



Published in final edited form as:

*Clin Neuropsychol.* 2016 February ; 30(2): 338–350. doi:10.1080/13854046.2016.1138504.

## Psychometric properties and normative data for the Brief Symptom Inventory-18 (BSI-18) in high school and collegiate athletes

Melissa A. Lancaster, Michael A. McCrea, and Lindsay D. Nelson

Departments of Neurosurgery and Neurology, Medical College of Wisconsin, Milwaukee, WI, USA

Melissa A. Lancaster: mlancaster@mcw.edu

### Abstract

**Objective**—Assessment of emotional functioning is important in sport-related concussion (SRC) management, although few standardized measures have been validated in this population, and appropriate normative data are lacking. We investigated the psychometric properties of the Brief Symptom Inventory-18 (BSI-18) in high school and collegiate athletes at risk of SRC and compiled normative data.

**Method**—Athletes ( $n = 2,031$ ) completed the BSI-18 and other measures of concussion symptoms, cognition, and psychological functioning. A subset of healthy individuals was re-evaluated at approximately 7, 30, 45, and 165 days. Psychometric analyses of test–retest reliability, internal consistency reliability, and concurrent validity were performed. Given significant differences between sexes and education levels (high school or college student) on the BSI-18 Global Severity Index and all subscales, normative conversion tables were produced after stratifying by these variables.

**Results**—The BSI-18 showed good internal consistency, fair to poor test–retest reliability, and good convergent validity with other measures of emotional functioning.

**Conclusions**—These data indicate that the BSI-18 may be a valuable measure of emotional state in concussed athletes and may provide unique information beyond post-concussive symptoms for research on the role of psychological factors in SRC recovery. The limited divergent validity of the BSI-18 depression and anxiety scales implies that they tap into general distress more so than specific mood or anxiety symptoms; therefore, BSI-18 scores should be not relied upon for differential diagnosis of mood and anxiety disorders. Normative data provided can be readily applied to clinical cases with high school and collegiate athletes.

### Keywords

Brief Symptom Inventory-18; sport-related concussion; concussion; emotional functioning; psychometric properties

---

Correspondence to: Melissa A. Lancaster, mlancaster@mcw.edu.

**Disclosure statement:** No potential conflict of interest was reported by the authors.

As the effects of sport-related concussion (SRC) become an increasing health concern, there is an even greater need for reliable and valid measures to assess recovery following injury. It is well documented that the common short-term sequelae of SRC include not only somatic (e.g., headache, dizziness, nausea) and cognitive (e.g., difficulty with memory, concentration, or information processing) symptoms, but also changes in psychological functioning such as anxiety, depression, and/or emotional lability. In addition, a positive relationship has been established between the presence of mood symptoms and length of recovery (Mooney, Speed, & Sheppard, 2005; Satz et al., 1998). It is therefore important that, in addition to cognitive and physical symptoms, emotional symptoms be assessed following SRC in order to guide intervention and improve outcome.

It has also been suggested that preexisting psychological symptoms such as anxiety and depression may increase the risk of sport-related injury, including SRC (Kelley, 1990; Yang et al., 2014). It is therefore important to identify these symptoms prior to the onset of participation in contact sports so that treatment options can be entertained if necessary. Although significant pre-season emotional distress should be treated, pre-injury psychological status is also important to measure in order to interpret post-injury concussion symptoms. In other words, pre-season measurement of emotional symptoms may help clarify the degree to which symptoms (e.g., irritability and sadness) are related to the recent injury or are more longstanding.

Despite the importance of measuring psychological factors in athletes at baseline and after SRC, few measures of emotional functioning have been validated for use in this population. In 2010, the Traumatic Brain Injury (TBI) Outcomes Workgroup, in an attempt to establish common outcome measures in TBI research, recognized emotional functioning as an important domain to be assessed following TBI (Wilde et al., 2010). Within each domain, recommended outcome measures were identified that were widely used in the TBI clinical and research communities and well documented in the scientific literature. Importantly, potential outcome measures needed documentation of sound psychometric properties and normative data. In addition, recommended measures needed to be applicable across a range of severity and individual functioning, publicly available, brief, and easy to administer. Emphasis was placed on designating a single measure that most thoroughly assessed each domain.

For measurement of emotional symptoms, the TBI Outcomes Workgroup recommended the Brief Symptom Inventory-18 (BSI-18; Derogatis, 2001) as the primary outcome measure. In addition, the BSI-18 is recommended by the National Institute of Neurological Disorders and Stroke (NINDS) common data elements (CDE) for assessment of psychological status in adults with concussion/mild TBI (Hicks et al., 2013). An 18-item abbreviated form of the original 90-item Symptom Checklist (SCL-90; Derogatis, 1977), the BSI-18 has been shown to demonstrate high internal consistency reliability, moderate test-retest reliability, strong convergent validity with measures of emotional functioning, and modest incremental validity in adults with moderate to severe traumatic brain injury (TBI; Meachen, Hanks, Millis, & Rapport, 2008). The BSI-18 subscales have also demonstrated moderate convergent validity in a sample of adults with mild to severe TBI (Williams, Rapport, Millis, & Hanks, 2014). Of note, an earlier abbreviation of the SCL-90, the Brief Symptom Inventory, a 53-item

measure (Derogatis, 1982), has been more widely used to measure emotional distress in mild TBI populations (Buchanan & Elias, 2001; Corrigan, Smith-Knapp, & Granger, 1998; Hart et al., 2014; Jellinek, Torkelson, & Harvey, 1982; Slaughter, Johnstone, Petroski, & Flax, 1999).

Although prior work supports the BSI-18's potential utility in the assessment of TBI broadly, there is little or no empirical data on its value for assessment and treatment of athletes with SRC. Limitations in the measure's standardization sample (all participants aged 20) make interpretation particularly difficult in adolescent athletes. Here, we investigated the psychometric properties of the BSI-18 in a sample of young, psychologically healthy high school and collegiate athletes engaged in contact sports (i.e., with higher incidence of SRC) and present normative data for clinicians and SRC researchers using this measure in adolescent and young adult athletes. Normative data in healthy young athletes will help determine if concussed athletes show increased emotional symptoms above and beyond what is expected based on a similar demographic group.

## Method

### Participants

Contact and collision sport student athletes from nine high schools and three colleges in southeastern Wisconsin enrolled in a larger study on the assessment of SRC (Project Head to Head) between August 2012 and October 2014. Informed consent was obtained for 2,154 participants, and 2,049 participated in baseline testing. Given that the aim of this study was to develop normative data, 12 participants were excluded due to a diagnosed psychiatric disorder and 6 due to missing demographic data (i.e., age), resulting in 2,031 total participants included in the current analyses (Table 1). Of the 12 students excluded for psychiatric reasons, the reported diagnoses were as follows: 7 with anxiety disorder, 3 with mood disorder, 1 with somatoform disorder, and 1 with post-traumatic stress disorder. As part of the parent study, a subgroup of 131 healthy athletes was selected to match concussed athletes on demographic characteristics. These healthy control athletes repeated the test procedures on several occasions in accordance with the parent study's protocol. These healthy, non-injured athletes were used in analyses of test-retest reliability.

Adult athletes and parents of minor athletes completed informed consent, and minor participants completed assent prior to baseline testing. Participants were compensated \$30 for their time and effort in completing the 90-min baseline assessments and \$50 for follow-up assessments. All testing procedures were approved by the Institutional Review Board at the Medical College of Wisconsin.

### Testing procedures

All participants underwent a one-on-one standardized health history interview by a research assistant. In order to screen for history of previous psychiatric diagnoses, participants were asked "Have you ever been told by a doctor or health professional that you had a psychiatric disorder?" If applicable, more information was obtained about specific diagnoses. Participants then underwent testing that included the Wechsler Test of Adult Reading

(WTAR; Wechsler, 2001), the Standardized Assessment of Concussion (SAC; McCrea, Kelly, Kluge, Ackley, & Randolph, 1997), the Sport Concussion Assessment Tool—3rd edition (SCAT3) Symptom Checklist (McCrorey et al., 2013), the Satisfaction With Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin 1985), the Balance Error Scoring System (BESS; Guskiewicz, 2001), and the BSI-18. As part of a separate study, participants also completed two computerized neuro-cognitive tests and Green's Medical Symptom Validity Test (MSVT; Green, 2003). Participants tested in August 2014 also completed a brief form of the Multidimensional Personality Questionnaire (MPQ; Patrick, Curtin, & Tellegen, 2002). Tests were individually proctored by a research assistant in quiet settings in groups ranging in size from 1 to 20 athletes. Testing sessions lasted approximately 90 min.

## Measures

**BSI-18**—The BSI-18 consists of 18 items on a 5-point (0–4) Likert scale and is designed to assess current psychological distress (over the past 7 days) in adult patients over age 20. In addition to a total score, referred to as the Global Severity Index (GSI; max = 72), separate scores can be calculated on three subscales: somatization (SOM), depression (DEP), and anxiety (ANX), with six questions contributing to each subscale (max = 24). Higher scores reflect greater distress.

**Satisfaction With Life Scale**—The SWLS (Diener et al., 1985) is a brief five-item measure of overall life satisfaction. Items address global assessment of satisfaction instead of referring to specific topics, which allow respondents to answer according to their own values (Pavot & Diener, 1993). Answers are recorded on a seven-point (1–7) Likert scale with higher scores (max = 35) reflecting greater life satisfaction. Good internal consistency was found in our sample ( $\alpha = .82$ ). This measure was included in the analysis of the BSI-18's concurrent validity due to previous studies demonstrating that the SWLS correlates negatively with well-established measures of negative affect, including all eight symptom dimensions assessed with the revised form of the SCL-90, from which the BSI-18 was created (Arrindell & Ettema, 1986; Blais, Vallerand, Pelletier, & Brière, 1989).

**Multidimensional Personality Questionnaire**—We assessed two higher order facets of personality, positive emotionality (PEM) and negative emotionality (NEM), using an abbreviated version of the MPQ brief form (Patrick et al., 2002). PEM and NEM are broadly conceptualized as temperamental dispositions toward positive and negative emotions, respectively (Patrick et al., 2002). The MPQ asks participants to rate the degree to which statements describe themselves (answers are mostly True/False, although some items require participants to select which of two statements best describes themselves). The version of the MPQ used in this study is comprised of 10 items that make up the NEM scale and 15 items assessing PEM. Each of these higher order scales is made up of 3–4-item subscales, with NEM comprised of the lower order scales of stress reaction, alienation, and aggression, and PEM made up of the subscales measuring well-being, social potency, achievement, and social closeness. Item scores (coded 1 if answered in the keyed direction, 0 if answered in the unkeyed direction) were averaged within each lower order facet and then averaged across the lower order facets for NEM and PEM separately. Cronbach's alpha in our sample was .70 and .49 for PEM and NEM, respectively. These measures were included for analysis of

concurrent validity of the BSI-18 subscales due to prior work demonstrating that anxiety and depression can be differentiated by their relations to NEM and (low) PEM, with low PEM a specific feature of depression and high NEM seen in both conditions (Clark & Watson, 1991; Tellegen, 1985).

### **Sport Concussion Assessment Tool—3rd Edition (SCAT3) Symptom Checklist**

—The SCAT 3 Symptom Checklist (McCrory et al., 2013) is a 22-item post-concussion symptom scale that assesses somatic, cognitive, and emotional functioning. The total score reflects the number of symptoms endorsed (max = 22), while the symptom severity score is obtained by summing the symptom score rated on a seven-point (0–6) Likert scale. Higher scores reflect greater symptom severity (max = 132). Good internal consistency was found in our sample ( $\alpha = .86$ ). The symptom severity score was used in the current analyses due to its greater range and variability.

## **Data analyses**

We investigated psychometric properties of the BSI-18 utilizing all 2,031 eligible participants from the baseline (Time 1) sample and 131 participants who completed repeat testing. Repeat testing was performed an average of 165 days (Time 2) after the first pre-season baseline (Time 1) and an average of 7 (Time 3), 14 (Time 4), 30 (Time 5), and 44 (Time 6) days after the first follow-up examination. These follow-up examinations were selected to match the repeat testing points of matched concussed athletes not included in these analyses. GSI and subscale scores showed significant skewness and kurtosis. As a result, Kruskal–Wallis tests were used to examine omnibus group differences, Mann–Whitney *U* tests were used for *post hoc* comparisons with Pearson's *r* values calculated to measure effect size, and Bonferroni correction was used to control for multiple comparisons. Data were log-transformed for reliability and concurrent validity analyses. Cronbach's  $\alpha$  was used to estimate internal consistency of each BSI-18 subscale and the GSI. Test–retest reliability and concurrent validity were estimated using Pearson product moment correlations. Test–retest reliability was computed for control subjects for 7- (Time 3 vs. Time 4), 30- (Time 4 vs. Time 6), 45- (Time 2 and Time 6), and 165 (Time 1 vs. Time 2)-day test–retest intervals. Concurrent validity analyses were conducted at baseline (Time 1) to capture the largest number of participants. Normative information was classified into percentile bands, given the significant skew in the data. Lower percentiles reflect greater emotional distress.

## **Results**

### **Sample characteristics and normative data**

Demographic characteristics of the total sample are provided in Table 1. No differences existed on the BSI-18 due to race, although Kruskal–Wallis tests revealed significant differences between gender and education level (high school or college student) on the GSI and all subscales (all  $ps < .001$ ). As a result, the sample was stratified by gender and education level: high school females, high school males, college females, and college males (Table 2; Figure 1). Post hoc comparisons revealed several differences. High school females scored higher on all BSI-18 subscales and the GSI compared to the other groups, with the

exception of SOM, where they only scored higher than college males. High school males scored higher than college males on SOM, college males and college females on DEP, college males on ANX, and college males on the GSI. College females scored higher than college males on SOM, ANX, and the GSI. For *post hoc* comparisons, effect sizes were small in all cases. Tables 3–6 contain normative tables for each individual group with raw score to percentile band conversions.

As a supplementary analysis, the effect of previous concussions on BSI-18 ratings was investigated in our sample. While no participants had suffered a recent concussion, some participants had a history of 1, 2, or 3 or more (3+) previous concussions (1 previous SRC,  $n = 418$ ; 2 previous SRCs,  $n = 123$ , 3 + previous SRCs,  $n = 58$ ). Kruskal-Wallis tests showed that significant differences in BSI-18 scores existed based on number of previous concussions on the DEP subscale ( $\chi^2 = 9.9, p = .019$ ) and the GSI ( $\chi^2 = 10.3, p = .016$ ). Post hoc analyses using Mann–Whitney  $u$  tests revealed that DEP scores were significantly higher for those with 2 or 3+ concussions compared to athletes with no concussion history, although these comparisons did not survive Bonferroni correction. GSI scores were significantly higher for those with 1 or 3+ concussions compared to athletes with no previous concussions, with the comparison between 0 and 3+ concussions surviving Bonferroni correction.

## Reliability

**Internal consistency**—The GSI demonstrated good internal consistency ( $\alpha = .83$ ). Internal consistency estimates for each subscale were somewhat lower (DEP:  $\alpha = .76$ ; ANX:  $\alpha = .66$ ; SOM:  $\alpha = .66$ ).

As a supplementary analysis, internal consistency was examined separately based on gender (males:  $N = 1562$ ; females:  $N = 469$ ). The median difference between coefficient alphas across the GSI and each subscale for females vs. males was .04, indicating that estimates were very similar.

**Test–retest**—Over seven days, reliability estimates were fair, ranging from .56 to .70. Reliability estimates over 30 days were poor to fair, ranging from .41 to .65. Estimates over 45 days were poor, ranging from .31 to .40. Estimates over 165 days (approximately 5.5 months) were also poor, ranging from .28 to .52. See Table 7 for complete information. Means and standard deviations for each BSI-18 scale at each time point are presented in Table 8.

A supplementary analysis also investigated test–retest reliability based on gender. Again, the median difference between stability coefficients across time points for females vs. males was .04, indicating that correlations were very similar; however, it should be noted that in these analyses, the number of females at each time point was very small (range of  $ns = 21$ – $25$ ).

## Validity

Table 9 contains information on criterion validity. Large correlations were seen between the BSI-18 GSI and subscale scores ( $r$ 's ranging from .74 to .81). All BSI-18 subscales were also

moderately correlated ( $r$ 's ranging from .38 to .49). The GSI was highly correlated with SCAT 3 Symptom Checklist symptom severity score ( $r = .58, p < .01$ ), and MPQ NEM ( $r = .43, p < .01$ ). All BSI-18 subtest scores also correlated moderately with symptom severity score ( $r$ 's ranging from .42 to .47). BSI-18 DEP ( $r = .36, p < .01$ ) and ANX ( $r = .39, p < .01$ ) also correlated with MPQ NEM. MPQ PEM was unrelated to any BSI-18 scale ( $r$ 's =  $-.06$  to  $.01$ ). There was a moderate negative association between BSI-18 GSI and SWLS total score ( $r = -.32, p < .01$ ). In addition, BSI-18 subscales correlated negatively with SWLS total score, with BSI-18 DEP showing the strongest (moderate) correlation ( $r$ 's ranging from  $-.20$  to  $-.37$ ).

Supplementary analyses investigated validity based on gender. Criterion validity ratings were very similar for males and females (the median difference between Pearson's  $r$  values was .02), although sample size was small for females when comparing the BSI-18 to the MPQ ( $n = 22$ ).

## Discussion

Despite recommendations that the BSI-18 be used as a CDE for measurement of emotional functioning in TBI research, the literature justifying this recommendation is limited mostly to adult patient samples with moderate and severe TBI. Here, we provided initial evidence that the BSI-18 may be useful in both the assessment and study of adolescents and young adults with sport-related concussion. In particular, the BSI-18 showed good internal consistency and demonstrated good criterion validity with other relevant personality and psychological functioning measures, supporting its use in measuring emotional functioning in high school and collegiate athletes.

The BSI-18 showed good internal consistency in this sample of high school and collegiate athletes, with the GSI demonstrating higher internal consistency ( $\alpha = .83$ ) than any of the BSI-18 subscales. The coefficient alpha values observed in our sample were somewhat lower than those published in the original standardization sample of 1,134 adults employed at a national US corporation (Derogatis, 2001) and a sample of adults tested 6 months to 15 years after traumatic brain injury (Meachen et al., 2008). However, our internal consistency estimates were highly similar to those found in a sample of adult inpatient TBI patients (Meachen et al., 2008), indicating that the internal consistency of the BSI-18 varies somewhat with the nature of the population assessed. Our finding that the GSI was the most reliable indicator of mood symptoms in high school and college athletes supports prior assertions that this higher order scale is the measure's best index of overall "emotional adjustment" (Derogatis, 2001); BSI-18 subscales are less reliable, indicating that within each subscale, items may not be measuring the same construct (e.g. "anxiety"); as such, each subscale should be interpreted with more caution (Burlingame, Lambert, Reisinger, Neff, & Mosier, 1995).

Test-retest coefficients ranged from fair after 7 days to poor after 165 days. To some degree, low stability across time is not unexpected for a measure of affective state, especially in younger adults, who have been shown to exhibit more fluctuation in affective states over short time intervals (Rocke, Li, & Smith, 2009). Yet, given the homogeneity of our sample

(comprised of young, psychologically healthy athletes), it is also possible that restriction of range may have impacted test–retest reliability. Consistent with these hypotheses, it has been demonstrated that subject characteristics may impact test–retest reliability more than the length of time between administrations (Dikmen, Heaton, Grant, & Temkin, 1999). For example, Meachen et al. (2008) found that in a sample of inpatients treated for TBI, test–retest coefficients ranging over 6 months to 2 years (ranging from .57 to .67) were higher for the BSI-18 than in the current sample. Despite explanations for this finding, however, poor test–retest reliability on the BSI-18 in a healthy athlete sample over one week, when most injured athletes show the most symptoms following SRC (McCrea et al., 2003), has implications for the BSI-18's value in tracking recovery longitudinally. In other words, if stability cannot be confirmed in healthy young athletes, it will be difficult to determine if change in emotional functioning over time following SRC is related to the injury or instead reflects normal variation or error.

The BSI-18 correlated with other validated measures used to assess the psychological effects of concussion. For example, BSI-18 GSI and subscale scores correlated moderately with scores on the SCAT 3 Symptom Checklist, implying some overlap in the content assessed but also the potential for the BSI-18 to offer unique information beyond the post-concussive symptoms that are routinely assessed in the context of SRC management. Similarly, the BSI-18 GSI and subscales correlated negatively with the SWLS such that greater emotional distress was related to lower overall life satisfaction. The BSI-18 correlated strongly with MPQ negative emotionality (NEM), a personality characteristic reflecting the general tendency to experience negative emotions such as anxiety and anger. Consistent with prior research of this construct (Clark & Watson, 1991), both BSI-18 ANX and DEP subscales showed moderate positive correlations with NEM. But counter to expectation, low positive emotionality (PEM) did not predict higher ratings on DEP. This implies that the DEP scale does not tap the anhedonia unique to major depression. Rather, elevations on the DEP and ANX subscales both appear to reflect general distress. This is consistent with prior assertions about other psychological symptom questionnaires of anxiety and depression, which tend to sample limited content and therefore assess general distress more so than specific anxiety and depression symptoms per se (Clark & Watson, 1991).

With respect to the general mental health of young athletes, our sample reported less psychological distress than the general population, as evident in normative differences on the BSI-18 between college athletes, who overlap with the original BSI-18 standardization sample, and published community norms (Derogatis, 2001). For example, in the community male sample, a GSI score of 5 corresponds to the 50th percentile, while in our collegiate male athlete sample, a GSI score of 5 would fall in the 10–25th percentile. While this is in part due to careful screening for previous psychiatric history, it is also possible that these differences reflect other cohort differences, as the original BSI-18 normative sample included no males and only 1 female under age 20. It has also been documented that young athletes show less mood disturbance than age-matched peers (Oler et al., 1994). Normative and psychometric differences between the current sample and other adult samples with and without a history of TBI (Derogatis, 2001; Meachen et al., 2008; Williams et al., 2014) indicate that healthy young athletes differ from these existing samples, underscoring the



importance of more appropriate norms and a separate investigation of BSI-18 reliability and validity for these individuals.

The normative data provided here for the BSI-18 can be readily applied to clinical cases with young athletes. Given that the current sample is comprised of non-injured athletes, clinicians can easily determine if an athlete, whether healthy or injured, falls in the expected or abnormal range. However, there are some limitations to these data that must be considered. First, the current sample was mostly Caucasian (12% African-American and 5% other) and from a limited Midwestern geographic area; as such, norms may not be as appropriate for minorities or individuals from other regions of the country. In addition, it should be noted that history of concussion was related to increased reporting of mood symptoms in our sample, especially in individuals who had experienced three or greater previous concussions. Clinicians should therefore assess for previous concussions and take this information into account when interpreting test data. Finally, limited divergent validity found between subscales implies that the BSI-18 is largely a measure of general negative affectivity and that clinicians interested in assessing for specific psychiatric conditions (e.g., major depression) in young adults should not rely solely on the BSI-18 to do so.

The psychometric properties of the BSI-18 and its utility in measuring outcome in individuals with traumatic brain injury have been demonstrated (Backhaus, Ibarra, Klyce, Trexler, & Malec, 2010; Dams-O'Connor et al., 2013; Meachen et al., 2008; Vangel, Rapport, & Hanks, 2011). The current study indicates that BSI-18 was moderately reliable and showed reasonable convergent validity in a sample of high school and collegiate athletes, suggesting that it may be valuable in the clinical assessment of athletes with SRC. Additional studies should explore the ability of other self-report measures to better tap specific psychiatric constructs of potential relevance to recovery from SRC (e.g., clinical depression and somatization) to determine if another outcome measure is more appropriate. In addition, more research will be needed to fully understand the interplay between pre-morbid and post-injury psychological factors and recovery from SRC; the current findings indicate that the BSI-18 may be one useful tool in these efforts.

## Acknowledgments

Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the US Army. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

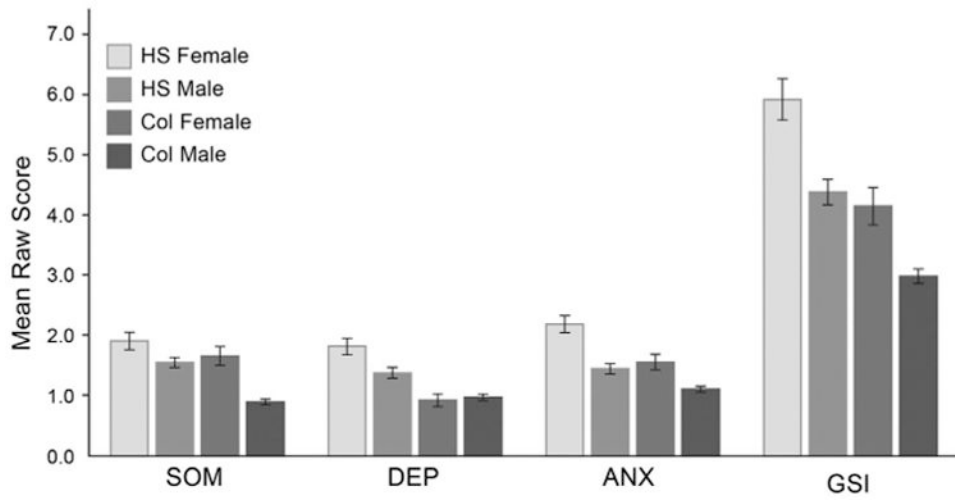
**Funding:** This work was supported by the US Army Medical Research and Materiel Command (US Department of Defense) [award number W81XWH-12-1-0004]; Clinical and Translational Science Institute at the Medical College of Wisconsin [grant number 1UL1-RR031973 (-01)]; National Center for Advancing Translational Sciences, National Institutes of Health [grant number 8UL1TR000055].

## References

- Arrindell, WA.; Ettema, JHM. Handleiding bijeen multidimensionele psychopathologie-indicator [SCL-90: manual for a multidimensional measure of psychopathology]. Lisse: Swets Test Services; 1986.
- Backhaus SL, Ibarra SL, Klyce D, Trexler LE, Malec JF. Brain injury coping skills group: A preventative intervention for patients with brain injury and their caregivers. *Archives of Physical*

- Medicine and Rehabilitation. 2010; 91:840–848. DOI: 10.1016/j.apmr.2010.03.015 [PubMed: 20510972]
- Blais MR, Vallerand RJ, Pelletier LG, Brière NM. L'échelle de satisfaction de vie: Validation canadienne-française du "Satisfaction with Life Scale" [The satisfaction scale: Canadian-French validation of the Satisfaction with Life Scale]. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*. 1989; 21:210–223. DOI: 10.1037/h0079854
- Buchanan KM, Elias LJ. Psychological distress and family burden following spinal cord injury: Concurrent traumatic brain injury cannot be overlooked. *Axone*. 2001; 22:16–17. [PubMed: 14625968]
- Burlingame GM, Lambert MJ, Reisinger CW, Neff WM, Mosier J. Pragmatics of tracking mental health outcomes in a managed care setting. *The Journal of Mental Health Administration*. 1995; 22:226–236. [PubMed: 10144458]
- Clark LA, Watson D. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. [Review]. *Journal of Abnormal Psychology*. 1991; 100:316–336. [PubMed: 1918611]
- Corrigan JD, Smith-Knapp K, Granger CV. Outcomes in the first 5 years after traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*. 1998; 79:298–305. [PubMed: 9523782]
- Dams-O'Connor K, Spielman L, Singh A, Gordon WA, Lingsma HF, Maas AI, et al. Investigators T.-T. The impact of previous traumatic brain injury on health and functioning: A TRACK-TBI study. *Journal of Neurotrauma*. 2013; 30:2014–2020. DOI: 10.1089/neu.2013.3049 [PubMed: 23924069]
- Derogatis, LR. *The SCL-90 Manual I: Scoring, administration and procedures for the SCL-90*. Baltimore, MD: Clinical Psychometric Research; 1977.
- Derogatis, LR. *The Brief Symptom Inventory (BSI): Administration, scoring, and procedures manual*. Baltimore, MD: Johns Hopkins University School of Medicine; 1982.
- Derogatis, LR. *Brief Symptom Inventory 18 (BSI-18): Administration, scoring, and procedures manual*. Bloomington, MN: Pearson; 2001.
- Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *Journal of Personality Assessment*. 1985; 49:71–75. DOI: 10.1207/s15327752jpa4901\_13 [PubMed: 16367493]
- Dikmen SS, Heaton RK, Grant I, Temkin NR. Test-retest reliability and practice effects of expanded Halstead-Reitan neuropsychological test battery. *Journal of the International Neuropsychological Society*. 1999; 5:346–356. [PubMed: 10349297]
- Green, P. *Green's medical symptom validity test for windows*. Edmonton: Green's Publishing; 2003.
- Guskiewicz KM. Postural stability assessment following concussion: One piece of the puzzle. *Clinical Journal of Sport Medicine*. 2001; 11:182–189. [PubMed: 11495323]
- Hart T, Benn EK, Bagiella E, Arenth P, Dikmen S, Hesdorffer DC, et al. Zafonte R. Early trajectory of psychiatric symptoms after traumatic brain injury: Relationship to patient and injury characteristics. *Journal of Neurotrauma*. 2014; 31:610–617. DOI: 10.1089/neu.2013.3041 [PubMed: 24237113]
- Hicks R, Giacino J, Harrison-Felix C, Manley G, Valadka A, Wilde EA. Progress in developing common data elements for traumatic brain injury research: Version two – the end of the beginning. *Journal of Neurotrauma*. 2013; 30:1852–1861. DOI: 10.1089/neu.2013.2938 [PubMed: 23725058]
- Jellinek HM, Torkelson RM, Harvey RF. Functional abilities and distress levels in brain injured patients at long-term follow-up. *Archives of Physical Medicine and Rehabilitation*. 1982; 63:160–162. [PubMed: 7082138]
- Kelley MJ Jr. Psychological risk factors and sports injuries. *Journal of Sports Medicine and Physical Fitness*. 1990; 30:202–221. [PubMed: 2205755]
- McCrea M, Kelly JP, Kluge J, Ackley B, Randolph C. Standardized assessment of concussion in football players. *Neurology*. 1997; 48:586–588. [PubMed: 9065531]
- McCrea M, Guskiewicz KM, Marshall SW, Barr W, Randolph C, Cantu RC, et al. Kelly JP. Acute effects and recovery time following concussion in collegiate football players: The NCAA Concussion Study. *JAMA*. 2003; 290:2556–2563. DOI: 10.1001/jama.290.19.2556 [PubMed: 14625332]
- McCrory P, Meeuwisse WH, Aubry M, Cantu B, Dvorak J, Echemendia RJ, et al. Turner M. Consensus statement on concussion in sport: The 4th International Conference on Concussion in

- Sport held in Zurich, November 2012. *British Journal of Sports Medicine*. 2013; 47:250–258. DOI: 10.1136/bjsports-2013-092313 [PubMed: 23479479]
- Meachen SJ, Hanks RA, Millis SR, Rapport LJ. The reliability and validity of the brief symptom inventory-18 in persons with traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*. 2008; 89:958–965. DOI: 10.1016/j.apmr.2007.12.028 [PubMed: 18452746]
- Mooney G, Speed J, Sheppard S. Factors related to recovery after mild traumatic brain injury. *Brain Injury*. 2005; 19:975–987. DOI: 10.1080/02699050500110264 [PubMed: 16263640]
- Oler MJ, Mainous AG 3rd, Martin CA, Richardson E, Haney A, Wilson D, Adams T. Depression, suicidal ideation, and substance use among adolescents. Are athletes at less risk? *Archives of Family Medicine*. 1994; 3:781–785. [PubMed: 7987512]
- Patrick CJ, Curtin JJ, Tellegen A. Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment*. 2002; 14:150–163. Research Support, Non-US Gov't Research Support, US Gov't, P.H.S. [PubMed: 12056077]
- Pavot W, Diener E. Review of the Satisfaction with life scale. *Psychological Assessment*. 1993; 5:164–172. DOI: 10.1037/1040-3590.5.2.164
- Roche C, Li SC, Smith J. Intraindividual variability in positive and negative affect over 45 days: Do older adults fluctuate less than young adults? *Psychology and Aging*. 2009; 24:863–878. DOI: 10.1037/a0016276 [PubMed: 20025402]
- Satz P, Forney DL, Zaucha K, Asarnow RR, Light R, McCleary C, et al. Becker D. Depression, cognition, and functional correlates of recovery outcome after traumatic brain injury. *Brain Injury*. 1998; 12:537–553. [PubMed: 9653518]
- Slaughter J, Johnstone G, Petroski G, Flax J. The usefulness of the Brief Symptom Inventory in the neuropsychological evaluation of traumatic brain injury. *Brain Injury*. 1999; 13:125–130. [PubMed: 10079957]
- Tellegen, A. Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In: Maser, AHTJD., editor. *Anxiety and the anxiety disorders*. Hillsdale, NJ: Erlbaum; 1985. p. 681-706.
- Vangel SJ Jr, Rapport LJ, Hanks RA. Effects of family and caregiver psychosocial functioning on outcomes in persons with traumatic brain injury. *Journal of Head Trauma Rehabilitation*. 2011; 26:20–29. DOI: 10.1097/HTR.0b013e318204a70d [PubMed: 21209560]
- Wechsler, D. Wechsler test of adult reading: WTAR. San Antonio, TX: The Psychological Corporation; 2001.
- Wilde EA, Whiteneck GG, Bogner J, Bushnik T, Cifu DX, Dikmen S, et al. von Steinbuechel N. Recommendations for the use of common outcome measures in traumatic brain injury research. *Archives of Physical Medicine and Rehabilitation*. 2010; 91:1650–1660 e1617. DOI: 10.1016/j.apmr.2010.06.033 [PubMed: 21044708]
- Williams MW, Rapport LJ, Millis SR, Hanks RA. Psychosocial outcomes after traumatic brain injury: Life satisfaction, community integration, and distress. *Rehabilitation Psychology*. 2014; 59:298–305. DOI: 10.1037/a0037164 [PubMed: 25019312]
- Yang J, Cheng G, Zhang Y, Covassin T, Heiden EO, Peek-Asa C. Influence of symptoms of depression and anxiety on injury hazard among collegiate american football players. *Research in Sports Medicine*. 2014; 22:147–160. DOI: 10.1080/15438627.2014.881818 [PubMed: 24650335]



**Figure 1.** Raw scores on Brief symptom inventory-18 scales by gender and education level. Error bars reflect 1 standard error of measurement. Col = college; HS = high school.

**Table 1**

Sample demographic characteristics by normative group.

	High School Females		High School Males		College Females		College Males	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age	16.2	1.1	15.9	1.0	18.9	1.1	19.1	1.3
WTAR standard score	105.2	10.6	101.5	13.3	103.6	11.1	100.0	12.8
Completed Education	9.7	1.0	9.7	1.0	12.8	1.0	12.6	.9
Race	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
American Indian/Alaska Native	0	.0	4	.7	1	.5	5	.5
Asian	5	2.0	12	2.1	1	.5	9	.9
Black	12	4.7	98	17.3	1	.5	132	13.3
Pacific Islander	2	.8	4	.7	3	1.4	3	.3
White	231	90.6	436	76.9	205	95.8	819	82.3
Unknown	5	2.0	13	2.3	3	1.4	22	2.2
Sport	0	.0	0	.0	0	.0	0	.0
Football	0	.0	319	56.3	0	.0	667	67.0
Soccer	166	65.1	148	26.1	151	70.6	191	19.2
Field hockey	37	14.5	0	.0	0	.0	0	.0
Wrestling	0	.0	76	13.4	0	.0	0	.0
Lacrosse	0	.0	11	1.9	40	18.7	99	9.9
Rugby	51	20.0	0	.0	0	.0	0	.0
Ice Hockey	1	.4	13	2.3	23	10.7	38	3.8

Note: *n* = 2,031. Sample size for each group is as follows: High School Female: *n* = 255; High School Male: *n* = 567; College Female: *n* = 214; and College Male: *n* = 955. WTAR = Wechsler Test of Adult Reading.

Table 2

Gender and education level differences in Brief Symptom Inventory-18 raw scale scores.

BSI-18 Scale	Group	M	SD	$\chi^2$	p	$\eta^2$	Post Hoc Comparisons (r)
SOM	1. High school females	1.9	2.3	107.6	<.001	.053	1>4 (.22); 2>4 (.21); 3>4 (.24)
	2. High school males	1.6	2.0				
	3. College females	1.7	2.3				
	4. College males	.9	1.5				
DEP	1. High school females	1.8	2.2	67.2	<.001	.033	1>2 (.14); 1>3 (.26); 1>4 (.21); 2>3 (.10); 2>4 (.11)
	2. High school males	1.4	2.2				
	3. College females	.9	1.6				
	4. College males	1.0	1.7				
ANX	1. High school females	2.2	2.2	71.9	<.001	.035	1>2(.19); 1>3 (.16); 1>4 (.23); 2>4 (.08); 3>4 (.09)
	2. High school males	1.4	2.1				
	3. College females	1.6	1.9				
	4. College males	1.1	1.6				
GSI	1. High school females	5.9	5.4	108.6	<.001	.053	1>2 (.16); 1>3 (.19); 1>4 (.27); 2>4 (.16); 3>4 (.18)
	2. High school males	4.4	5.1				
	3. College females	4.1	4.6				
	4. College males	3.0	3.8				

Note: Kruskal–Wallis omnibus tests were used to compare groups. Mann–Whitney tests were computed to determine significant *post hoc* comparisons, and multiple comparisons were controlled by Bonferroni correction. Effect sizes for *post hoc* comparisons are presented as Pearson's *r* values. ANX = Anxiety; BSI-18 = Brief Symptom Inventory-18; DEP = Depression; GSI = Global Severity Index; and SOM = Somatization. Sample size for each group is as follows: High School females:  $n = 255$ ; High School males:  $n = 567$ ; College females:  $n = 214$ ; and College males:  $n = 955$ .

**Table 3**

Brief Symptom Inventory-18 percentile bands for high school female athletes.

Brief Symptom Inventory-18 Raw Score					
Percentile	SOM	DEP	ANX	GSI	Percentile
<1	9-24	10-24	10-24	24-72	<1
1-2	8	8-9	8-9	21-23	1-2
3-9	5-7	5-7	6-7	14-20	3-9
10-25	3-4	3-4	3-5	9-13	10-25
26-75	0-2	0-2	0-2	2-8	26-75
>75	-	-	-	0-1	>75

Note:  $n = 255$ . ANX = Anxiety; DEP = Depression; GSI = Global Severity Index; and SOM = Somatization.

**Table 4**

Brief Symptom Inventory-18 percentile bands for high school male athletes.

Brief Symptom Inventory-18 Raw Score					
Percentile	SOM	DEP	ANX	GSI	Percentile
<1	9-24	11-24	10-24	25-72	<1
1-2	7-8	9-10	8-9	22-34	1-2
3-9	4-6	4-8	4-7	11-21	3-9
10-25	2-3	2-3	2-3	6-10	10-25
26-75	0-1	0-1	0-1	1-5	26-75
>75	-	-	-	0	>75

Note:  $n = 567$ . ANX = Anxiety; DEP = Depression; GSI = Global Severity Index; and SOM = Somatization.



**Table 5**

Brief Symptom Inventory-18 percentile bands for collegiate female athletes.

Brief Symptom Inventory-18 Raw Score						
Percentile	SOM	DEP	ANX	GSI	Percentile	Percentile
<1	9-24	8-24	7-24	16-72	<1	<1
1-2	8	6-7	6	15	1-2	1-2
3-9	4-7	3-5	4-5	10-14	3-9	3-9
10-25	2-3	1-2	2-3	6-9	10-25	10-25
26-75	0-1	0	0-1	1-5	26-75	26-75
>75	-	-	-	0	>75	>75

Note:  $n = 214$ . ANX = Anxiety; DEP = Depression; GSI = Global Severity Index; and SOM = Somatization.

**Table 6**

Brief Symptom Inventory-18 percentile bands for collegiate male athletes.

Brief Symptom Inventory-18 Raw Score						
Percentile	SOM	DEP	ANX	GSI	Percentile	Percentile
<1	6-24	8-24	7-24	18-72	<1	<1
1-2	5	6-7	6	14-17	1-2	1-2
3-9	3-4	3-5	3-5	8-13	3-9	3-9
10-25	1-2	1-2	2	4-7	10-25	10-25
26-75	0	0	1	0-4	26-75	26-75
>75	-	-	0	-	>75	>75

Note:  $n = 955$ . ANX = Anxiety; DEP = Depression; GSI = Global Severity Index; and SOM = Somatization.

**Table 7**

Internal consistency and test–retest reliability estimates of the Brief Symptom Inventory-18 scale.

	Internal Consistency	Test–Retest Reliability			
		7 days	30 days	45 days	165 days
GSI	.83	.69	.65	.37	.52
SOM	.66	.70	.60	.40	.28
ANX	.66	.62	.41	.31	.46
DEP	.76	.56	.52	.38	.44

Note: internal consistency:  $n = 2,031$ . Test–retest reliability:  $rs$  at 7, 30, 45, and 165 days = 119, 99, 103, and 131, respectively. ANX = anxiety; Dep = Depression; gsi = global severity index; and SOM = somatization.

**Table 8**

Means and standard deviations for Brief Symptom Inventory-18 raw scale scores at each time point.

	Time 1 (Baseline)		Time 2 (165 Days Post Baseline)		Time 3 (7 days Post Time 2)		Time 4 (14 days Post Time 2)		Time 5 (44 days Post Time 2)	
	<i>n</i> = 131		<i>n</i> = 131		<i>n</i> = 125		<i>n</i> = 123		<i>n</i> = 103	
BSI-18 scale	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
GSI	3.79	4.44	2.34	3.38	1.75	2.98	1.24	2.31	1.17	1.82
DEP	1.36	2.21	.80	1.52	.59	1.21	.39	.98	.50	1.31
ANX	1.38	1.83	.79	1.40	.53	1.25	.32	.87	.29	.59
SOM	1.05	1.52	.74	1.35	.63	1.36	.53	1.23	.38	.82

Note: ANX = Anxiety; DEP = Depression; GSI = Global Severity Index; and SOM = Somatization.

Correlation matrix of log-transformed Brief Symptom Inventory- 18 (BSI-18) scales and other measures of emotional functioning, personality, and concussion symptoms.

**Table 9**

	<i>M</i>	<i>SD</i>	BSI-18 GSI	BSI-18 DEP	BSI-18 ANX	BSI-18 SOM
BSI-18 GSI	3.86	4.59				
BSI-18 DEP	1.19	1.94	.74**			
BSI-18 ANX	1.39	1.89	.81**	.49**		
BSI-18 SOM	1.29	1.89	.76**	.38**	.47**	
SCAT 3	5.34	6.79	.58**	.42**	.47**	.46*
SWLS	27.68	4.61	-.32**	-.36**	-.24**	-.20**
MPQ PEM	-.04	.67	.01	-.06	-.03	.04
MPQ NEM	-.01	.62	.43**	.36**	.39**	.17

Note: ANX = Anxiety, DEP = Depression; GSI = Global Severity Index; MPQ = Multidimensional Personality Questionnaire; NEM = Negative Emotionality; PEM = Positive Emotionality; SCAT 3 = Sport Concussion Assessment Tool--3rd Edition; SOM = Somatization; and SWLS = Satisfaction with Life Scale. For BSI and SCAT 3 symptom score correlations,  $n = 2031$ . For correlations with SWLS,  $n = 2027$ . For correlations with MPQ,  $n = 86$ .

\* denotes  $p < .05$ ;

\*\* denotes  $p < .01$ .