Modeling the Structure of Acute Sport-Related Concussion Symptoms: A Bifactor Approach

Lindsay D. Nelson,[1] Mark D. Kramer,[2] Christopher J. Patrick,[3] AND Michael A. McCrea[1]

[1]Department of Neurosurgery, Medical College of Wisconsin, Milwaukee, Wisconsin
[2]Steve Hicks School of Social Work, University of Texas at Austin, Austin, Texas
[3]Department of Psychology, Florida State University, Tallahassee, Florida

Abstract

Objectives: Concussions cause diverse symptoms that are often measured through a single symptom severity score. Researchers have postulated distinct dimensions of concussion symptoms, raising the possibility that total scores may not accurately represent their multidimensional nature. This study examined to what degree concussion symptoms, assessed by the Sport Concussion Assessment Tool 3 (SCAT3), reflect a unidimensional versus multidimensional construct to inform how the SCAT3 should be scored and advance efforts to identify distinct phenotypes of concussion. **Methods:** Data were aggregated across two prospective studies of sport-related concussion, yielding 219 high school and college athletes in the acute (<48 hr) post-injury period. Item-level ratings on the SCAT3 checklist were analyzed through exploratory and confirmatory factor analyses. We specified higher-order and bifactor models and compared their fit, interpretability, and external correlates. **Results:** The best-fitting model was a five-factor bifactor model that included a general factor on which all items loaded and four specific factors reflecting emotional symptoms, torpor, sensory sensitivities, and headache symptoms. The bifactor model demonstrated better discriminant validity than the counterpart higher-order model, in which the factors were highly correlated ($r = .55–.91$). **Conclusions:** The SCAT3 contains items that appear unidimensional, suggesting that it is appropriate to quantify concussion symptoms with total scores. However, evidence of multidimensionality was revealed using bifactor modeling. Additional work is needed to clarify the nature of factors identified by this model, explicate their clinical and research utility, and determine to what degree the model applies to other stages of injury recovery and patient subgroups. (*JINS*, 2018, *24*, 793–804)

Keywords: Head injury, Concussion, Sports, Exploratory factor analysis, Confirmatory factor analysis, Bifactor model, Hierarchical model, Higher-order factor model, Psychometric properties, Clinical phenotypes

INTRODUCTION

Concussion gives rise to a wide array of somatic, cognitive, emotional, and other symptoms that persist for weeks or longer in a small subset of athletes and a sizable percentage of non-athlete patients (McCrea et al., 2003; McMahon et al., 2014). Symptom checklists, such as the Sport Concussion Assessment Tool 3 (SCAT3; McCrory et al., 2013) and the more recent SCAT5 (McCrory et al., 2017), are widely used in clinical settings to identify injury sequelae and have been proposed to be useful for detecting differences between patients in injury characteristics.[1]

In particular, several studies conducted in athlete and non-athlete settings have suggested that there are separable dimensions of concussion symptoms (e.g., cognitive, somatic, emotional; Ayr, Yeates, Taylor, & Browne, 2009; Joyce, Labella, Carl, Lai, & Zelko, 2015; Kontos et al., 2012; Merritt & Arnett, 2014; Meterko et al., 2012; Piland, Motl, Ferrara, & Peterson, 2003; Potter, Leigh, Wade, & Fleminger, 2006; Sady, Vaughan, & Gioia, 2014; Vanderploeg et al., 2015; Waljas et al., 2012) and that attending to these distinct facets of clinical presentation may reveal ways to tailor clinical management approaches to individual athletes (Collins, Kontos, Reynolds, Murawski, & Fu, 2014). Identifying distinct clinical phenotypes of concussion could lead to more precision medicine approaches to managing concussed patients. However, finding that multiple distinct constructs underlie concussion symptoms would call into question the validity of the widespread practice of summing symptom ratings into a single total score. Thus, evaluating the degree to which concussion symptoms are

---

Correspondence and reprint requests to: Lindsay Nelson, 8701 West Watertown Plank Road, Milwaukee, WI. E-mail: linelson@mcw.edu

[1] The name of the edition of the SCAT that was in use when most of the data were collected (i.e., SCAT3) is used throughout the study, but because the items, rating scale, and examinee instructions were unchanged from its predecessor, SCAT2, and equivalent to its successor, SCAT5, the findings are applicable to these other versions.

unidimensional or multidimensional is important both to validate how concussion symptom checklists are currently used and to inform the question of whether clinical management practices should shift toward more individualized approaches.

Although numerous studies have suggested the presence of distinct dimensions underlying concussion symptom reporting, there is reason to believe that the degree to which the dimensions identified are truly distinct has been overstated. For example, studies using latent trait models (typically lower- or higher-order factor models) and simpler data reduction strategies (principal component analysis [PCA]) with concussion symptom checklists have yielded models with marginal fit (Ayr et al., 2009; Joyce et al., 2015; Meterko et al., 2012), significant cross-loadings of items on different factors (Merritt & Arnett, 2014), a dominant first factor (Joyce et al., 2015; Kontos et al., 2012; Waljas et al., 2012), and highly correlated factors (Joyce et al., 2015; Piland et al., 2003; Potter et al., 2006; Sady et al., 2014; Vanderploeg et al., 2015). Furthermore, few studies have presented data to validate the factor structures identified, although the high inter-factor correlations suggest limited potential for purportedly distinct factors to demonstrate discriminant validity (e.g., Franke, Czarnota, Ketchum, & Walker, 2015). Taken together, these findings could indicate that concussion symptoms are essentially unidimensional, with some multidimensional features, and that alternative modeling approaches might better represent their underlying structure.

One alternative conceptualization of the structure of concussion symptoms that may better account for the evidence of both a dominant over-arching factor and multi-faceted subdimensions is a bifactor modeling approach. Figure 1 depicts a more traditional higher-order model in comparison to a bifactor model. These models account for covariance among item responses in different ways. In a higher-order model, item covariance is represented by an over-arching second-order factor (or factors) accounting for covariance between lower-order factors, with lower-order factors further parsed into responses to the items that load on these factors. In a bifactor model, each item is saturated by a general factor, meaning that a broad general factor directly influences scores on each item, and the residual variance unaccounted for by the general factor evinces structure such that certain items

make up specific factors representing correlated residuals among these items. The general factor and specific factors of the bifactor model are parameterized to be independent from one another, whereas correlations between lower-order factors of the higher-order model are accounted for by higher- or second-order factors.

Distinctions in interpretation of these models have implications for understanding the structure of concussion symptoms and for their clinical assessment. For example, the covariance accounted for by factors in a higher-order model between lower-order factors and their constituent items indicates that an individual scoring highly on the higher-order factor (or factors) would likely score highly on the lower-order factors loading on that dimension as well, which would also have bearing on other factors of the model (because they are allowed to correlate). In contrast, the independence of the general and specific factors of the bifactor model indicates that an individual's standing on the latent general factor has no bearing on their endorsement of symptoms marking subfactors of the model. In other words, a person with a low total score may actually score highly on one or another subfactor, and this endorsement pattern would be masked by relying solely on the general factor or total score. In the context of self-report symptom scales, the specific factors are often referred to as "method" factors, and these factors may represent constructs of limited substantive value (e.g., as tends to be the case for items that are similarly worded) or in some cases represent clinically relevant constructs in their own right (such as acquiescence or other forms of response bias; Campbell & Fiske, 1959).

Bifactor models have become popular in psychiatric and medical research (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012; Ebesutani et al., 2011; Gignac & Watkins, 2013; Patrick, Hicks, Nichol, & Krueger, 2007; Thomas, 2012) due to the practical and theoretical advantages of modeling broad traits of interest in these research domains (Reise, 2012; Reise, Morizot, & Hays, 2007; Thomas, 2012). For example, the Brief Symptom Inventory (BSI; Derogatis, 1993), a self-report measure of internalizing symptoms most commonly conceptualized as measuring correlated lower-order factors of depression, anxiety, and somatization, has been modeled in a bifactor framework. In particular, Thomas (2012) argued that the lack of consistent fit of the
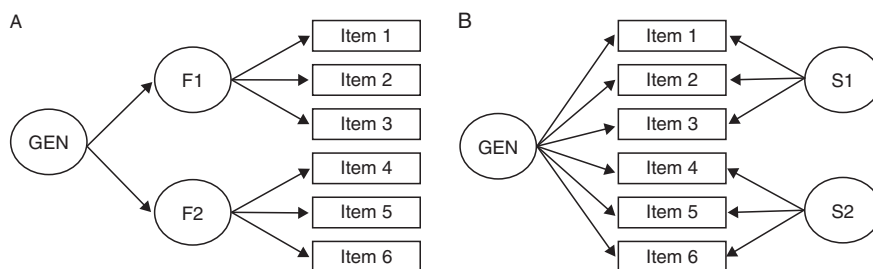
Fig. 1. (A) Higher-order factor model represented by lower-order factors (F1, F2) on which each item loads and a higher-order factor (GEN) representing a broader construct encompassed by the two correlated lower-order factors. (B) Bifactor model in which all items load on a single general factor (GEN), with two additional specific factors (S1, S2) that account for the residual covariance among items after extraction of the general factor.

higher-order model across studies and high between-factor correlations suggests that the BSI predominantly captures a single underlying dimension of internalizing symptoms/emotional distress. Modeling the BSI with a bifactor model formally acknowledges that a large general factor influences item responding and also yields domain-specific factors. That bifactor model factors better predict many DSM-IV-TR diagnoses as compared to those of the corresponding higher-order model suggests that the bifactor model may have value as an alternative way to conceptualize and measure aspects of BSI symptomatology (Thomas, 2012).

An example of another clinical construct for which bifactor modeling has proven valuable is psychopathy. Originally conceptualized as a unitary condition reflecting the presence of affective-interpersonal features (e.g., grandiosity, deceitfulness, shallow affect, superficial charm) coupled with impulsive-antisocial proclivities, this unitary concept has been difficult to reconcile with findings of highly distinct patterns of association for these two symptom components of psychopathy with other personality and behavioral measures. Bifactor modeling proved valuable in this case for separating out distinct trait-like subdimensions from general antisocial behavior (Patrick et al., 2007). This modeling work paved the way for a new model of psychopathy (Patrick, Fowles, & Krueger, 2009) that conceptualizes this condition in terms of core dispositions termed boldness (reflecting fearlessness), meanness (reflecting callous-unemotionality), and disinhibition (reflecting proneness to impulse control problems), with the former two constructs representing psychopathy-specific dimensions, and the latter a more general externalizing psychopathology dimension (Nelson & Foell, 2018).

Our objective was to investigate the factor structure of the SCAT3 symptom checklist to inform its clinical use and advance efforts to empirically derive distinct clinical phenotypes of concussion. We aggregated data from two recent prospective longitudinal studies of sport-related concussion in high school and collegiate athletes who were assessed in the acute (<48 hr) post-injury period, from which we derived and compared a range of plausible structural models (lower-order, higher-order, bifactor) to determine the best model and establish to what extent the SCAT3 captures a unidimensional versus multidimensional construct. This acute post-injury timepoint was the focus of this study given the rapid typical course of recovery from sport-related concussion and our goal to characterize the structure of concussion symptoms.

With the luxury of concussed patients who completed extensive clinical assessments before and after injury, we also evaluated relations between SCAT3 factors identified and various clinical assessment measures to provide preliminary validation of the models derived. We hypothesized that, in comparison to a higher-order factor model, a bifactor model would provide the best fit to the SCAT3 item set while also yielding factors with more distinctive relationships with other patient and injury variables. We predicted that the bifactor model would fit best based on (a) the evidence, reviewed above, that concussion symptoms appear to predominantly reflect a strong general factor and (b)

the fact that, for measurement reasons, bifactor models tend to be favored over higher-order models (Cucina & Byle, 2017; Mansolf & Reise, 2017). For this latter reason, we present results for both the bifactor model and a counterpart higher-order model, to be able to evaluate the two models in other ways apart from comparative fit.

METHODS

Participants

The sample consisted of participants from two prospective longitudinal studies of sport-related concussion: Project Head to Head 1 and 2. Project Head to Head 1 (2012–2014) recruited 2148 high school and collegiate contact sports athletes at preseason, of which 166 were followed serially after concussions (Nelson, LaRoche, et al., 2016; Nelson, Tarima, et al., 2016). Project Head to Head 2 (2015–present) is recruiting high school and collegiate football players; current analyses used data for 912 athletes (69 of whom had incurred concussions) enrolled at the time this analytic work was undertaken.

Datasets for these two samples were aggregated and trimmed of repeat cases, yielding $N = 219$ concussed individuals for current analyses. Both studies were approved by the Medical College of Wisconsin's institutional review board, and participants in each provided informed consent (or assent if under 18 years of age) before participation. Sample, injury, and recovery characteristics are listed in Table 1.

Concussion Definition and Characteristics

Both studies from which data were aggregated for this project used the same definition of concussion and common inclusion/exclusion criteria. In particular, concussion was defined based on that of the study sponsor (the U.S. Department of Defense): "an injury to the brain resulting from an external force and/or acceleration/deceleration mechanism from an event such as a blast, fall, direct impact, or motor vehicle accident which causes an alteration in mental status typically resulting in the temporally related onset of symptoms such as headache, nausea, vomiting, dizziness/balance problems, fatigue, insomnia/sleep disturbances, drowsiness, sensitivity to light/noise, blurred vision, difficulty remembering, and/or difficulty concentrating" (Helmick et al., 2006).

Diagnoses were made by athletic trainers at participating schools and confirmed by the research staff at the acute (24–48 hr post-injury) assessment. Only four participants underwent clinical brain imaging following their injuries (all head CTs), and none indicated the presence of any acute intracranial abnormalities. The acute injury characteristics were distributed as follows: 5.5% loss of consciousness, 11.5% posttraumatic amnesia, and 7.4% retrograde amnesia. Duration of unconsciousness, when present, was $M = 29$ s ($SD = 36$; range 2 s to 2 min). Duration of posttraumatic

Table 1. Sample demographics and injury characteristics (N = 219)

	N (%) or M (SD)
Male (vs. female)	194 (88.6%)
College (vs. high school)	141 (64.4%)
Age	17.6 (1.9)
Race	
White	175 (79.9%)
Black	38 (17.4%)
Asian	2 (0.9%)
Not reported	5 (1.8%)
Sport	
Football	164 (74.9%)
Soccer	36 (16.4%)
Lacrosse	7 (3.2%)
Wrestling	4 (1.8%)
Ice hockey	4 (1.8%)
Rugby	3 (1.4%)
Field hockey	1 (0.5%)
Years of participation in study sport	8.01 (3.94)
Prior concussion(s)	124 (57.1%)
ADHD	22 (10.1%)
Learning disability	6 (2.8%)
Psychiatric disorder	2 (0.9%)
Acute injury characteristics	
Loss of consciousness	12 (5.5%)
Posttraumatic amnesia	25 (11.5%)
Retrograde amnesia	16 (7.4%)
SCAT3 Symptom Severity	
Baseline (pre-season)	5.9 (9.7)
24–48 hours	24.5 (19.2)
Day 8	6.8 (13.7)
Day 15	3.0 (7.9)
Day 45	1.7 (3.9)
Symptom duration (days)	6.0 (median)

Note. Within the football cohort, participants were distributed by position as follows: 25% offensive line, 23% defensive line, 13% linebacker, 13% wide receiver, 12% defensive back, 8% running back, 6% quarterback, and 1% kicker. The 36 soccer players were distributed by position as follows: 31% midfield, 28% defense, 25% forward, and 17% goalkeeper. ADHD = attention deficit hyperactivity disorder.

amnesia, when present, was $M = 46$ min ($SD = 97$ min; range 1 min to 7 hr).

Baseline and Post-Injury Test Battery

Athletes in both studies completed pre-season baseline examinations and post-injury clinical assessments at 24–48 hr and 8, 15, and 45 days post-injury. The baseline assessments for each included demographic and health history questionnaires and the Wechsler Test of Adult Reading (Wechsler, 2001). Pre- and post-injury assessments also included a neuropsychological assessment battery including the Sport Concussion Assessment Tool 3 (SCAT3) symptom checklist (McCrory et al., 2013), the Standardized Assessment of Concussion (McCrea et al., 1998), the Balance Error Scoring System (Guskiewicz, Ross, & Marshall, 2001), and the Brief Symptom Inventory – 18 (BSI-18; Derogatis, 2001).

Additional assessment measures were available for a subset of the combined sample: 168 (78%) athletes completed a computerized neurocognitive test battery, the Immediate Postconcussion and Cognitive Testing (ImPACT, online version; ImPACT Applications Inc.), at each assessment; 24 (11%) took the King-Devick Test (Galetta et al., 2011); and 66 (30%) took additional self-report measures of personality and emotional functioning (the latter two types of measures were only in the second study protocol). Relevant to the current study, measures of the latter type included a brief (29-item) version of the Multidimensional Personality Questionnaire (which provides estimates of trait negative [NEM] and positive emotionality [PEM]; Patrick, Curtin, & Tellegen, 2002) and the following somatic and cognitive symptom subscales of the adult and adolescent Restructured Forms of the Minnesota Multiphasic Personality Inventory (i.e., MMPI-2-RF and MMPI-2-A): Somatic Complaints (RC1), Malaise, Gastrointestinal Complaints, Head Pain Complaints, Neurologic Complaints, and Cognitive Complaints (Archer, Handel, Ben-Porath, & Tellegen, 2014; Ben-Porath & Tellegen, 2008). Participants who were at least 18 years old at the preseason assessment (78% of the Project Head to Head 2 sample) completed the MMPI-2-RF; those below age 18 completed the MMPI-2-A. The numbers of items shared between counterpart scales of the two test forms are: 19/30 RC1, 8/8 Malaise, 4/5 Gastrointestinal Complaints, 4/6 Head Pain Complaints, 7/10 Neurologic Complaints, 4/11 Cognitive Complaints. Test manual norms were used to convert raw scores to uniform T scores before aggregating scores across the two test forms for analysis.

Statistical Analysis

Our primary goal was to investigate the factor structure of the 22-item SCAT3 symptom checklist, focusing on the acute (24–48 hr post-injury) time point given that this is when symptom severity (and thus variance in reported symptomatology, and potential covariance among differing symptoms) is generally the highest. First, we reviewed item-level descriptive statistics (Table 2) and inter-item correlations (see online Supplementary Table). The Kaiser-Meyer-Olkin coefficient and Bartlett's test, computed in IBM SPSS Statistics for Windows (Version 24), were used to test for sampling adequacy and sphericity, respectively.

All factor analyses were conducted in *Mplus* (7th edition; Muthén & Muthén, 1998–2015) using WLSMV estimation (given the ordinal nature of the items). Geomin rotation, an oblique rotation method, was used in factor analyses that allowed factors to be correlated. First, we conducted exploratory factor analyses (EFA) and exploratory bifactor analyses (bi-EFA) to inform the specification of confirmatory factor analytic (CFA) models. Inspection of the eigenvalues and the scree plot from an initial EFA indicated the presence of a dominant first factor (eigenvalues one–five were 11.79, 1.62, 1.14, .92, and .79, respectively), with a subtle break

Table 2. SCAT3 item-level descriptive statistics (24–48 hours post-injury)

	M	*SD*	Median	% Endorsed	Skewness	Kurtosis
Headache	2.46	1.66	2	83.6%	.02	−1.11
Pressure in head	1.82	1.51	2	75.8%	.51	−.67
Neck pain	1.17	1.41	1	52.8%	.98	−.16
Nausea or vomiting	.54	.98	0	30.9%	1.97	3.17
Dizziness	.92	1.21	0	47.3%	1.27	1.05
Blurred vision	.52	.97	0	30.0%	2.27	5.41
Balance problems	.65	1.00	0	38.6%	1.60	1.79
Sensitivity to light	1.23	1.40	1	56.8%	1.01	.21
Sensitivity to noise	1.00	1.32	0	47.5%	1.29	1.01
Feeling slowed down	1.74	1.64	1	69.5%	.62	−.75
Feeling like in a fog	1.42	1.47	1	62.7%	.87	.00
Don't feel right	1.83	1.55	2	73.2%	.48	−.64
Difficulty concentrating	1.70	1.57	1	69.5%	.63	−.48
Difficulty remembering	1.01	1.25	1	52.7%	1.37	1.89
Fatigue	1.81	1.65	2	68.2%	.47	−.91
Confusion	.59	.98	0	33.2%	1.69	2.33
Drowsiness	1.45	1.52	1	60.9%	.80	−.21
Trouble falling asleep	.67	1.33	0	25.9%	2.07	3.45
More emotional	.47	1.02	0	22.7%	2.31	4.71
Irritability	.60	1.09	0	32.3%	2.14	4.68
Sadness	.34	.73	0	21.8%	2.29	4.64
Nervous/anxious	.51	.98	0	28.2%	2.30	6.15

Note. "Endorsement" was operationalized as any non-zero rating (i.e., rating of 1–6).

evident in the scree plot between the second and third eigenvalues.

Given prior work suggesting the presence of more than two factors (often three–four) underlying responses on similar symptom checklists, we explored a range of factor solutions (one–five lower-order factors for higher-order models; one–five specific factors for bifactor EFA models) and used a combination of model fit, factor interpretability, and factor external correlates to select the optimal model(s).

Next, CFAs were run for higher-order and bifactor models configured according to results from the exploratory analyses. A four-factor higher-order model was not included at this stage, as the four-factor EFA included a factor on which all items loaded more strongly on other factors. For the higher-order models (involving one, two, three, or five factors), each item was specified as loading on the factor on which it loaded most highly in the corresponding EFA. In all five higher-order EFA models, all items but four loaded at least .4 on a factor, with loadings above .29 for the remaining four. For bifactor models, all items were parameterized to load on the general factor, with items parameterized to load as well on a specific factor if they loaded at least .30 on that factor in the bi-EFA. A small number of items that met this criterion were ultimately dropped from the final CFA models, given that they were not conceptually related to other items associated with the specific factor, and did not improve model fit or load significantly on the specific factor at the CFA stage.

Model-based reliability measures of omega, omega hierarchical (omegaH), and relative omega (omega/omegaH)

were computed for the best-fitting (five-factor bifactor) model, with a particular focus on omegaH (and relative omega) as an estimate of the proportion of observed variance (or proportion of reliable variance, in the case of relative omega) in total and subscale scores that could be attributed to the underlying general and specific factors (Rodriguez, Reise, & Haviland, 2016). Finally, regression-based factor scores for models of major interest (i.e., five-factor higher-order factor model and five-factor bifactor model) were computed for use in analyses examining associations of model factors with a range of clinical criterion measures available from the pre- and post-injury assessments.

RESULTS

Item-Level Descriptive Analyses

Item-level descriptive statistics are presented in Table 2, while inter-item correlations are in the Supplementary Table. Across all items, endorsement frequencies (i.e., instances of a non-zero rating) ranged from 21.8% (sadness) to 83.6% (headache), with item medians ranging from zero to two. Correlations between items (rho coefficients) ranged from .08 to .78, and all items correlated at least .29 with one or more items. Values of the Kaiser-Meyer-Olkin coefficient (.92) and Bartlett's test of sphericity ($\chi^2[231] = 2995.09$; $p < .001$) were clearly acceptable and suggestive of sufficient shared variance among items to conduct factor analyses.

Confirmatory Factor Analyses

Higher-order models

All CFA model fit statistics are presented in Table 3. The two-factor model separated items with emotional content (emotional, irritability, sadness, nervousness/anxiousness) from all other items, with a correlation between factors of $r = .67$. This emotional factor remained robust in solutions with more factors, whereas other concussion symptoms were split up to varying degrees across factors in other models. Beyond the clear emotional factor, other factors were often difficult to interpret and were very highly correlated ($r = .76–.91$).

Model fit improved with extraction of more factors (e.g., comparative fit index [CFI] values for the one-, two-, three-, and five-factor higher-order solutions were .939, .967, .975, and .981, respectively). The high factor inter-correlations suggested that a higher-order factor explaining the covariance among items might fit well, or that the measure might be better represented by a bifactor model. For the five-factor model, we explicitly modeled a higher-order factor as accounting for the covariance among the five lower-order factors, which improved model fit by a nominal amount (CFI went from .981–.982).

Bifactor models

Fit statistics for the bifactor CFA model are shown in Table 3. Loadings for the two- and five-factor models appear in Table 4. All solutions contained a specific factor (S1) reflecting covariance among the residuals for the four emotional-content items. Additional specific factors (S2, S3, S4) reflected torpor (feeling slowed down, feeling as if in a "fog," don't feel right, fatigue; four items), sensory sensitivity (i.e., to light and noise; two items), and headache/pressure in head (two items). The five-factor bifactor model (depicted in Figure 2) fit best among the bifactor solutions and exhibited modest improvement over the best-fitting higher-order model (CFI = .989 *vs*. .982 for the five-factor higher-order model). The six-factor bifactor model split the four-item torpor specific factor into two factors, which slightly attenuated model fit (CFI = .987). In each of these bifactor models, all 22 SCAT3 items loaded substantially on the general factor (GEN; ∼ .5 or higher; see Table 4), providing evidence for considerable unidimensionality to the scale.

Values of omega for GEN, S1 (emotional), S2 (torpor), S3 (sensory sensitivities), and S4 (headache) were .97, .89, .94, .87, and .83, respectively. OmegaH for the factors was .94, .45, .12, .33, .28, respectively, and values of relative omega were .96, .51, .13, .38, .33. The very high relative omega value for the general factor suggests that most (96%) of the reliable variance in the SCAT3 can be attributed to this factor and that the correlation between the general factor and observed SCAT3 symptom severity scores is nearly perfect (sqrt[omegaH] = .97) (Rodriguez et al., 2016).

External Correlates of Factors

To examine the potential clinical value of the CFA models specified by our analyses, we computed correlations between factor scores for the best-fitting (five-factor) higher-order and bifactor models and available clinical criterion measures; these correlations are shown in Table 5. Of note, although the five-factor higher-order and bifactor models evidenced the best fit, a concern with each of these models is the possibility of over-extraction of factors. To address this issue, we specified alternative two-factor versions of the higher-order and bifactor models. The factors that were shared across two- and five-factor higher-order models were very highly correlated across model types ($rs > .99$ in each case), and, therefore, we present data on the external correlates for only the five-factor models.

As could be expected from the high correlations among factors in the higher-order model, there was substantial overlap in the clinical correlates of the five factors from this model. In contrast, the bifactor model provided higher specificity, with scores on the general factor correlating robustly with self-report (somatic symptoms, distress) and performance-based (cognitive, balance) clinical criterion

Table 3. CFA model fit statistics

	χ^2	k	RMSEA (90% CI)	CFI	TLI
One-factor/higher-order models					
1-factor	693.03	141	.103 (.094, .011)	.939	.933
2-factor	473.99	142	.076 (.067, .086)	.967	.963
3-factor	403.72	144	.066 (.057, .076)	.975	.972
5-factor	353.92	151	.060 (.049, .070)	.981	.977
5-factor + G	350.35	146	.057 (.047, .067)	.982	.979
Bifactor models					
2-factor (1 G, 1 S)	435.14	145	.072 (.062, .081)	.971	.967
3-factor (1 G, 2 S)	417.10	143	.068 (.059, .078)	.974	.970
4-factor (1 G, 3 S)	374.59	144	.061 (.051, .071)	.979	.976
5-factor (1 G, 4 S)	287.44	151	.045 (.033, .056)	.989	.987
6-factor (1 G, 5 S)	301.94	149	.048 (.036, .059)	.987	.985

Note. CFI = comparative fit index; G = general factor; S = specific factor; RMSEA = root mean square error of approximation; TLI = Tucker-Lewis index. All χ^2 *p*-values were < .001. k = number of free parameters.

Table 4. Loadings for selected higher-order and bifactor CFA models

| | Higher-order models | | | | | | | Bifactor models | | | | | | |
| | 2-Factor | | 5-Factor | | | | | 2-Factor | | 5-Factor | | | | |
	F1	F2	F1	F2	F3	F4	F5	GEN	S1	GEN	S1	S2	S3	S4
Headache	.72						.86	.72		.70				.48
Pressure in head	.72						.84	.71		.69				.48
Neck pain	.48					.51		.49		.50				
Nausea or vomiting	.65				.75			.65		.67				
Dizziness	.78					.83		.78		.79				
Blurred vision	.62					.66		.62		.63				
Balance problems	.76			.77				.76		.77				
Sensitivity to light	.73				.83			.73		.69			.54	
Sensitivity to noise	.74				.84			.74		.70			.54	
Feeling slowed down	.89			.90				.89		.82		.49		
Feeling like in a fog	.87			.88				.87		.83		.27		
Don't feel right	.86			.87				.86		.82		.27		
Difficulty concentrating	.88			.89				.88		.91				
Difficulty remembering	.71					.75		.71		.73				
Fatigue	.85					.86		.85		.78		.21		
Confusion	.81			.86				.81		.83				
Drowsiness	.82			.83				.82		.78				
Trouble falling asleep	.59					.67		.59		.60				
More emotional		.87	.87					.54	.67	.56	.68			
Irritability		.83	.83					.60	.41	.61	.40			
Sadness		.85	.85					.50	.85	.51	.85			
Nervous/anxious		.67	.77					.57	.37	.58	.36			

Note. F = factor within a higher-order model; S = specific factor within a bifactor model. In the two-factor higher-order model, $r = .67$ between factors. In the five-factor higher-order model, $r = .55–.67$ between F1 and the other factors, while correlations between F2–F5 ranged from $r = .78–.91$. The loadings of F1–F5 on a higher-order general factor were .83, .95, .85, .95, and .69, respectively. Factors in the bifactor model are uncorrelated by design.

measures assessed at the same acute post-injury time point. The implication is that the general factor best captured the acute clinical effects of concussion, consistent with the intended purpose of this inventory as a whole. However, this factor also correlated positively with pre-injury ratings of somatic symptoms and general distress (as assessed by the BSI-18), as well as with time to complete the King-Devick Test of oculomotor functioning at baseline.

In contrast, the headache specific factor (S4) correlated in a selective manner with acute post-injury neurocognitive impairment as indexed by ImPACT, and showed a single significant correlation with one self-report measure (baseline MMPI-2-RF Malaise). The emotional specific factor (S1) correlated most strongly with ratings of pre-injury emotional distress (BSI-18 anxiety/depression), and the torpor specific factor (S2) correlated to a modest degree with pre-injury ratings of agentic PEM. The sensory sensitivity specific factor (S3) was relatively uncorrelated with these clinical criterion measures.

DISCUSSION

The current work provides a novel approach to representing the structure of concussion symptoms. In contrast to prior investigations using only conventional EFA, CFA, or PCA approaches in conceptualizing the structure of concussion, we present evidence that a bifactor modeling approach could accommodate findings of both a dominant general factor and multidimensionality underlying this measure. In particular, we demonstrated that the vast majority of variance in concussion symptom reporting is accounted for by a single general factor that appears to influence ratings on all 22 items of the SCAT3 checklist. This finding, alongside the strong fit of a bifactor model, suggests that the SCAT3 is essentially unidimensional and, accordingly, that it is appropriate for clinician users of the scale to continue to use total scores when a single index of overall concussion symptom severity is desired.

On the other hand, we also found that representing the underlying structure of SCAT3 symptoms through a single factor is insufficient, as there is substantial covariation in symptom reporting not accounted for by a single factor alone. The bifactor model presented suggests that there are additional constructs beyond the general factor that influence concussion symptom reporting. Additional investigation is needed to clarify the clinical and research utility of these specific factors. Given consistency across the iterations, these findings also apply to the newer SCAT5.

Additional research is needed to understand the constructs represented by the specific factors identified in our best-fitting bifactor model. Given the similarities in the content of
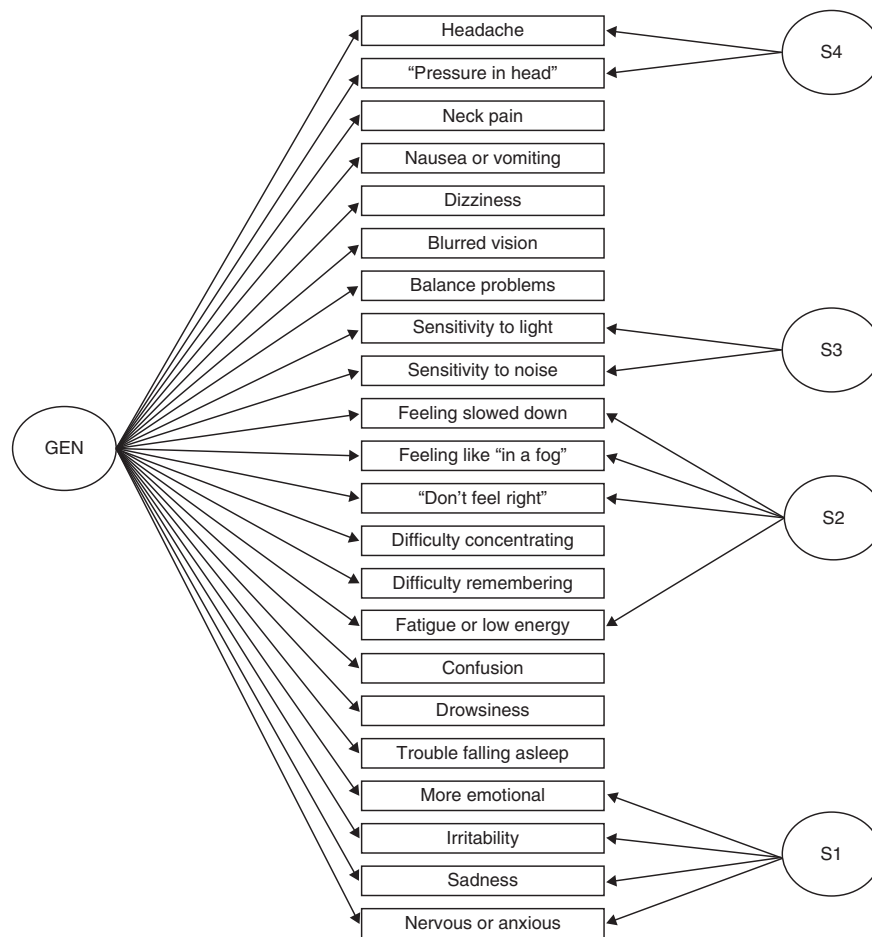
Fig. 2. Structure of the best-fitting bifactor model of acute concussion symptoms, derived from Sport Concussion Assessment Tool 3 (SCAT3) symptom ratings at 24-48 hr postconcussion.

items comprising the sensory sensitivity and headache (S3 and S4) factors, one might wonder if these two factors in particular simply reflect non-substantive "method" factors. If this proves to be the case, redundant items could be merged or eliminated to achieve a more abbreviated and empirically informed self-report measure of the general factor that underlies concussion symptoms. Work to investigate this question should be informed by modern test theory (item response theory) to ensure the retention of items that precisely measure concussion symptom severity across the continuum.

In addition, reliability estimates for the specific factors (especially S2–S4) were rather low, which might call into question their utility. Nonetheless, specific factors did manifest significant relationships with various criterion measures, such as injury-related neurocognitive impairment (S4) and pre-injury emotional distress (S1), suggesting that they could represent meaningful and distinct constructs. The finding that the emotional specific factor loaded preferentially on pre-injury distress suggests that the bifactor model may provide a useful way to explicitly measure dispositional factors that affect concussion symptom reporting from post-injury data (a challenge in the civilian TBI literature where pre-injury baseline assessments are not readily obtained).

Future research should explicate the constructs underlying these specific factors and determine if they can be more reliably measured with adjustments to the model or inclusion of additional items or variables within the model.

We also demonstrated that SCAT3 symptom ratings can alternatively be represented through a (more traditional) multidimensional higher-order model, which some readers may have theoretical reasons for favoring over a bifactor model. Our findings illustrate, however, that when distinct dimensions of the SCAT3 are represented by a higher-order model, their correlations are so high that the practical value of parsing them apart is questionable. In particular, we demonstrated that factors within the higher-order model have limited discriminant validity. The high correlation of factors in this model, consistent with other work (Joyce et al., 2015; Potter et al., 2006; Sady et al., 2014; Vanderploeg et al., 2015), suggests that researchers should avoid statistically forcing lower-order factors to be uncorrelated (i.e., through orthogonal rotation), as it parses symptoms that are highly coherent and thus properly represented by correlated dimensions (as in the higher-order model of our study) or a broad common factor (as in the bifactor model we presented). In contrast to a bifactor model, these types of models may mask the importance/dominance of a single, broad dimension

Table 5. External correlates of factors from the best-fitting higher-order and bifactor models

	Higher-order model						Bifactor model				
	GEN	F1	F2	F3	F4	F5	GEN	S1	S2	S3	S4
Pre-season baseline											
BSI-18 Somatization	**.39**	**.38**	**.40**	**.38**	**.36**	**.38**	**.39**	.12	−.10	.08	.12
BSI-18 Depression	**.20**	**.29**	**.21**	**.16**	**.18**	**.20**	**.20**	**.22**	−.08	−.04	.07
BSI-18 Anxiety	**.24**	**.28**	**.24**	**.18**	**.23**	**.26**	**.24**	**.18**	−.02	−.07	.12
BSI-18 Global severity	**.33**	**.38**	**.34**	**.28**	**.30**	**.34**	**.33**	**.21**	−.08	−.02	.13
MPQ NEM	.16	.16	.17	.15	.15	.12	.16	.03	.00	.03	−.07
MPQ PEM	−.05	−.05	−.06	−.04	−.04	−.12	−.06	.01	.22	.07	−.13
MPQ Agentic PEM	−.05	−.04	−.06	−.06	−.02	−.12	−.06	.03	**.28**	.00	−.14
MPQ Communal PEM	−.13	−.11	−.13	−.11	−.13	−.13	−.15	.02	.10	.09	.00
MMPI-2-RF RC1	.08	.05	.08	.07	.07	.17	.08	−.03	−.11	−.09	.21
MMPI-2-RF Malaise	.13	.15	.10	.13	.12	**.26**	.11	.08	.05	.02	**.31**
MMPI-2-RF GIC	.10	.04	.08	.13	.12	.09	.10	.04	.15	.04	−.02
MMPI-2-RF HPC	.10	.10	.10	.06	.09	.16	.10	.07	−.02	−.17	.15
MMPI-2-RF NUC	.14	.06	.11	.17	.14	.19	.14	−.11	−.02	.07	.14
MMPI-2-RF COG	−.07	−.09	−.06	−.09	−.08	.03	−.06	−.04	−.16	−.15	.19
SAC	.03	.04	.01	.00	.05	.06	.02	.05	**.15**	−.07	.09
BESS	−.03	.01	−.03	−.01	−.04	.01	−.04	.07	−.03	.05	.09
ImPACT VERM	−.14	−.04	**−.17**	**−.17**	−.13	−.09	−.15	.16	.11	−.11	.09
ImPACT VISM	−.11	.01	−.13	−.16	−.07	−.10	−.11	**.18**	.17	−.14	−.02
Im VMS	**.19**	**.19**	**.19**	.15	**.19**	**.18**	**.18**	.11	.06	−.02	.05
Im RT	−.08	−.03	−.09	−.04	−.09	−.08	−.09	.01	.02	.13	−.02
King-Devick Test	**.41**	**.45**	**.43**	.25	**.40**	**.34**	**.41**	.19	.10	−.23	−.08
24-48 hours post-injury											
BSI-18 Somatization	**.38**	**.28**	**.38**	**.31**	**.39**	**.32**	**.38**	−.03	.10	**−.18**	−.02
BSI-18 Depression	**.23**	**.28**	**.22**	**.22**	**.24**	**.18**	**.23**	**.18**	.14	.02	−.02
BSI-18 Anxiety	.14	**.19**	.14	**.15**	.14	.06	**.15**	.09	.04	.01	**−.14**
BSI-18 Global severity	**.32**	**.32**	**.32**	**.29**	**.33**	**.24**	**.32**	.10	.12	−.06	−.08
MMPI-2-RF RC1	**.39**	**.35**	**.39**	**.33**	**.38**	**.32**	**.40**	.06	.03	−.01	−.07
MMPI-2-RF Malaise	**.27**	.14	**.29**	.20	**.29**	.16	**.29**	−.13	.02	−.11	−.21
MMPI-2-RF GIC	.15	.16	.16	.11	.15	.12	.17	.00	−.02	−.21	−.07
MMPI-2-RF HPC	**.32**	.17	**.31**	**.33**	**.32**	**.33**	**.33**	−.15	−.08	.17	.07
MMPI-2-RF NUC	.20	.20	.24	.12	.20	.13	.20	.10	.07	−.16	−.14
MMPI-2-RF COG	**.42**	**.39**	**.43**	**.37**	**.41**	**.35**	**.42**	.12	−.02	.05	−.08
SAC	**−.28**	**−.24**	**−.29**	**−.25**	**−.28**	**−.29**	**−.27**	−.05	−.05	−.02	−.13
BESS	**.17**	.13	**.17**	.11	**.17**	**.18**	**.18**	−.07	−.05	−.11	.08
ImPACT VERM	**−.27**	**−.23**	**−.24**	**−.26**	**−.26**	**−.37**	**−.26**	−.06	.00	−.07	**−.31**
ImPACT VISM	**−.23**	−.16	**−.20**	**−.20**	**−.24**	**−.29**	**−.23**	.04	−.03	.06	**−.21**
Im VMS	**−.41**	**−.30**	**−.39**	**−.36**	**−.40**	**−.46**	**−.40**	−.03	.00	−.03	**−.26**
Im RT	**.24**	.16	**.23**	.15	**.25**	**−.29**	**.23**	.02	.09	−.13	**.19**
King-Devick Test	.43	.48	.43	.35	.42	.49	.45	−.01	.01	−.13	.12

Note. Bold where *p* < .05. The sample size with available King-Devick Test data was small (*n* = 16 for pre-injury and *n* = 14 for post-injury measures). 24–48-hour assessment variables = baseline difference scores. BESS = Balance Error Scoring System; BSI-18 = Brief Symptom Inventory 18-item; COG = Cognitive Complaints scale; GIC = Gastrointestinal Complaints scale; GSI = Global Severity Index; HPC = Head Pain Complaints scale; ImPACT = Immediate Postconcussion and Cognitive Testing; MMPI-2-RF = Minnesota Multiphasic Personality Inventory-2-Restructured Form; MPQ = Multidimensional Personality Questionnaire; NEM = negative emotionality; NUC = Neurologic Complaints scale; PEM = positive emotionality; RC1 = RC1 Somatic Complaints scale; RT = Reaction Time composite; SAC = Standardized Assessment of Concussion; VERM = Verbal Memory composite; VISM = Visual Memory composite; VMS = Visual Motor Speed composite

accounting for large proportions of variance in the domain. Irrespective of identification of a best-fitting model, ultimately, the true test of a model's utility lies in demonstrating that it advances scientific understanding or informs clinical practice.

Given that SCAT3 symptoms were assessed in the acute postconcussive period in our study sample and correlated robustly with other subjective and objective markers of injury, one might draw the obvious conclusion that the general factor specified in the bifactor model represents concussion symptoms. With this in mind, it is interesting that the general factor also correlated robustly with several *pre-injury* clinical variables, including pre-injury baseline somatic and distress symptoms. This finding is consistent with a wealth of

evidence indicating a direct link between pre-injury emotional functioning and response to concussion which in turn points to the need for more investigation into the mechanisms by which pre-injury risk and resilience factors influence patients' responses to injury and courses of recovery (Broshek, De Marco, & Freeman, 2015; Nelson, Tarima, et al., 2016; Root et al., 2016).

Some limitations of the current study are important to acknowledge. One is that the sample size, although large for a prospective study of acute concussion, was modest for a factor analytic study. A second limitation is that, given the sample size, we did not evaluate the applicability of models specified in the current work to other post-injury time points or test for their emergence in subgroups posited to vary in symptom reporting and recovery (e.g., differing age or gender groups; Dick, 2009; Field, Collins, Lovell, & Maroon, 2003). Ongoing multicenter studies of sport-related concussion (Broglio et al., 2017) will provide opportunities to cross-validate findings from the current study in larger samples and to extend them to different stages of recovery and patient groups. Consideration of alternative quantitative modeling approaches would also be worthwhile, such as those designed to identify subgroups of patients with qualitatively distinct patterns of symptom presentation (Muthen & Asparouhov, 2006). Additionally, consideration of a larger item pool beyond those of the SCAT3 would be useful to enable identification of factors for which there were too few items in this scale alone. Finally, some questionnaire measures used as criterion measures in this study (e.g., scales from the MMPI-2-RF) were administered in isolation rather than as part of the entire measure as they were originally validated. This could diminish the validity of these scales and, by extension, the interpretability of correlations between these measures and the SCAT3 factors.

CONCLUSION

The current study provides evidence that the symptom checklist on the SCAT3 (and its more recent iteration, the SCAT5) captures an essentially unidimensional construct related to concussion symptom reporting, suggesting that current practices to sum SCAT3/5 items into a single severity score are appropriate. We also presented preliminary evidence that a bifactor model provides a useful way to identify distinct dimensions that contribute to SCAT3/5 symptoms, although additional research is needed to establish the clinical utility of the specific factors identified and to refine their measurement.

Future work examining the stability of this model across different post-injury time points and patient groups will be useful for validating the structure and psychometric properties of the SCAT3/5 for a range of clinical applications. Additional studies comparing a broader range of structural models and incorporating clinical and neurobiological variables into these models may reveal distinct phenotypes of concussion with relevance to clinical management decisions and translational research.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1355617718000462

REFERENCES

Archer, R.P., Handel, R.W., Ben-Porath, Y.S., & Tellegen, A. (2014). *MMPI-A-RF Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.

Ayr, L.K., Yeates, K.O., Taylor, H.G., & Browne, M. (2009). Dimensions of postconcussive symptoms in children with mild traumatic brain injuries. *Journal of the International Neuropsychological Society*, *15*(1), 19–30. doi: 10.1017/S1355617708090188

Ben-Porath, Y.S., & Tellegen, A. (2008). *Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF)*. Minneapolis, MN: University of Minnesota Press.

Broglio, S.P., McCrea, M., McAllister, T., Harezlak, J., Katz, B., & Hack, D., . . . Investigators, Care Consortium. (2017). A national study on the effects of concussion in collegiate athletes and US military service academy members: The NCAA-DoD Concussion Assessment, Research and Education (CARE) Consortium Structure and Methods. *Sports Medicine*, *47*(7), 1437–1451. doi: 10.1007/s40279-017-0707-1

Broshek, D.K., De Marco, A.P., & Freeman, J.R. (2015). A review of post-concussion syndrome and psychological factors associated with concussion. *Brain Injury*, *29*(2), 228–237. doi: 10.3109/02699052.2014.974674

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105.

Chen, F.F., Hayes, A., Carver, C.S., Laurenceau, J.P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219–251. doi: 10.1111/j.1467-6494.2011.00739.x

Collins, M.W., Kontos, A.P., Reynolds, E., Murawski, C.D., & Fu, F.H. (2014). A comprehensive, targeted approach to the clinical care of athletes following sport-related concussion.

Knee Surgery, Sports Traumatology, Arthroscopy, *22*(2), 235–246. doi: 10.1007/s00167-013-2791-6

Cucina, J., & Byle, K. (2017). The bifactor model fits better than the higher-order model in more than 90% of comparisons of mental abilities test batteries. *Journal of Intelligence*, *5*, 27.

Derogatis, L.R. (1993). *BSI, Brief Symptom Inventory: Administration, scoring, and procedures manual* (4th ed.). Minneapolis, MN: National Computer Systems.

Derogatis, L.R. (2001). *Brief Symptom Inventory 18 (BSI-18): Administration, scoring, and procedures manual*. Bloomington, MN: Pearson.

Dick, R.W. (2009). Is there a gender difference in concussion incidence and outcomes? *British Journal of Sports Medicine*, *43* (Suppl. 1), i46–i50. doi: 10.1136/bjsm.2009.058172

Ebesutani, C., Smith, A., Bernstein, A., Chorpita, B.F., Higa-McMillan, C., & Nakamura, B. (2011). A bifactor model of negative affectivity: Fear and distress components among younger and older youth. *Psychological Assessment*, *23*(3), 679–691. doi: 10.1037/a0023234

Field, M., Collins, M.W., Lovell, M.R., & Maroon, J. (2003). Does age play a role in recovery from sports-related concussion? A comparison of high school and collegiate athletes. *Journal of Pediatrics*, *142*(5), 546–553. doi: 10.1067/mpd.2003.190

Franke, L.M., Czarnota, J.N., Ketchum, J.M., & Walker, W.C. (2015). Factor analysis of persistent postconcussive symptoms within a military sample with blast exposure. *Journal of Head Trauma Rehabilitation*, *30*(1), E34–E46. doi: 10.1097/HTR.0000000000000042

Galetta, K.M., Brandes, L.E., Maki, K., Dziemianowicz, M.S., Laudano, E., Allen, M., … Balcer, L.J. (2011). The King-Devick test and sports-related concussion: Study of a rapid visual screening tool in a collegiate cohort. *Journal of the Neurological Sciences*, *309*(1-2), 34–39. doi: 10.1016/j.jns.2011.07.039

Gignac, G.E., & Watkins, M.W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research*, *48*(5), 639–662. doi: 10.1080/00273171.2013.804398

Guskiewicz, K.M., Ross, S.E., & Marshall, S.W. (2001). Postural stability and neuropsychological deficits after concussion in collegiate athletes. *Journal of Athletic Training*, *36*(3), 263–273.

Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J.G. (2009). Research electronic data capture (REDCap)–A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. doi: 10.1016/j.jbi.2008.08.010

Helmick, K., Guskiewicz, K., Barth, J., Cantu, R., Kelly, J., McDonald, E., … Warden, D. (2006). *Defense and Veterans Brain Injury Center Working Group on the Acute Management of Mild Traumatic Brain Injury in Military Operational Settings: Clinical practice guideline and recommendations*. Washington, DC: Defense and Veteran Brain Injury Center. Retrieved from http://www.pdhealth.mil/downloads/clinical_practice_guideline_recommendations.pdf.

Joyce, A.S., Labella, C.R., Carl, R.L., Lai, J.S., & Zelko, F.A. (2015). The Postconcussion Symptom Scale: Utility of a three-factor structure. *Medicine and Science in Sports and Exercise*, *47*(6), 1119–1123. doi: 10.1249/MSS.0000000000000534

Kontos, A.P., Elbin, R.J., Schatz, P., Covassin, T., Henry, L., Pardini, J., & Collins, M.W. (2012). A revised factor structure for the post-concussion symptom scale: Baseline and postconcussion

factors. *American Journal of Sports Medicine*, *40*(10), 2375–2384. doi: 10.1177/0363546512455400

Mansolf, M., & Reise, S.P. (2017). When and why the second-order and bifactor models are distinguishable. *Intelligence*, *61*, 120–129.

McCrea, M., Guskiewicz, K.M., Marshall, S.W., Barr, W., Randolph, C., Cantu, R.C., … Kelly, J.P. (2003). Acute effects and recovery time following concussion in collegiate football players: The NCAA Concussion Study. *Journal of the American Medical Association*, *290*(19), 2556–2563. doi: 10.1001/jama.290.19.2556

McCrea, M., Kelly, J.P., Randolph, C., Kluge, J., Bartolic, E., Finn, G., & Baxter, B. (1998). Standardized assessment of concussion (SAC): On-site mental status evaluation of the athlete. *Journal of Head Trauma Rehabilitation*, *13*(2), 27–35.

McCrory, P., Meeuwisse, W., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., … Vos, P.E. (2017). Consensus statement on concussion in sport-the 5th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, *51*, 838–847. doi: 10.1136/bjsports-2017-097699

McCrory, P., Meeuwisse, W.H., Aubry, M., Cantu, B., Dvorak, J., Echemendia, R.J., … Turner, M. (2013). Consensus statement on concussion in sport: The 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine*, *47*(5), 250–258. doi: 10.1136/bjsports-2013-092313

McMahon, P., Hricik, A., Yue, J.K., Puccio, A.M., Inoue, T., & Lingsma, H.F., . . . TRACK-TBI Investigators. (2014). Symptomatology and functional outcome in mild traumatic brain injury: Results from the prospective TRACK-TBI study. *Journal of Neurotrauma*, *31*(1), 26–33. doi: 10.1089/neu.2013.2984

Merritt, V.C., & Arnett, P.A. (2014). Premorbid predictors of postconcussion symptoms in collegiate athletes. *Journal of Clinical and Experimental Neuropsychology*, *36*(10), 1098–1111. doi: 10.1080/13803395.2014.983463

Meterko, M., Baker, E., Stolzmann, K.L., Hendricks, A.M., Cicerone, K.D., & Lew, H.L. (2012). Psychometric assessment of the Neurobehavioral Symptom Inventory-22: The structure of persistent postconcussive symptoms following deployment-related mild traumatic brain injury among veterans. *Journal of Head Trauma Rehabilitation*, *27*(1), 55–62. doi: 10.1097/HTR.0b013e318230fb17

Muthen, B., & Asparouhov, T. (2006). Item response mixture modeling: Application to tobacco dependence criteria. *Addictive Behaviors*, *31*(6), 1050–1066. doi: 10.1016/j.addbeh.2006.03.026

Muthén, L.K., & Muthén, B.O. (1998-2015). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nelson, L.D., & Foell, J. (2018). Externalizing proneness and psychopathy. In C.J. Patrick (Ed.), *Handbook of psychopathy* (2nd ed.). New York, NY: Guilford Press.

Nelson, L.D., LaRoche, A.A., Pfaller, A.Y., Lerner, E.B., Hammeke, T.A., Randolph, C., … McCrea, M.A. (2016). Prospective, head-to-head study of three computerized neurocognitive assessment tools (CNTs): Reliability and validity for the assessment of sport-related concussion. *Journal of the International Neuropsychological Society*, *22*(1), 24–37. doi: 10.1017/S1355617715001101

Nelson, L.D., Tarima, S., LaRoche, A.A., Hammeke, T.A., Barr, W.B., Guskiewicz, K., … McCrea, M.A. (2016). Preinjury

somatization symptoms contribute to clinical recovery after sport-related concussion. *Neurology*, *86*, 1856–1863.

Patrick, C.J., Curtin, J.J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment*, *14*(2), 150–163.

Patrick, C.J., Fowles, D.C., & Krueger, R.F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, *21*(3), 913–938.

Patrick, C.J., Hicks, B.M., Nichol, P.E., & Krueger, R.F. (2007). A bifactor approach to modeling the structure of the psychopathy checklist-revised. *Journal of Personality Disorders*, *21*(2), 118–141. doi: 10.1521/pedi.2007.21.2.118

Piland, S.G., Motl, R.W., Ferrara, M.S., & Peterson, C.L. (2003). Evidence for the Factorial and Construct Validity of a Self-Report Concussion Symptoms Scale. *Journal of Athletic Training*, *38*(2), 104–112.

Potter, S., Leigh, E., Wade, D., & Fleminger, S. (2006). The Rivermead Post Concussion Symptoms Questionnaire: A confirmatory factor analysis. *Journal of Neurology*, *253*(12), 1603–1614. doi: 10.1007/s00415-006-0275-z

Reise, S.P. (2012). Invited paper: The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. doi: 10.1080/00273171.2012.715555

Reise, S.P., Morizot, J., & Hays, R.D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, *16*(Suppl. 1), 19–31. doi: 10.1007/s11136-007-9183-7

Rodriguez, A., Reise, S.P., & Haviland, M.G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. doi: 10.1037/met0000045

Root, J.M., Zuckerbraun, N.S., Wang, L., Winger, D.G., Brent, D., Kontos, A., & Hickey, R.W. (2016). History of somatization Is associated with prolonged recovery from concussion. *Journal of Pediatrics*, *174*, 39–44 e31. doi: 10.1016/j.jpeds.2016.03.020

Sady, M.D., Vaughan, C.G., & Gioia, G.A. (2014). Psychometric characteristics of the postconcussion symptom inventory in children and adolescents. *Archives of Clinical Neuropsychology*, *29*(4), 348–363. doi: 10.1093/arclin/acu014

Thomas, M.L. (2012). Rewards of bridging the divide between measurement and clinical theory: Demonstration of a bifactor model for the Brief Symptom Inventory. *Psychological Assessment*, *24*(1), 101–113. doi: 10.1037/a0024712

Vanderploeg, R.D., Silva, M.A., Soble, J.R., Curtiss, G., Belanger, H.G., Donnell, A.J., & Scott, S.G. (2015). The structure of postconcussion symptoms on the Neurobehavioral Symptom Inventory: A comparison of alternative models. *Journal of Head Trauma Rehabilitation*, *30*(1), 1–11. doi: 10.1097/HTR.0000000000000009

Waljas, M., Iverson, G.L., Hartikainen, K.M., Liimatainen, S., Dastidar, P., Soimakallio, S., … Ohman, J. (2012). Reliability, validity and clinical usefulness of the BNI fatigue scale in mild traumatic brain injury. *Brain Injury*, *26*(7-8), 972–978. doi: 10.3109/02699052.2012.660511

Wechsler, D. (2001). *Wechsler test of adult reading: WTAR*. San Antonio, TX: The Psychological Corporation.