate outliers were identified using the Mahalanobis distance at $P < .01$.^{19,23}

Confirmatory Factor Analysis. We conducted the CFA using maximum likelihood estimation in AMOS software across the sample. We conducted 2 additional CFAs for sample subgroups (ie, in-season traditional athletes and collegiate dancers) because the original ASBQ was designed for elite athletes and may be more commonly used while athletes are participating in sport (ie, in-season training). The proposed 3-factor structure was assessed using a priori model fit indices. The goodness-of-fit indices computed were the Comparative Fit Index (CFI; ≥ 0.95), Goodness of Fit Index (GFI; ≥ 0.95), Tucker-Lewis Index (TLI; ≥ 0.95), root mean square error of approximation (RMSEA; ≤ 0.05), and Bollen Incremental Fit Index (IFI; ≥ 0.95).^{19,24} Model fit was also evaluated for localized areas

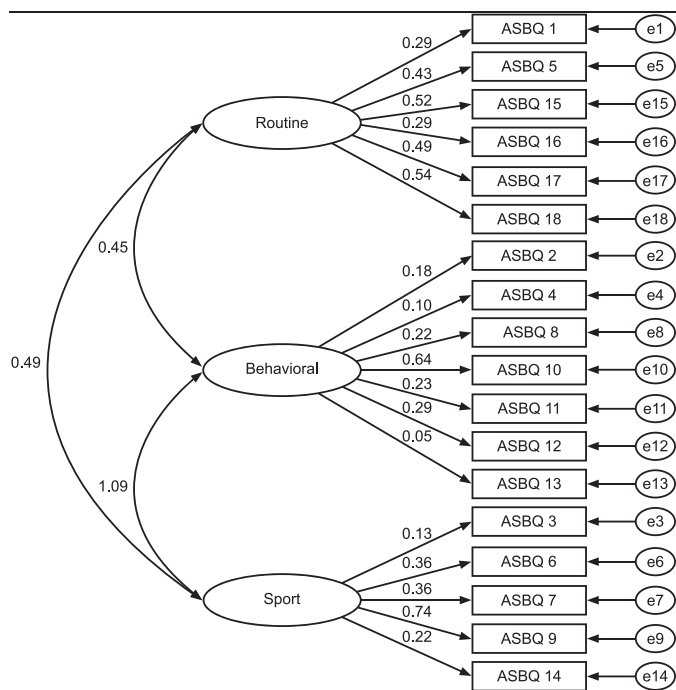


Figure 1. The Athlete Sleep Behavior Questionnaire (ASBQ) hierarchical confirmatory factor analysis measurement model with standardized loadings. See Table 1 for item descriptions. Abbreviation: e, error.

of strain, including statistical significance of parameter estimates.^{19,25}

Alternate Model Generation. A PCA with varimax rotation was conducted using SPSS to identify a more parsimonious model if the criteria were not met for the original CFA. A PCA with varimax rotation (ie, orthogonal rotation) was applied to replicate the original analytic procedures used to establish the scale.¹² Items were removed one at a time, and the PCA was repeated until a solution that met the recommendations was determined. Item removal was guided by the statistical (eg, low loadings [<0.40] or high cross-loadings [≥ 0.30] with other items),¹⁶ theoretical (eg, does the item make sense with the other items that have factored?), and survey design (eg, double barreled) concerns identified across items at each step of the iterative PCA extraction process.^{12,16,24} The PCA process was repeated by 2 investigators (E.N.M. and M.C.) to ensure that the same final solution was reached. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (≥ 0.70) and Bartlett test of sphericity ($P < .001$) were assessed for violations,¹⁶ and the extraction was fixed to retain 3 components as specified by previous researchers.¹²

We then assessed the alternate model identified during the PCA 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. Parameter estimates and modification indices were evaluated for local strain identification. Bivariate correlation analyses were conducted using the cumulative (ie, total) and construct scores from the 18-item ASBQ scale and the cumulative and construct scores from the newly proposed modified 9-item ASBQ; a priori thresholds (<0.1 , *trivial*; 0.1 – 0.3 , *small*; 0.3 – 0.5 , *moderate*; 0.5 – 0.7 , *large*; 0.7 – 0.9 , *very large*; and 0.9 – 1.0 , *almost perfect*) were used to

characterize the magnitude of correlation between scales and constructs.¹²

Multigroup Invariance Testing. Multigroup CFA invariance testing (ie, configural, metric, and scalar) was planned for the original or alternate model, depending on which model met contemporary recommendations. Multigroup invariance testing was performed using AMOS software to assess model fit across different subgroups, as follows: sex (ie, male or female), sport type (ie, collegiate dancer or traditional student-athlete), self-reported injury status (ie, healthy or injured), and level of competition (ie, NCAA Division I or lower division athlete). The CFI difference test (CFI_{DIFF}) and χ^2 difference test (χ^2_{DIFF}) were used to evaluate model fit, with a cutoff of $P = .01$.^{16,19,26} We placed more emphasis on the CFI_{DIFF} test because of the sensitivity of the χ^2_{DIFF} test to sample size^{19,26}; if a model exceeded the χ^2_{DIFF} test but met the CFI_{DIFF} test, invariance testing continued.

RESULTS

A total of 614 individuals completed the survey; 9 individuals (1.5%) were missing responses to $>10\%$ 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 ≥ 3.4) or multivariate (Mahalanobis distance ≥ 33) outliers and were subsequently removed from the data set.^{22,26} These participants included both sexes, 3 injury categories (ie, healthy, acute, or persistent), and various levels of competition (eg, Division I or NAIA). Removing these respondents 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 men, 452 women; age = 19.84 ± 1.62 years; age range = 16–32 years; Table 3). Most respondents ($n = 325/491$, 66.2%) were involved at the NCAA Division I level and were classified as healthy ($n = 412/551$, 74.8%; Table 3).

Confirmatory Factor Analysis

The CFA of the 3-factor, 18-item ASBQ did not meet recommended model fit indices for the full sample data ($\chi^2 = 600.900$, $P < .001$, CFI = 0.586, GFI = 0.888, TLI = 0.520, RMSEA = 0.080, IFI = 0.593; Figure 1). The model also failed to meet recommendations for in-season traditional athletes ($\chi^2 = 216.09$, $P < .001$, CFI = 0.562, TLI = 0.492, RMSEA = 0.087, IFI = 0.605) and collegiate dancers ($\chi^2 = 372.29$, $P < .001$, CFI = 0.602, TLI = 0.539, RMSEA = 0.078, IFI = 0.616). Moreover, several potential fit concerns arose in the full sample: factor loadings that were not different ($P > .05$), low item loadings (<0.40), and standardized path coefficients >1 . The latent variable correlations were small (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 3-component PCA solution included items with low loadings (<0.40) and 1 item with a high cross-loading (Table 4). In total, 9 items with low loadings,

Table 4. Initial Fixed 3-Factor Extraction

Athlete Sleep Behavior Questionnaire Item ^a	Factor		
	1	2	3
10.	0.752 ^b	-0.054	0.014
9.	0.727 ^b	0.090	0.065
6.	0.524 ^b	0.228	0.007
5.	0.432 ^b	0.283	0.108
16.	0.409 ^b	0.174	0.109
8.	0.380 ^b	-0.008	-0.003
18.	0.032	0.744 ^b	-0.113
17.	-0.011	0.671 ^b	-0.044
15.	0.278	0.494 ^b	-0.075
13.	-0.221	0.436 ^b	0.236
1.	0.158	0.389 ^b	-0.014
14.	0.096	0.342 ^b	0.218
12.	0.262	0.332 ^b	0.111
11.	0.159	0.280 ^b	0.255
3.	0.059	-0.137	0.711 ^b
2.	0.097	-0.006	0.673 ^b
4.	-0.069	0.148	0.580 ^b
7.	0.413	0.002	0.430 ^b
Eigenvalue	2.96	1.79	1.47

^a See Table 1 for item descriptions.

^b Loading for each component.

high cross-loadings, or high interitem correlations were removed during subsequent analyses of the PCA procedures. The resulting 3-component, 9-item solution contained items with loadings >0.49 and without substantial cross-loadings. The solution accounted for 23.4% of the variance, with Cronbach α ranging from 0.47 to 0.52 (Table 5). Total scores on the 9-item modified ASBQ demonstrated a very large correlation ($r = 0.850$; $R^2 = 0.722$) with the scores for the 18-item ASBQ. Large relationships were found between the behavioral constructs ($r = 0.635$; $R^2 = 0.403$) and routine constructs ($r = 0.643$; $R^2 = 0.413$) of the modified 9-item ASBQ and the 18-item ASBQ, whereas a small relationship was found between the sport constructs of the 2 scales ($r = 0.134$; $R^2 = 0.018$).

Covariance Model Refined ASBQ

Covariance modeling of the refined 3-factor, 9-item scale showed improved model fit, with goodness-of-fit indices meeting nearly all criteria ($\chi^2 = 43.018$, $P = .01$, CFI = 0.951, GFI = 0.983, TLI = 0.926, RMSEA = 0.038, IFI =

Table 5. Fixed 3-Factor Principal Component Analysis

Component	Athlete Sleep Behavior Questionnaire Item ^a	Factor		
		1	2	3
Routine	6.	0.768 ^b		
	5.	0.697 ^b		
	16.	0.611 ^b		
Behavioral	2.		0.713 ^b	
	3.		0.683 ^b	
	4.		0.671 ^b	
Sport	18.			0.816 ^b
	17.			0.776 ^b
	13.			0.492 ^b
Eigenvalue		1.83	1.52	1.26
Cronbach α		0.48	0.47	0.52

^a See Table 1 for item descriptions.

^b Item loaded within the criteria on the factor.

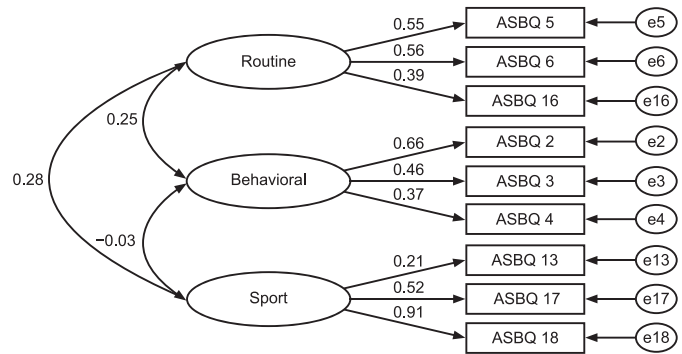


Figure 2. The Athlete Sleep Behavior Questionnaire (ASBQ) 9-item covariance model with standardized loadings. See Table 1 for item descriptions. Abbreviation: e, error.

0.952; Figure 2). The latent variable correlations were small (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 different, ranging from 0.21 to 0.91, and modification indices did not reveal any meaningful cross-loadings or error covariance specifications.

Multigroup Invariance Testing Across Subgroups

Sex. All 556 individuals in the sample reported their sex (men = 104, women = 452). Individual CFAs by sex met some but not all recommended fit criteria for men (CFI = 0.944, TLI = 0.916, RMSEA = 0.039) and women (CFI = 0.96, TLI = 0.939, RMSEA = 0.034; Table 6). The configural model (ie, equal form) met most model fit indices ($\chi^2 = 64.72$, CFI = 0.957, RMSEA = 0.025; Table 6). The metric model (ie, equal loadings) passed both the CFI_{DIFF} test (CFI_{DIFF} = 0.001) and the χ^2_{DIFF} test ($\chi^2_{DIFF} = 5.38$). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variances model passed the CFI_{DIFF} test (CFI_{DIFF} = 0.001) and the χ^2_{DIFF} test ($\chi^2_{DIFF} = 8.78$), indicating that the variances were not different between groups. The scalar model (ie, equal indicator intercepts) did not pass the CFI_{DIFF} test (CFI_{DIFF} = 0.014) but passed the χ^2_{DIFF} test ($\chi^2_{DIFF} = 17.04$). As such, completing the subsequent steps of the multigroup invariance testing process (ie, testing of means) was not deemed appropriate.

Sport Type. Of the 556 individuals in the sample, 543 (97.7%) reported their athletic classification (collegiate dancers = 303, traditional student-athletes = 240) and were included in the analysis. Individual CFAs by individual sport type indicated that the model did not meet the recommended fit for traditional student-athletes (CFI = 0.843, TLI = 0.764, RMSEA = 0.063) but did meet the recommended fit for collegiate dancers (CFI = 0.982, TLI = 0.972, RMSEA = 0.022; Table 7). The configural model (ie, equal form) did not meet the recommended model fit indices ($\chi^2 = 74.26$, CFI = 0.921, RMSEA = 0.032; Table 7). Therefore, completing the subsequent steps of the multigroup invariance testing process (eg, metric and equal latent means) was not deemed appropriate.

Injury Status. Of the 556 individuals in the sample, 551 (99.1%) provided their injury status (healthy = 412, injured = 139) and were included in the analysis. Individual CFAs

Table 6. Goodness-of-Fit Indices for Measurement Invariance Analyses By Sex

Variable	χ^2	χ^2 Difference Test			Comparative Fit Index	Tucker-Lewis Index	Root Mean Square Error of Approximation
		Degrees of Freedom	(Degrees of Freedom Difference)	Comparative Fit Index			
Men (n = 104)	27.79	24	NA	0.944	NA	0.916	0.039
Women (n = 452)	36.81	24	NA	0.96	NA	0.939	0.034
Configural model (equal form)	64.72	48	NA	0.957	NA	0.935	0.025
Metric model (equal loadings)	70.1	54	5.38 (6)	0.958	0.001	0.944	0.023
Equal factor variances model ^a	73.58	57	8.78 (9)	0.957	0.001	0.940	0.023
Scalar model (equal indicator intercepts)	81.76	60	17.04 (12)	0.943	0.014 ^b	0.932	0.026
Equal latent means ^a	NA	NA	NA	NA	NA	NA	NA

Abbreviation: NA, value not calculated.

^a Indicates substantive question.

^b Indicates value did not meet cutoff criteria.

by injury status showed that the model met some of the recommended fit criteria for the healthy group (CFI = 0.951, TLI = 0.926, RMSEA = 0.039) and all the fit criteria for the injured group (CFI = 0.991, TLI = 0.986, RMSEA = 0.014; Table 8). The configural model (ie, equal form) met most model fit indices ($\chi^2 = 63.61$, CFI = 0.958, RMSEA = 0.024; Table 8). The metric model (ie, equal loadings) did not pass the CFI_{DIFF} test (CFI_{DIFF} = 0.014). Thus, completing the subsequent steps of the multigroup invariance (eg, scalar and equal latent means) testing process was not warranted.

Level of Competition. Of the 556 individuals in the sample, 491 (88.3%) stated their level of competition (Division I = 325, lower division of competition = 166) and were included in the analysis. Individual CFAs by the level of competition indicated that the models did not meet the recommended fit criteria for NCAA Division I athletes (CFI = 0.932, TLI = 0.897, RMSEA = 0.046) or the lower-division athletes (CFI = 0.867, TLI = 0.801, RMSEA = 0.069; Table 9). The configural model (ie, equal form) did not meet the model fit indices ($\chi^2 = 83.38$, CFI = 0.907, RMSEA = 0.039; Table 9). As such, the multigroup invariance testing process (eg, metric and scalar) was not warranted.

DISCUSSION

The first purpose of our study was to examine the psychometric properties of the originally proposed ASBQ in a broader athletic population (ie, collegiate traditional student-athletes and collegiate dancers). Because model fit indices were not met, the secondary purpose was to use

PCA and alternate model generation to determine if a modified ASBQ could be identified from the item pool for use in the collegiate athlete population. Contemporary psychometric analysis methods were applied to assess the model fit of the ASBQ and the alternate model to guide recommendations for use in future research and clinical practice. Our results suggested that the original ASBQ has poor psychometric properties and should not be used in collegiate athlete populations. The alternate model met many fit recommendations; however, given the scale concerns and psychometric testing results that did not meet all recommended criteria, further exploration is warranted before adoption in clinical practice and research.

The CFA of the ASBQ

Our CFA findings did not support the model scale structure proposed in the original study.¹² Model fit was poor, with specific concerns related to low item loadings (<0.40) and model misspecification as evidenced by the standardized loading between the behavioral and sport latent constructs being >1 ($r = 1.09$).¹⁹ Furthermore, high latent variable correlation values (ie, ≥ 0.95)²⁷ indicated potential multicollinearity and a lack of unique constructs being measured; accordingly, item removal or modification of the items was warranted.^{19,25} The instrument may be improved by condensing the scale, rewording items, or developing new items to more effectively measure the originally proposed dimensions.¹⁵ Further testing (ie, invariance analyses) on the original ASBQ was not supported in our sample; therefore, alternate models were explored.²⁶ Our findings did not support the use of the

Table 7. Goodness-of-Fit Indices for Measurement Invariance Analyses By Sport Type

Variable	χ^2	χ^2 Difference Test			Comparative Fit Index	Tucker-Lewis Index	Root Mean Square Error of Approximation
		Degrees of Freedom	(Degrees of Freedom Difference)	Comparative Fit Index			
Nondancers ^a (n = 240)	46.79	24	NA	0.843	NA	0.764	0.063
Dancers (n = 303)	27.46	24	NA	0.982	NA	0.972	0.022
Configural model (equal form)	74.26 ^c	48 ^c	NA	0.921 ^c	NA	0.881 ^c	0.032 ^c
Metric model (equal loadings)	NA	NA	NA	NA	NA	NA	NA
Equal factor variances model ^b	NA	NA	NA	NA	NA	NA	NA
Scalar model (equal indicator intercepts)	NA	NA	NA	NA	NA	NA	NA
Equal latent means ^b	NA	NA	NA	NA	NA	NA	NA

Abbreviation: NA, value not calculated.

^a The term *nondancers* was used to describe the traditional student-athletes in the goodness-of-fit index.

^b Indicates substantive question.

^c Indicates value did not meet cutoff criteria.

Table 8. Goodness-of-Fit Indices for Measurement Invariance Analyses By Injury Status

Variable	χ^2	Degrees of Freedom	χ^2 Difference Test		Comparative Fit Index	Tucker-Lewis Index	Root Mean Square Error of Approximation	
			(Degrees of Freedom)	Difference				
Healthy (n = 412)	38.91	24	NA		0.951	NA	0.926	0.039
Injured (n = 139)	24.66	24	NA		0.991	NA	0.986	0.014
Configural model (equal form)	63.61	48	NA		0.958	NA	0.938	0.024
Metric model (equal loadings)	64.61	54	1.00 (6)		0.972	0.014 ^b	0.962	0.019
Equal factor variances model ^a	NA	NA	NA		NA	NA	NA	NA
Scalar model (equal indicator intercepts)	NA	NA	NA		NA	NA	NA	NA
Equal latent means ^a	NA	NA	NA		NA	NA	NA	NA

Abbreviation: NA, value not calculated.

^a Indicates substantive question.

^b Indicates value did not meet cutoff criteria.

originally proposed 3-factor, 18-item ASBQ instrument in a collegiate athlete population or in subgroups of physically active individuals who were actively engaged in training (ie, in-season traditional athletes and collegiate dancers); consequently, we do not recommend using it in this population without alteration.^{19,25} Moreover, on further review of the items identified in the original ASBQ, most items appeared to be formative (ie, a change to the indicator is associated with variation of the latent construct) and not reflective (ie, the item reflects a change of the latent construct).²⁷⁻²⁹ This outcome could indicate misspecification that may be leading to poor fit of the model.²⁷⁻²⁹

Alternate Model Generation and Multigroup Invariance Testing

The alternate model produced a similar 3-factor structure as the original ASBQ scale¹²; however, model fit (eg, low factor loadings and Cronbach α levels less than recommended levels) and instrument design concerns remained.^{15,19,25,29} Furthermore, 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 (ie, routine, behavioral, and sport) did not factor into the originally proposed construct. For the routine construct, 4 of the original items (items 5, 16, 17, and 18) were retained in the 9-item model; only 2 of these items (items 5 and 16) loaded on the routine construct, and the other 2 (items 17 and 18) loaded on the sport construct. For the behavioral construct, 3 of the original items (items 2, 4, and 13) were retained in the 9-item model; 2 of these items (items 2 and 4) loaded on the behavioral construct, and 1 item (item 13) loaded on the sport construct. Only 2 items

(items 3 and 6) from the original sport construct were retained in the final 9-item model, but neither loaded on the sport construct; 1 item (item 3) loaded on the behavioral construct, and 1 item (item 6) loaded on the routine construct.

Despite item removal, the very large correlation ($r = 0.850$) between total scores on the modified ASBQ and the original ASBQ suggested the 9-item version accounted for most of the variance in participant responses in the 18-item version and indicated the modified scale captured a theoretical measurement of sleep behaviors similar to that of the original scale. Moreover, the total score correlation findings suggested a similar phenomenon was being measured across the 2 scales and served as evidence of item redundancy in the 18-item model. The large construct correlations between the routine constructs ($r = 0.643$) and behavioral constructs ($r = 0.635$) across scales also reflected measurement of a similar phenomenon in each construct across the 2 scales. Yet the correlation between the sport constructs of the 2 scales was small and indicated that a different phenomenon was primarily assessed across these 2 constructs. The strong but not perfect correlations for the total score, routine constructs, and behavioral constructs of the 2 scale versions were expected because of the similar items (ie, the modified ASBQ contains 9 of the original 18 items and 2 of the 3 items in the modified routine and behavioral constructs were retained from the original constructs). Additionally, a strong correlation between scale versions was expected because the redundant and poor fitting items that resulted in measurement error and variation in the original scale were removed. The small correlation for the sport constructs was also anticipated because none of the items were shared between the 2

Table 9. Goodness-of-Fit Indices for Measurement Invariance Analyses By Level of Competition

Variable	χ^2	Degrees of Freedom	χ^2 Difference Test		Comparative Fit Index	Tucker-Lewis Index	Root Mean Square Error of Approximation	
			(Degrees of Freedom)	Difference				
NCAA Division I (n = 325)	40.32	24	NA		0.932	NA	0.897	0.046
Lower division (n = 166)	43.02	24	NA		0.867	NA	0.801	0.069
Configural model (equal form)	83.38	48	NA		0.907	NA	0.861	0.039
Metric model (equal loadings)	NA	NA	NA		NA	NA	NA	NA
Equal factor variances model ^a	NA	NA	NA		NA	NA	NA	NA
Scalar model (equal indicator intercepts)	NA	NA	NA		NA	NA	NA	NA
Equal latent means ^a	NA	NA	NA		NA	NA	NA	NA

Abbreviations: NA, value not calculated; NCAA, National Collegiate Athletic Association.

^a Indicates substantive question.

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 (ie, item removal) were necessary to address the previously discussed fit (eg, multicollinearity between the latent variables and 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.^{16,19} Our findings were confirmed in the covariance model, which had a substantially improved model fit and reduced latent variable correlations (Figure 2). Of note, 8 of the 9 items (89%) removed presented as formative indicators. Because most of the items removed were formative, the removal of these items may have reduced model misspecification, which may also explain the increase in model fit statistics.

However, concerns with the alternate model were noted when multigroup invariance testing between subgroups was performed. Baseline models were assessed between men and women; the model fit criteria met some but not all contemporary model fit recommendations.^{16,19} Group differences in variances were not observed for poor sleep behaviors between sexes. Because the scalar model did not meet contemporary model fit recommendations, further invariance testing for group mean differences was not supported. With the modified ASBQ, the scalar model results indicated that men and women did not conceptualize sleep behavior similarly. Therefore, group mean differences between men and women on the modified scale should not be interpreted as true group differences until further testing with a larger sample is conducted to refute our multigroup invariance test findings.

Multigroup invariance testing was also conducted on the modified ASBQ across traditional student-athletes and collegiate dancer subgroups. The model fit criteria were satisfied for collegiate dancers but not for traditional student-athletes. The failure of the configural model suggested that collegiate dancers and traditional student-athletes did not conceptualize sleep behavior similarly across the modified ASBQ items. Further testing should be done with larger samples from both groups to refute our results. Additionally, researchers may want to rewrite or modify items in the scale to better suit different populations.

The modified ASBQ was then subjected to invariance testing across injury status (ie, healthy versus injured). For those who self-reported being healthy, the model fit met some, but not all, model fit criteria; however, for those who were injured, the model fit met all contemporary model fit recommendations. The configural model indicated that the model fit criteria were satisfied when both groups were included in the model; yet the model fit indices for the metric invariance model were not met, reflecting an inconsistent factor structure.^{16,19} 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. The use of this scale in these groups (ie, injured or healthy) is not recommended without testing in another sample of collegiate athletes. Lastly, multigroup invariance testing was also performed across 2 levels of competition: NCAA Division I athletes versus 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 the model fit criteria were not satisfied.^{16,19}

Potential fit and design concerns should be considered beyond the multigroup invariance results despite the alternate model having an improved fit for most fit indices (ie, CFI, GFI, IFI, and RMSEA). First, Cronbach α values ranged from 0.47 to 0.52; nonetheless, these values were lower than the recommended values of ≥ 0.70 and ≤ 0.89 .¹⁶ The low internal consistency may have signified that the content of the items was too heterogeneous or the items were not relevant to the sample of individuals who responded.²⁶ Second, the TLI value was less than the recommended value (< 0.95),¹⁹ which may have been related to model misspecification from the combination of omitting cross-loadings²⁷ and low factor loadings (eg, item 13 with a loading of 0.21).³⁰ Model misspecification may introduce bias by not accounting for all parameters, correlations, or other pertinent values. Best-practice recommendations for survey item development (eg, avoiding double-barreled items and redundancy between items)¹⁵ and analytic procedures (eg, exploratory factor analysis to allow for an oblique solution and parallel analysis to determine factor retention)^{16,31} may be used to produce a more parsimonious and valid scale. These changes can allow for the scale to be more easily understood and consistently answered, which may result in improved model fit and a more precise assessment of sleep behaviors in collegiate athletes.

The cited 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 multigroup invariance testing recommendations. Still, other plausible explanations may help us understand the multigroup invariance testing results that indicated that the model fit for the modified ASBQ exceeded contemporary recommendations when tested in female respondents and collegiate dancers (who were also primarily women). The model failed to meet recommendations when tested in subgroups more heavily dominated by men or traditional student-athletes. Group differences for other variables (eg, grade point average [GPA] and academic preparation) may have affected how each group interpreted or responded to items, which might indicate that the item design was not effective for a specific subgroup and resulted in the subsequent failure of the model to meet multigroup invariance testing requirements conducted using other descriptive variables (eg, sport type). For example, researchers^{32,33} have demonstrated that women tended to maintain higher grades than their male counterparts and that nonathletes entered college with higher test scores than traditional athletes. In our study, men reported having a lower GPA (3.28) than women (3.57), and those who danced had a higher GPA (3.62) than traditional student-athletes (3.35; Table 3). Thus, their previous academic preparation may have influenced how respon-

dents in our sample understood or responded to items, which may mean that certain item characteristics (eg, double-barreled questions, item reading level, and item bias) influenced some groups more than others, producing increased response variation among certain subgroups. Therefore, the modified ASBQ might be sufficient to use in certain populations but not in others without further refinement to address survey design concerns.

Implementation in Clinical Practice and Research

Measuring sleep patterns in the collegiate athlete population is an important component in athletic training; however, we do not recommend using the original version of the ASBQ in the collegiate athletic population. The modified ASBQ may be a viable alternative with certain populations because of the improved model fit for several fit indices (ie, CFI, GFI, RMSEA, and IFI); yet caution is warranted as further research is needed to confirm in which subgroups and in which contexts (eg, repeated testing) their scores may be interpreted. For example, low factor loadings, poor internal consistency, a lower than recommended TLI (<0.95), and poor multigroup invariance testing findings raise concerns. The invariance testing results, along with the design of the items not following many recommended best-practice standards,¹⁵ provided 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, we did not perform other necessary steps in scale development, such as longitudinal invariance testing and the assessment of scale responsiveness. These should be completed to inform clinicians on how to use and interpret scale results before 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 the construct and total scores.

Furthermore, given the multifactorial nature of sleep,^{4,5} it is unlikely that this modified version successfully captures sleep behavior patterns in collegiate student-athletes. Hence, it would be prudent to develop an instrument that adequately assesses the multifactorial nature of sleep behaviors in athletes, regardless of sex, sport type, injury status, or collegiate competition level.

Limitations

Our study had several limitations. First, although the fit indices for the alternate model met most of the recommended standards, we were unable to assess responsiveness, measure test-retest reliability, or perform longitudinal invariance testing because data collection occurred at a single time point. Second, our large and diverse sample included a similar number of participants per subgroup (ie, collegiate dancers and traditional student-athletes), and more diversity in terms of sex, ethnicity, competition level for the student-athletes, and injury types would have been beneficial. The analyses we used also perform better with large sample sizes; thus, certain multigroup invariance testing procedures should be tested again in larger and more equally represented samples. For example, most (74.8%)

participants in our study self-reported as healthy; conducting multigroup invariance testing with larger samples of healthy and injured respondents with more similar 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 evaluated in this study may have affected the way the scale was interpreted between groups and should be considered by future researchers. Based on the measurement properties and theoretical design of the original ASBQ, we identified concerns with whether the scale adequately measures the proposed constructs; although 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 (ie, sleep behaviors, sleep routine, and sport), and a further criterion validity (eg, correlating scores on the constructs of the modified scale to other instruments or items thought to test the same) assessment is needed.

Third, we conducted our analysis of the ASBQ by using a reflective measurement model to remain consistent with earlier investigations^{12,17}; however, our evaluation of item content suggested the ASBQ may be better appraised using formative or mixed-model analysis.^{26,28} Therefore, future authors should consider both the individual items in the ASBQ and the best course of analysis (eg, reflective, formative, or mixed model) to create a scale that accurately assesses sleep behaviors in athletes. Our purpose was to replicate the original ASBQ testing¹² with the addition of CFA and multigroup invariance testing procedures; nonetheless, in future studies, researchers should consider using other recommended procedures (eg, exploratory factor analysis with oblique solutions, parallel analysis, or formative models) to test these items or develop new items to establish a psychometrically sound instrument to measure the multifactorial nature of sleep behaviors in collegiate athletes.^{15,19,24,33}

CONCLUSIONS

The CFA of the original 3-factor, 18-item ASBQ did not meet contemporary fit recommendations. We performed alternate model generation, which led to the creation of a 3-factor, 9-item model. Although the model fit was substantially improved, more examination is needed to ensure that 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 Jr, Zee PC, Walsh JK; Boards of Directors of the American Academy of Sleep Medicine and the Sleep Research Society. Sleep: a health imperative. *Sleep*. 2012;35(6):727–734. doi:10.5665/sleep.1846
2. Williamson A, Lombardi DA, Folkard S, Stutts J, Courtney TK, Connor JL. The link between fatigue and safety. *Accid Anal Prev*. 2011;43(2):498–515. doi:10.1016/j.aap.2009.11.011
3. Sitzler B. Sleep handout now available. National Athletic Trainers' Association. Published March 21, 2017. Accessed March 8, 2020. <https://www.nata.org/blog/beth-sitzler/sleep-handout-now-available>
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. doi:10.1080/17461391.2014.932016

5. Samuels C. Sleep, recovery, and performance: the new frontier in high-performance athletics. *Neurol Clin.* 2008;26(1):169–180, ix–x. doi:10.1016/j.ncl.2007.11.012
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(suppl 1):S43–S52. doi:10.1080/02640414.2012.703784
7. Lee A, Galvez JC. Jet lag in athletes. *Sports Health.* 2012;4(3):211–216. doi:10.1177/1941738112442340
8. O'Toole ML. Overreaching and overtraining in endurance athletes. In: Kreider RB, Fry AC, O'Toole ML, eds. *Overtraining in Sport.* Human Kinetics; 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 Sports Sci.* 2012;30(6):541–545. doi:10.1080/02640414.2012.660188
10. Sinnerton S, Reilly T. Effects of sleep loss and time of day in swimmers. In: MacLaren D, Reilly T, Lees A, eds. *Biomechanics and Medicine in Swimming: VI.* Taylor & Francis; 1992:399–404.
11. Reilly T, Edwards B. Altered sleep-wake cycles and physical performance in athletes. *Physiol Behav.* 2007;90(2–3):274–284. doi:10.1016/j.physbeh.2006.09.017
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 Sci.* 2018;11(1):37–44. doi:10.5935/1984-0063.20180009
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. doi:10.1016/j.jsams.2015.02.003
14. Johnson U. Athletes' experiences of psychosocial risk factors preceding injury. *Qual Res Sport Exerc Health.* 2011;3(1):99–115. doi:10.1080/19398441.2011.547690
15. Dillman DA, Smyth JD, Christian LM. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method.* 4th ed. John Wiley & Sons; 2014.
16. Leech NL, Barrett KC, Morgan GA. *IBM SPSS for Intermediate Statistics: Use and Interpretation.* 5th ed. Routledge; 2014.
17. Darendeli A, Diker G, Çınar Z. Athlete sleep behavior questionnaire—Turkish version: study of validity and reliability. *J Turk Sleep Med.* 2019;6(2):43–48. doi:10.4274/jtms.galenos.2019.58076
18. Bryant FB, Yarnold PR. Principal components analysis and exploratory and confirmatory factor analysis. In: Grimm LG, Yarnold PR, eds. *Reading and Understanding Multivariate Statistics.* American Psychological Association; 1995:99–136.
19. Kline RB. *Principles and Practice of Structural Equation Modeling.* 4th ed. Guilford Press; 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. doi:10.4085/1062-6050-355-17
21. Vela LI, Denegar CR. The disablement in the physically active scale, part II: the psychometric properties of an outcomes scale for musculoskeletal injuries. *J Athl Train.* 2010;45(6):630–641. doi:10.4085/1062-6050-45.6.615
22. Ullman JB. Structural equation modeling. In: Tabachnick BG, Fidell LS, eds. *Using Multivariate Statistics.* 4th ed. Pearson; 2001:653–771.
23. Tabachnick BG, Fidell LS. *Using Multivariate Statistics.* 4th ed. Allyn & Bacon; 2001.
24. Grimm LG, Yarnold PR. *Reading and Understanding Multivariate Statistics.* American Psychological Association; 1995.
25. Brown TA. *Confirmatory Factor Analysis for Applied Research.* 2nd ed. Guilford Press; 2015.
26. Byrne BM. *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming.* 2nd ed. Routledge; 2010.
27. Kim S. Testing a revised measure of public service motivation: reflective versus formative specification. *J Public Adm Res Theory.* 2011;21(3):521–546. doi:10.1093/jopart/muq048
28. Coltman T, Devinney TM, Midgley DF, Venaik S. Formative versus reflective measurement models: two applications of formative measurement. *J Bus Res.* 2008;61(12):1250–1262. doi:10.1016/j.jbusres.2008.01.013
29. 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. doi:10.1207/S15328007SEM1003_1
30. Shi D, Lee T, Maydeu-Olivares A. Understanding the model size effect on SEM fit indices. *Educ Psychol Meas.* 2019;79(2):310–334. doi:10.1177/0013164418783530
31. Matsunaga M. How to factor-analyze your data right: do's, don'ts, and how-to's. *Int J Psychol Res.* 2010;3(1):97–110. doi:10.21500/20112084.854
32. Turner RW II, Vissa K, Hall C, et al. Sleep problems are associated with academic performance in a national sample of collegiate athletes. *J Am Coll Health.* 2021;69(1):74–81. doi:10.1080/07448481.2019.1655027
33. Aries E, McCarthy D, 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. doi:10.1023/B:RIHE.0000040264.76846.e9

Address correspondence to Emilie N. Miley, DAT, ATC, University of Florida, 3450 Hull Rd, Gainesville, FL 32607. Address email to mileyen@ortho.ufl.edu.