Title

Invariance of the Bifactor Structure of Mild Traumatic Brain Injury (mTBI) Symptoms on the Rivermead Postconcussion Symptoms Questionnaire Across Time, Demographic Characteristics, and Clinical Groups: A TRACK-TBI Study

Journal

Assessment, 28(6)

ISSN

1073-1911

Authors

Agtarap, Stephanie
Kramer, Mark D
Campbell-Sills, Laura
et al.

Publication Date

2021-09-01

DOI

10.1177/1073191120913941

Peer reviewed

Published in final edited form as:

Assessment. 2021 September ; 28(6): 1656–1670. doi:10.1177/1073191120913941.

Invariance of the Bifactor Structure of Mild Traumatic Brain Injury (mTBI) Symptoms on the Rivermead Post-Concussion Symptoms Questionnaire across Time, Demographic Characteristics, and Clinical Groups: A TRACK-TBI Study

Stephanie Agtarap[1], **Mark D Kramer**[2], **Laura Campbell-Sills**[1], **Esther Yuh**[3], **Pratik Mukherjee**[3], **Geoffrey T Manley**[3], **Michael A McCrea**[4], **Sureyya Dikmen**[5], **Joseph T Giacino**[6], **Murray B Stein**[1], **Lindsay D Nelson**[4], **TRACK-TBI Investigators**[*]

[1]Defense and Veterans Brain Injury Center, Naval Medical Center, San Diego, CA, USA.

[2]Independent Consultant, USA.

[3]University of California San Francisco, San Francisco, CA, USA.

[4]Medical College of Wisconsin, Milwaukee, WI, USA.

[5]University of Washington, Seattle, WA, USA.

[6]Massachusetts General Hospital, Boston, MA, USA.

Abstract

This study aimed to elucidate the structure of the Rivermead Post Concussion Symptoms Questionnaire (RPQ) and evaluate its longitudinal and group variance. Factor structures were developed and compared in 1,011 patients with mild traumatic brain injury (mTBI; i.e., Glasgow Coma Scale score 13–15) from the Transforming Research and Clinical Knowledge in TBI

Corresponding Author: Lindsay D. Nelson, PhD, ABPP, Department of Neurosurgery, Medical College of Wisconsin, 8701 West Watertown Plank Road, Milwaukee, WI 53226, Phone: (414) 955-7308, Fax: (414) 955-0193, linelson@mcw.edu.

[*]**The TRACK-TBI Investigators:** Opeolu Adeoye, MD, University of Cincinnati; Neeraj Badjatia, MD, University of Maryland; Kim Boase, University of Washington; Yelena Bodien, PhD, Massachusetts General Hospital; M. Ross Bullock, MD PhD, University of Miami; Randall Chesnut, MD, University of Washington; John D. Corrigan, PhD, ABPP, Ohio State University; Karen Crawford, University of Southern California; Ramon Diaz-Arrastia, MD PhD, University of Pennsylvania; Ann-Christine Duhaime, MD, Mass General Hospital; Richard Ellenbogen, MD, University of Washington; V Ramana Feeser, MD, Virginia Commonwealth University; Adam R. Ferguson, PhD, University of California, San Francisco; Brandon Foreman, MD, University of Cincinnati; Raquel Gardner, University of California, San Francisco; Etienne Gaudette, PhD, University of Southern California; Dana Goldman, PhD, University of Southern California; Luis Gonzalez, TIRR Memorial Hermann; Shankar Gopinath, MD, Baylor College of Medicine; Rao Gullapalli, PhD, University of Maryland; J Claude Hemphill, MD, University of California, San Francisco; Gillian Hotz, PhD, University of Miami; Sonia Jain, PhD, University of California, San Diego; Frederick K. Korley, MD, PhD, University of Michigan; Joel Kramer, PsyD, University of California, San Francisco; Natalie Kreitzer, MD, University of Cincinnati; Harvey Levin, MD, Baylor College of Medicine; Chris Lindsell, PhD, Vanderbilt University; Joan Machamer, MA, University of Washington; Christopher Madden, MD, UT Southwestern; Alastair Martin, PhD, University of California, San Francisco; Thomas McAllister, MD, Indiana University; Randall Merchant, PhD, Virginia Commonwealth University; Laura B. Ngwenya, MD, PhD, University of Cincinnati; Florence Noel, PhD, Baylor College of Medicine; David Okonkwo, MD PhD, University of Pittsburgh; Eva Palacios, PhD, University of California, San Francisco; Daniel Perl, MD, Uniformed Services University; Ava Puccio, PhD, University of Pittsburgh; Miri Rabinowitz, PhD, University of Pittsburgh; Claudia Robertson, MD, Baylor College of Medicine; Jonathan Rosand, MD, MSc, Massachusetts General Hospital; Angelle Sander, PhD, Baylor College of Medicine; Gabriella Satris, University of California, San Francisco; David Schnyer, PhD, UT Austin; Seth Seabury, PhD, University of Southern California; Mark Sherer, PhD, TIRR Memorial Hermann; Sabrina Taylor, PhD, University of California, San Francisco; Nancy Temkin, PhD, University of Washington; Arthur Toga, PhD, University of Southern California; Alex Valadka, MD, Virginia Commonwealth University; Mary Vassar, RN MS, University of California, San Francisco; Paul Vespa, MD, University of California, Los Angeles; Kevin Wang, PhD, University of Florida; John K. Yue, MD, University of California, San Francisco; Ross Zafonte, Harvard Medical School

(TRACK-TBI) study, using RPQ data collected at 2 weeks, and 3, 6, and 12 months post-injury. A bifactor model specifying a general factor and emotional, cognitive, and visual symptom factors best represented the latent structure of the RPQ. The model evinced strict measurement invariance over time and across sex, age, race, psychiatric history, and mTBI severity groups, indicating that differences in symptom endorsement were completely accounted for by these latent dimensions. While highly unidimensional, the RPQ has multidimensional features observable through a bifactor model, which may help to differentiate symptom expression patterns in the future.

Keywords

mild TBI; bifactor; invariance; post-concussive symptoms; Rivermead Post-Concussion Symptoms Questionnaire; traumatic brain injury

In the United States, nearly three million individuals are affected by traumatic brain injury (TBI) every year (Taylor et al., 2017), the vast majority of whom are classified as having mild TBI (mTBI), based on a Glasgow Coma Scale (GCS) of 13–15 (Rimel et al., 1981). Patients with mTBI report a wide array of cognitive, emotional, and physical symptoms following injury, including headaches, dizziness, nausea/vomiting, noise sensitivity, sleep disturbance, fatigue, irritability, depression/tearfulness, frustration/impatience, forgetfulness, poor concentration, slowed thinking, blurred vision, light sensitivity, double vision, and restlessness (International Classification of Diseases 10th edition [ICD-10], 1992).

Although the prevalence of persistent mTBI symptoms is debated, it is well-documented that these symptoms are typically most severe acutely, diminish for most patients over time, and can vary in presentation and duration across individuals (Arciniegas et al., 2005; Bazarian et al., 2009; Boake, et al., 2005; Hiploylee et al., 2017; Mitten & Strauman, 2000; Voormolen et al., 2018). Based on the wide variability in recovery course and outcomes after mTBI and the limited treatment options available, there has been increasing interest in identifying distinct patterns of symptom presentation with the goal of designing precision medicine treatments for mTBI. Motivated by this general goal, the present study leveraged latent variable modeling techniques (factor analysis) to evaluate the degree to which distinct dimensions underlying mTBI symptoms (i.e., “clinical phenotypes”) exist that may inform clinical and translational research.

Rivermead Post-Concussion Symptoms Questionnaire

The Rivermead Post-Concussion Symptoms Questionnaire (RPQ), developed by King et al. (1995), is a widely used self-report questionnaire of mTBI symptoms in mTBI research settings. Although conceptually derived from published literature as a measure of universal TBI symptoms (King et al., 1995), where item ratings are typically summed to create a total score, prior psychometric investigations have suggested that the RPQ may have a multidimensional structure (i.e., measure more than one underlying construct). For example, Rasch analysis performed by two independent groups using data from adults with head injury in the chronic (3+ months post-injury) recovery course were interpreted as indicating that the RPQ does not measure a unidimensional construct (Eyres et al., 2005; Lannsjö et

al., 2011). One of these studies led to the practice used by some researchers to divide the instrument into two subscales, one measuring the purportedly acute symptoms of headache, dizziness, and nausea, and the other reflecting the remaining 13 items of the RPQ (Eyres et al., 2005).

Other investigations provided evidence of additional factors that accounted for RPQ item endorsement. Confirmatory factor analyses (CFA) by Potter, Leigh, Wade and Fleminger (2006) yielded evidence of three highly correlated factors (i.e. cognitive, emotional, somatic) that explained variance in symptoms reported 6 months after head injury, although similar fit was obtained when the emotional and somatic factors were collapsed. Another CFA using RPQ data from patients with mild-to-moderate TBI yielded a different three-factor model with cognitive/emotional, somatic, and visual factors; however, the participants were patients who had both sustained TBI within the last year and reported post-TBI depression (Hermann et al., 2009).

A more recent CFA of the RPQ among military service members following blast exposure found the structure from Potter et al. to be ill-fitting, whereas a 4-factor structure comprising emotional, cognitive, visual, and vestibular domains was favored (Franke et al., 2015). However, a substantial portion of the sample endorsed a high level of posttraumatic stress disorder symptoms and depression following injury (Franke et al., 2015). Lastly, because the four factors were highly correlated, correlation patterns with external variables were very similar across factors. Taken together, this prior work is consistent in elucidating some degree of multidimensionality in the latent structure of the RPQ, but the number and nature of its dimensions are less well delineated.

The Bifactor Model as an Alternative Conceptualization of RPQ Symptom Structure

The strong correlations among factors identified in prior structural models of the RPQ suggest the presence of a superordinate or general factor that may reflect overall TBI symptom severity. As such, second order factor models and bifactor models may each account for a largely unidimensional domain - while reconciling elements of multidimensionality - in ways correlated factor models used in prior studies have not. These approaches have distinctive implications for the latent structure of a domain, and account for multidimensionality differently. Second order factor models posit that a superordinate, unidimensional factor influences item endorsement via intermediate first order factors that account for the shared variance among groups of items comprising differentiable symptom dimensions. In a bifactor model, proportion of variance in each item is accounted for by a general, unidimensional factor (e.g., general self-reported mTBI symptom severity); residual item covariances that are unaccounted for by the general factor exhibit additional structure, such that independent subdomains or specific factors account for remaining variance in the domain (Nelson et al., 2018). Figure 1 illustrates the comparison between a second order and bifactor model. In contrast, a correlated factor model posits that separable but related (to varying degrees) factors account for the shared variance between sets of items and that these correlated dimensions account for the latent structure and, therefore, item endorsement

patterns within the domain. In our recent investigation of the structure of another symptom checklist widely used in athletes with sport-related mTBI (the Sport Concussion Assessment Tool, or SCAT), we found that a bifactor structure (rather than correlated or second order factor models) best delineated the structure of mTBI symptoms and manifested better discriminant validity across external clinical variables (Nelson et al., 2018), supporting the consideration of a bifactor structure for other mTBI symptom measures and patient populations.

Evaluating the Comparability of Structural Model Factors Using Measurement Invariance Analyses

In addition to evaluating the structure of the RPQ through representation by a wider array of models, it is important to evaluate the degree to which the model's factors can be comparably interpreted over time and in different groups. Mild TBI recovery is a dynamic process where distinct neurophysiological events occur over time (Giza & Hovda, 2014) and the predominant symptoms can change. For example, it has been suggested that physiological symptoms are often most prominent in the acute post-injury period (King, 1997), whereas cognitive and emotional symptoms may predominate in the chronic recovery phase (Gordon et al., 2000). Similarly, dimensions of mTBI symptoms can differ across distinct groups of individuals. For example, factors like sex (i.e., female; Rabinowitz et al., 2015), greater age (King, 2014), race (Caucasians vs. non-Caucasians), presence of psychiatric history (Hermann et al., 2009; Lishman, 1988; Stein et al., 2015), and TBI severity (Dikmen et al., 2017) have all been found to predict distinct recovery courses and symptom outcomes in mTBI samples.

Thus, it is conceivable that mTBI symptom checklists such as the RPQ measure different constructs over time or between groups - questions addressable through formal tests of measurement invariance. Should the RPQ measure different constructs over time and/or across groups, then biases in the test would yield uninterpretable results with regard to recovery and/or group mean differences. However, should invariance analyses show that the structural model factors of the RPQ are directly comparable over time and/or across groups, then the measure would be much more useful in future work elaborating upon the sources of these differences. To our knowledge, no study has formally investigated the integrity of the RPQ structural model parameters across time or groups of interest via measurement invariance testing, which assesses the degree to which structural model constructs are measured on the same scales (Reise, Widaman, & Pugh, 1993). In addition to informing a richer understanding of the structure of mTBI symptoms, evaluating the longitudinal and group invariance of the model best reflective of the RPQ symptom domain may inform whether or not this instrument will be valuable in future clinical and research efforts to delineate phenotypes differentiated by self-reported mTBI symptoms.

Current Study

The purposes of this study were to elucidate the factor structure of the RPQ and to evaluate the replicability of the model's parameters across time and groups in a large, longitudinal sample of patients with community-acquired mTBI enrolled in the multicenter Translating

Research and Clinical Knowledge in TBI (TRACK-TBI) study. TRACK-TBI participants were recruited from Level I trauma centers within 24 hours of injury and completed the RPQ at 2 weeks and 3, 6, and 12 months post-injury. Leveraging the large TRACK-TBI sample with acute and post-injury follow-up assessments, we investigated the interpretability of the model's factors across sex, age, race, psychiatric history, and mTBI severity as indicated by brain scan, as well as over time (i.e., from 2 weeks through 12 months post-injury). Based on the aforementioned work, we hypothesized that a bifactor (rather than a correlated factor or a second order) model would best represent the structure of the RPQ, and that the model's factors would reflect quantitative (i.e., severity) as opposed to qualitative (e.g., biased item responses) differences over the recovery period and across patient subgroups.

Method

TRACK-TBI Study

Subjects were identified and recruited through the TRACK-TBI study – a prospective, multicenter study of patients across 11 Level I trauma centers in the United States (Yue et al., 2013). Eligibility criteria for TRACK-TBI included presentation within 24 hours of injury with clinical indication for a head computed tomography (CT) scan (e.g., acute intracranial abnormalities, bleeding, or other neuroimaging evidence of brain injury due to recent trauma), and showing or reporting evidence of alterations in consciousness or amnesia. Exclusion criteria included being in custody, being pregnant, having non-survivable physical trauma or a debilitating mental disorder (i.e., schizophrenia, bipolar disorder or other that would interfere with follow-up and validity of outcome assessments), or being non-English or non-Spanish speaking. All participants or legal representatives gave written informed consent to participate in the study. Outcome data were collected from each patient at 2 weeks and 3, 6, and 12 months after injury. Patient interviews were conducted either by phone follow-up (3 months) or in person (2 weeks, 6 months, 12 months), with some assessments performed over the phone at time points other than 3 months to limit missing data. Study protocols were approved by the institutional review boards of each respective site.

Participants

Mild TBI patients enrolled between 2/26/2014 and 5/4/2016 were considered for analyses. A total of 1,352 patients with TBI were enrolled—of these, 1,155 met our inclusion criterion for mTBI (defined as Glasgow Coma Score of 13–15 upon Emergency Department arrival). Of the 1,155 participants with mTBI, the percentage with complete outcome data at each follow-up assessment was 81.1% (2 weeks), 75.6% (3 months), 71.5% (6 months), and 66.0% (12 months). Item-level missingness in the RPQ was very low (<0.5%) within patients who followed-up at each timepoint. Therefore, patients were included in analyses if they provided any RPQ data at any follow-up assessment, yielding 1,011 participants in the analyses reported below.

Table 1 shows descriptive statistics for the sample. Participants with outcome data did not differ in sex, cause of injury, psychiatric history, or prior TBI history at any of the 4 time points compared to those without. At 2 weeks post-injury, participants with outcome data

were an average of 5 years older than participants lost to follow-up ($p < .001$), but no age differences were observed at other time points. At 12 months post-injury, participants with outcome data were more likely to be non-White than those lost to follow-up ($p = .012$), but no race differences were observed at other time points. In terms of ethnicity, participants with outcome data at 3 of 4 time points (3 months through 12 months), were more likely to be non-Hispanic ($p < .001$) than participants without outcome data at those points. Across all time points, participants with outcome data were somewhat more educated than those lost to follow-up ($\Delta M = 1$ year, ps $\leq .009$). Finally, participants with outcome data at 2 weeks were more likely to have a positive CT scan ($p = .028$).

Measures

RPQ.—Developed by King et al. (1995), the RPQ elicits self-report ratings of the severity of symptoms over the last 24 hours compared to pre-injury levels. The response options for each of the 16 items are 0 (*not experienced at all*), 1 (*no more of a problem than pre-injury*), 2 (*mild problem*), 3 (*moderate problem*), or 4 (*severe problem*). Because scores of 0 and 1 both represent the absence of injury-related symptoms, these responses were collapsed into 0 (*not experienced at all/no more of a problem*; King et al., 1995).

Demographic and injury-related variables.—Demographic, history, and injury-related variables submitted to measurement invariance analyses included age, race, sex, self-reported history of psychiatric disorder, and acute neuroimaging findings. Race, sex, and history of psychiatric disorder were self-reported. Because data were received in de-identified fashion with ages > 90 given a nonspecific code, two participants with ages > 90 years old were coded as 90.5 years old for continuous analyses of age. Age was categorized into terciles for measurement invariance analyses. Due to limited representation of other races in the sample, invariance testing with race only included participants identifying as African American or Caucasian and compared as a binary variable. Acute neuroimaging findings comprised the outcome of CT scans as read by a single board-certified neuroradiologist, coded as positive or negative for acute intracranial abnormalities utilizing the recommended TBI Common Data Elements (e.g., subarachnoid hemorrhage, contusion; Duhaime et al., 2010).

Data Analytic Plan

As our first aim was to assess the factor structure of the RPQ, we used a multi-step factor analytic approach to the data. First, due to the ordinal nature of the items, we observed the relationship between items using polychoric correlations (for inter-item correlation matrix for all timepoints, see Model Invariance Supplementary Output). Second, 1/3 of the sample was randomly selected to conduct exploratory factor and bifactor analyses (EFA, bi-EFA, respectively) of the 16 items to identify candidate structures for each respective time point. Third, we applied these candidate structures to the remaining 2/3 of the sample in a series of CFAs, including correlated factor, second order, and bifactor models, to examine the fit of various structural models at each time point. Lastly, we conducted measurement invariance analyses to determine the extent to which structural model parameters were equivalent across clinically relevant groups (e.g., sex) and across different stages (i.e., assessment time

points) of the mTBI recovery period. All factor and measurement invariance analyses were conducted in *Mplus* (7th edition; Muthen & Muthen, 1998–2017).

Factor analyses.—Due to the ordinal nature of the items, item skewness, and the low level of item-level missingness within the data, factor analyses were conducted using mean- and variance-adjusted weighted least squares (WLSMV) estimation (Beauducel & Herzberg, 2006; Li, 2016). Geomin rotation, an oblique rotation method which allows for factors to be correlated was used in EFA, whereas factors were uncorrelated in bi-EFA. Eigenvalues, scree plots, and factor loadings were considered to determine candidate EFA and bi-EFA structures at each follow-up time point. The root mean square error of approximation (RMSEA) was used to assess absolute fit, whereas the comparative fit index (CFI) and Tucker Lewis index (TLI) were used to assess incremental fit. Models with RMSEA < .05 and CFI/TLI > .95 were considered well-fitting, and RMSEA < .08 was considered acceptably fitting (Hu & Bentler, 1999; MacCallum, Browne, & Sugawara, 1996). For correlated and second order CFA models, each item was specified to load on the factor on which it loaded most strongly in the corresponding EFA (except when specifying a relatively consistent model at each time point, see below). For bifactor CFA models, each item was parameterized to load on the general factor, and items were parameterized to additionally load on a specific factor if they loaded at least .30 on that factor in the corresponding bi-EFA.

Measurement invariance analyses.—Formal factorial measurement invariance modeling (Liu et al., 2017; Meredith, 1993; Putnick & Bornstein, 2016) allows for tests of the degree to which facets of the model reflect the same constructs in different groups or across different assessment time points. There are several levels of measurement invariance that impose increasingly stringent constraints on the equivalence of the model across groups/time; when fit statistics suggest that the more stringent model fits no worse than the one with more relaxed assumptions, one cannot reject the model with more constraints placed upon parameters specifying more equivalency in parameters across groups/time. The first two levels, configural and weak invariance, constrain parameters such that their interpretations yield evidence for qualitative differences across groups/time. The latter two levels, strong and strict invariance, yield evidence for quantitative (i.e., severity) differences across groups/time. Configural invariance specifies models such that the same overall structure can be imposed across groups/time. Weak invariance specifies that the same items load on each respective factor (i.e., configural invariance) and, additionally, that loadings are equal in magnitude across groups/time. Strong invariance holds that configural and weak invariance are present and, additionally, that item thresholds are invariant across groups/time. Finally, strict invariance additionally constrains the residual variances of items to be equal across groups/time, indicating that individual differences in endorsement of items is completely accounted for by the items' respective latent factors.

Longitudinal measurement invariance tests differ from group-difference invariance tests considering that the dependence of observations at different time points is accounted for through allowing facets of the model to be correlated over time and the residual correlations among items across time points to be freely estimated. Based on recommendations by

Cheung & Rensvold (2002) and Chen (2007), models with increasing stringency were not rejected if they resulted in < .01 decline in CFI or < .015 in RMSEA, respectively. Additionally, Chi-square difference tests were also observed but less favored owing to their over-sensitivity in large samples even within the context of no change in CFI and TLI across models (Hu, Bentler, & Kano, 1999). Model-based reliability measures of omega, omega hierarchical, and relative omega were also computed (Reise, Bonifay & Haviland, 2013).

Results

Exploratory Factor Analyses

EFA estimated parameters for models with up to six factors, and bi-EFA included models with one general factor and up to five subfactors. The first eigenvalues ranged between 9.28 and 11.44 across the time points, with first to second eigenvalue ratios of 7.70, 9.12, 10.29, and 12.24 for 2 week, 3 month, 6 month, and 12 month assessments, respectively, suggesting a dominant first factor.

In the EFA, four-factor solutions at each time point evinced interpretable loading patterns and good fit, with some consistency in the factors across time. The EFA factor structure at 2 weeks yielded acute somatic (headaches, dizziness, nausea), cognitive (forgetful, poor concentration, longer to think), visual (blurred vision, light sensitivity, and double vision), and other somatic/emotional symptoms (sleep disturbance, feeling frustrated, fatigue, irritability, depression, noise sensitivity, restlessness) factors. In contrast, at 3 months the four-factor EFA suggested emotional (irritable, depressed, frustrated), cognitive, and visual factors along with a factor representing the remaining somatic symptoms. Emotional and cognitive factors also emerged in the four-factor EFA solution at six months, but one item of the visual factor loaded on a separate somatic factor with dizziness and nausea loading on the former. Lastly, the 12-month, four-factor EFA included emotional and visual factors, a collapsed somatic/cognitive factor, and a two-item factor defined by sleep disturbance and restlessness. Given some consistency across assessments in the emergence of emotional, cognitive, visual, and somatic factors, the 3-month EFA model was viewed as one that should be specified in CFA with data from other time points, in addition to four-factor models with loading patterns based on the time-specific EFA results described above. Five- and six-factor models emerging in EFA were not tested in CFA owing to evidence for factor over-extraction (e.g., Heywood cases, factors with zero or one significant loading).

In contrast to EFA where solutions varied somewhat across time points, four-factor bi-EFA models in particular demonstrated greater consistency in loading patterns at separate assessments. These models supported the presence of a strong general factor (e.g., at 2 weeks loadings of the items on the general factor ranged between .55 and .85). The subfactors of the four-factor bi-EFA model reflected emotional (irritability, depression, frustration), cognitive (forgetfulness, poor concentration, taking longer to think), and visual (blurred vision, sensitivity to light, and double vision) symptom dimensions varying independently from the general factor and from each other. Models with one general factor and one or two subfactors, as well as four or five subfactors, evinced less consistency across assessments than the four-factor bi-EFA model. Thus, only the four-factor bi-EFA model

with one general factor and emotional, cognitive, and visual subfactors was specified in CFA.

Confirmatory Factor Analyses

Higher order factor models.—Fit statistics for higher order factor models (i.e., first order/correlated factor models and second order models) are presented in Table 2. One-factor CFAs are presented for baseline comparison and evaluation of this model's fit given evidence for a prominent factor saturating the domain. Additionally, four-factor correlated factor models derived from EFAs at each respective time point are presented. As noted above, EFA models derived at the 3-month follow-up were included as correlated four-factor model (i.e., emotional, cognitive, visual, and somatic factors) variants in CFA analyses as well. Second order models with one superordinate factor and four first order factors loading upon this second order factor were specified, based on the EFA results for each time point and the 3 month time point.

Absolute fit of one-factor models was not adequate at any time point (RMSEA ≥ .085). Four-factor correlated factor models fit better than one-factor models at each time point, but those specified based on the EFA from the respective time point each fit less well than those based on the 3 month EFA model. The second order model specified with the latter model's first order factors (i.e., emotional, cognitive, visual, and somatic) fit comparably well to these four-factor correlated factor models. Notably, and speaking to the saturation of the mTBI symptom domain by a dominant dimension, the loadings of the somatic, emotional, cognitive, and visual factors on the second order factor ranged between .97 – .99, .90 – .93, .86 – .93, and .81 – .88, respectively, across time points. These findings indicate that, with the possible exception of the visual factor, the first order factors of the five-factor second order model are nearly isomorphic with the superordinate factor. Finally, nested within the four-factor bifactor model were second order models with one superordinate factor, three first order factors (i.e., emotional, cognitive, and visual), and remaining items loading directly on the second order factor. These models did not fit as well as the four-factor bifactor model at any time point (see below).

Bifactor models.—Fit statistics for the four-factor bifactor model, specifying one general factor reflecting the shared variance among all items and three specific, independent subfactors reflecting the structure of remaining variance in certain items are presented in Table 2. The four-factor bifactor model demonstrated excellent fit at each time point (CFI .986 – .995; TLI .982 – .994; RMSEA .041 – .059). Thus, this model showed excellent fit at each point of the recovery period, superiority of fit relative to the second order variant of the model, and greater distinctiveness between its facets than first order factors of the five-factor second order model. At each of the four time points, all items showed strong, significant loadings on the General factor, and three specific factors comprised residual covariance among items assessing Emotional, Cognitive, and Visual symptoms.

Total omega reliability for the General factor, and Emotional, Cognitive, and Visual subfactors, respectively, were .97, .92, .98, and .96 at 2 weeks; .97, 93, .94, .91 at 3 months; .96, .94, .95, .92 at 6 months; and .98, .94, .96, and .93 at 12 months. OmegaH scores

were .93, .18, .25, and .32 at 2 weeks; .94, .16, .21, .27 at 3 months; .96, .13, .15, .23 at 6 months; and .95, .11, .20, and .30 at 12 months, respectively. Table 3 shows a summary of related omega scores. The high omega values, paired with the percentage of common variance explained by the General factor (ECV [New] ranged from 82 – 87%), imply that most of the variance within the RPQ can be attributed to this unidimensional construct. However, some substantive variance was explained by the subfactors, particularly the Visual subfactor, at each time point. Given that 1) the General factor of the four-factor bifactor model accounted for a large proportion of shared variance among RPQ items, 2) specific factors of the bifactor model accounted for non-trivial proportions of variance unaccounted for by the General factor, 3) loadings of first order factors on the second order factor of the second order model were so high that the first order factors were nearly indistinguishable from each other and from the second order factor, and 4) the four-factor bifactor RPQ model demonstrated excellent fit, we favored the four-factor bifactor model as best representing the latent structure of the RPQ.

Measurement Invariance of the Bifactor Model

We compared the fit of invariance models across various demographic (i.e., sex, age, race) and clinical (i.e., reported psychiatric history prior to injury, positive vs. negative head CT findings) groups for each assessment time point to examine how well the constructs reflect the same individual difference dimensions across these groups. Additionally, we conducted longitudinal measurement invariance analyses to evaluate how well factors and subfactors of the four-factor bifactor RPQ model indexed the same constructs at different points in the recovery period in the sample as a whole.

Group invariance.—Measurement invariance of the bifactor model was compared across the demographic groups of sex, age group (i.e., roughly equal groups of 17 – 28, 29 – 48, and 49 – 90+ year-olds), and Caucasian American vs. African American race. For each set of invariance models (i.e., configural, weak, strong, and strict invariance), across each of the demographic and clinical groupings, and at each assessment time point (i.e., 2 weeks, 3 months, 6 months, 12 months) the strict measurement invariance model was deemed the best fitting. Fit statistics for the strict invariance models for each group at each time point are presented in Table 4. For the sake of brevity, fit statistics for all of the four levels of invariance for each group at the four time points are presented in the Supplement (Table 1). One can compare fit statistics of the invariance models for each group at each time point, and in no comparison did any model imposing increasing constraints demonstrate a decrease in CFI or TLI of .01 or greater. Both tables demonstrate that the magnitudes of absolute (all RMSEA < .05) and comparative (all CFI > .990; all TLI > .990) strict invariance model fit indices were excellent as well.

Strict measurement invariance across groups indicates that group differences in endorsement of specific RPQ items are accounted for at the latent factor level. In other words, once controlling for the model's factors, endorsement of each item was equal across groups. This implies that groups differed in severity on the model's factors and did not differ in qualitatively distinctive ways (e.g., in different patterns of endorsement, in salience of specific items to factors).

The strict invariance model allows for estimation of group mean-level differences on each of the factors as well, the magnitudes of which are directly comparable and readily interpretable. To evaluate the potential clinical significance of the bifactor model dimensions, we investigated the group means (from the strict invariance model output) on each factor, reported as standard deviation (*SD*) units in Table 5. Women were higher than men on the General factor by .56, .45, .41, and .33 *SD*s (all *p*s <.001) at the 2-week, 3-month, 6-month, and 12-month time points, respectively. African Americans were consistently higher on the General factor than Caucasian Americans by .37 – .54 *SD*s (all *p*s < .001). Age is presented in Table 6, where the middle age group (i.e., 29 – 48) scored somewhat more highly (~ .25 *SD*, all *p*s < .05) on the General factor than the youngest age group (i.e., 17 – 28) at each time point. The oldest age group (i.e., 49 – 88) was lower than the youngest group at 2 weeks (−.31 *SD*, $p = .040$) and 3 months on the Emotional subfactor (−.58 *SD*, $p = .001$), higher than the youngest group on the Cognitive subfactor at 3 months (.34 *SD*, $p = .021$) and 6 months (.63 *SD*, $p < .001$), and higher than the youngest group on the Visual subfactor at 6 months (.37 *SD*, $p = .038$) and 12 months (.34 *SD*, $p = .047$). Positive psychiatric history was associated with higher General factor means than negative psychiatric history by .39 – .56 *SD* units at each time point (all *p*s <.001). Finally, individuals with positive CT scan results were consistently higher than those with negative findings on the Visual subfactor (.30 – .44 *SD* units, all *p*s <.05).

Longitudinal Invariance.—Longitudinal measurement invariance analyses were conducted to determine to what extent facets of the bifactor model could be interpreted in the same way at the four post-injury time points (i.e., 2 weeks, 3 months, 6 months, 12 months). Fit statistics for the configural, weak, strong, and strict longitudinal invariance models are presented in Supplement Table 1. Neither CFI nor TLI decreased by .01 or more from the configural through strict invariance models, the absolute fit was excellent for each model (RMSEA < .025), and the incremental fit was excellent for each model (CFI > .99; TLI > .99). Therefore, the most parsimonious and excellent fitting strict longitudinal measurement invariance model was deemed the best fitting. Figure 2 presents loading parameters of this model.

Table 5 includes mean-level differences on the bifactor model's General factor and Emotional, Cognitive, and Visual subfactors over time, relative to the 2-week assessment time point. Expectedly, scores on the General factor decreased from the 2-week to 3-month assessment (−.46 *SD*, $p <.001$), from the 2-week to 6-month assessment (−.60 *SD*, $p < .001$), and from the 2-week to 12-month assessment (−.64, $p < .001$). The greatest decline in the General factor was from 2 weeks to 3 months, followed by a decline smaller in magnitude from 3 months to 6 months, and leveling off from 6 months to 12 months. Interestingly, the Emotional subfactor increased from 2 weeks to 3 months (.75 *SD*, $p < .001$), dropped by .10 *SD* from 3 months to 6 months, and remained at that level through 12 months. The Cognitive subfactor increased from 2 weeks to 3 months (.36 *SD*, $p < .001$), doubled to .74 *SD* ($p < .001$) by the 6-month assessment, and dropped a small degree by the 12-month assessment (.59 *SD*, $p <.001$). Finally, the Visual subfactor did not show significant change in mean levels across the four time points.

Discussion

The purposes of this study were to elucidate the factor structure of the RPQ and to determine the degree to which its structure was invariant across time and groups of interest among individuals with mTBI. Whereas prior work has been limited to correlated factor structures, we additionally considered second order factor models and bifactor models. The four-factor bifactor model provided excellent fit to the data and explicitly accounted for the largely unidimensional structure of the RPQ while simultaneously delineating specific dimensions reflecting remaining multidimensional components of the domain not adequately explained by the General factor. The results of our study provide evidence that, as expected, the RPQ is largely unidimensional. This conclusion is supported by an excellent fitting bifactor model and omega estimates indicating that ≥ 82% of the variance in RPQ total scores is explained by the General factor. In addition to supporting the continued use of total scores as an index of overall mTBI symptom burden, our model helps illustrate why one-factor models of the RPQ are ill-fitting and explains why correlated factor models of the RPQ yield factors with high correlations and minimal discriminant validity.

Additionally, we tested the equivalency of bifactor model parameters across group variables and over time. Measurement invariance analyses revealed that parameters of the four-factor bifactor model of the RPQ were equivalent over time and across sex, age, race, psychiatric history, and mTBI severity (CT+/−) groups. This is perhaps surprising, given the dynamic nature of mTBI recovery. However, strict measurement invariance indicates that differences in RPQ item endorsement over time and across groups are completely attributable to the latent factors of the model. In the case of the General factor, for example, differences in total scores on the RPQ across recovery phases and diverse patient subgroups likely represent valid time/group differences as opposed to instrument-related measurement discrepancies. In addition to finding that a dominant General factor underlies RPQ symptom endorsement, strict measurement invariance models indicate that differences over time and between patient groups were quantitative rather than qualitative – that is, patients with mTBI vary in severity of symptom dimensions more so than by patterns of symptom endorsement. This has implications for current research efforts to develop precision medicine approaches. In particular, the strong covariance among symptoms due to the presence of a dominant general factor may make it difficult to identify distinct clinical phenotypes of mTBI on the basis of observed symptoms. The bifactor model proposed here detected distinct symptom phenotypes that could better develop precision-medicine approaches to stratify and treat patients with mTBI.

Because observed mTBI symptom ratings vary in degree (i.e., severity) more than in kind (i.e., clusters of symptom endorsement), efforts to distinguish among subtypes of mTBI on the basis of self-reported symptom endorsement patterns alone may be unproductive. Clinical phenotyping efforts may instead benefit from using the distinct bifactor model dimensions identified in this work or integrating symptom dimensions with a more diverse array of clinical assessment data from additional modalities.

Despite the strong General factor underlying RPQ symptom ratings, a significant minority of variance was attributable to other, independent dimensions affecting ratings of Emotional,

Cognitive, and Visual symptoms. To better understand the clinical value of the specific factors identified by the bifactor model, we estimated mean differences in factor scores across time and between groups defined by demographic and clinical characteristics. Consistent with the natural history of recovery from mTBI (Carroll et al., 2014), general factor scores showed a steady decline in mTBI symptoms over time, with the most dramatic change occurring within the first three months post-injury. General factor scores also correlated with demographic variables in expected ways (i.e., with similar patterns to what has been reported for observed mTBI symptom severity scores). In particular, General factor scores were higher in females, individuals with pre-injury psychiatric disorders, and African Americans, consistent with research on mTBI outcomes after trauma (Dischinger et al., 2009; Preiss-Farzanegan, et al., 2009). Visual specific factor scores were higher for patients with objective evidence of brain injury/more severe mTBI (CT+) and remained stable over time, implying perhaps that this subfactor provides a symptom-based marker of injury severity. Though specific factors within our model offered sensible constructs of TBI symptom sets, the low internal consistency reliability estimates for each subfactor suggest that their clinical and research utility may be limited. Further research may be fruitful to determine the degree to which these specific factors could be made more reliable (e.g., with the addition of other like-items). Given their limited reliability, it is surprising and perhaps promising that we observed meaningful, significant group differences in specific factor score estimates.

Several study limitations should be noted. Our data may have been missing at random (MAR) or not missing at random (NMAR), types of missingness that can bias WLSMV estimates. However, because of the ordinal, highly non-normally distributed aspects of our data, estimates based on full information maximum likelihood estimation (more robust to MAR data) would have been biased due to not meeting distributional assumptions (Beauducel et al., 2006; Savalei, 2010). Another limitation may be that 3-month follow-up appointments were conducted via phone (vs. predominantly in-person at other time points). Mitigating this concern somewhat, the presence of strict measurement invariance across time provides evidence that the RPQ consistently measures mTBI symptom dimensions regardless of mode of assessment. In addition, the sample was restricted to patients with admission GCS 13–15, and the extent to which these findings generalize to individuals with more severe TBIs is unknown. Moreover, because the study only enrolled patients who presented to a Level I trauma center and had a head CT ordered by the treating physician, it is unclear to what degree the findings would generalize to the broader mTBI population (e.g., those who do not have a head CT or do not present to a trauma center for care, or military populations exposed to additional environmental threats). However, given our relatively broad definition of mTBI and demonstration of strict invariance of the resulting RPQ structural model across time and patient groups, one might expect the model to fit in other subgroups, individuals with dual-diagnoses (e.g., PTSD) and in the mTBI population as a whole.

Future Directions

This study provides further evidence that the RPQ is sufficiently unidimensional to continue to use total scores to represent overall mTBI symptom burden, although there are

multidimensional aspects of the domain that additionally account for smaller proportions of variance in symptom ratings. These findings align with our work showing a high degree of unidimensionality of mTBI symptoms among young athletes with sport-related injuries who were assessed with a different symptom checklist (SCAT; Nelson et al., 2018), suggesting that this general finding may hold across diverse assessment measures and mTBI subpopulations. The consistency of the RPQ bifactor model's parameters across time and patient characteristics demonstrates its utility in comparing scores across key demographic and clinical subgroups, as well as across various phases of mTBI recovery. Despite a dominant first factor, the four-factor bifactor model was able to parse meaningful subsidiary factors that contributed more modestly to RPQ ratings in the domains of Cognitive, Emotional, and Visual symptoms. Further exploration of these specific symptom dimensions is warranted; in particular, Visual factor scores may represent a novel symptom-based index associated with objective measures of mTBI severity, such as positive head CT scans. Overall, these findings, alongside related work with the SCAT in young athletes with mTBI, imply that patients vary more in severity than type of symptoms, which may hinder attempts to identify categorically distinctive groups of mTBI patients from observed symptom ratings. Efforts to characterize individuals based on profiles of the bifactor model dimensions identified in this study may facilitate a more precision medicine-based approach to classify and treat patients with mTBI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The TRACK-TBI study (PI: Manley) was funded by the U.S. National Institutes of Health, National Institute of Neurological Disorders and Stroke (Grant # U01 NS1365885), OneMind, and NeuroTrauma Sciences, LLC. This secondary data analysis project was supported by National Institutes of Health grants 1R03NS100691 and R01NS110856 (PI: Nelson) and the U.S.

Department of Defense TBI Endpoints Development (TED) Initiative (Grant # W81XWH-14-2-0176; PI: Manley).

References

Bazarian JJ, Wong T, Harris M, Leahey N, Mookerjee S, & Dombovy M (1999). Epidemiology and predictors of post-concussive syndrome after minor head injury in an emergency population. Brain Injury, 13(3), 173–189. [PubMed: 10081599]

Beauducel A, & Herzberg PY (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. Structural Equation Modeling, 13(2), 186–203.

Binder LM (1986). Persisting symptoms after mild head injury: a review of the postconcussive syndrome. Journal of Clinical and Experimental Neuropsychology, 8(4), 323–346. [PubMed: 3091631]

Boake C, McCauley SR, Levin HS, Pedroza C, Contant CF, Song JX, … & Diaz-Marchan PJ (2005). Diagnostic criteria for postconcussional syndrome after mild to moderate traumatic brain injury. The Journal of Neuropsychiatry and Clinical Neurosciences, 17(3), 350–356. [PubMed: 16179657]

Carroll L, Cassidy JD, Peloso P, Borg J, Von Holst H, Holm L, … & Pépin M (2004). Prognosis for mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. Journal of Rehabilitation Medicine, 36(0), 84–105.

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*: A Multidisciplinary Journal, 14(3), 464–504.

Cnossen MC, Winkler EA, Yue JK, Okonkwo DO, Valadka AB, Steyerberg EW, … & Manley GT (2017). Development of a prediction model for post-concussive symptoms following mild traumatic brain injury: A TRACK-TBI pilot study. Journal of Neurotrauma, 34(16), 2396–2409. [PubMed: 28343409]

Dikmen S, Machamer J, & Temkin N (2017). Mild traumatic brain injury: longitudinal study of cognition, functional status, and post-traumatic symptoms. Journal of neurotrauma, 34(8), 1524–1530. [PubMed: 27785968]

Dischinger PC, Ryb GE, Kufera JA, & Auman KM (2009). Early predictors of postconcussive syndrome in a population of trauma patients with mild traumatic brain injury. Journal of Trauma and Acute Care Surgery, 66(2), 289–297.

Duhaime AC, Gean AD, Haacke EM, Hicks R, Wintermark M, Mukherjee P, Brody D, Latour L, & Riedy G (2010). Common data elements in radiologic imaging of traumatic brain injury. Archives of Physical Medicine and Rehabilitation, 91(11), 1661–1666. [PubMed: 21044709]

Eyres S, Carey A, Gilworth G, Neumann V, & Tennant A (2005). Construct validity and reliability of the Rivermead post-concussion symptoms questionnaire. Clinical Rehabilitation, 19(8), 878–887. [PubMed: 16323387]

Giza CC, & Hovda DA (2014). The new neurometabolic cascade of concussion. Neurosurgery, 75(suppl_4), S24–S33. [PubMed: 25232881]

Gordon WA, Haddad L, Brown M, Hibbard MR, & Sliwinski M (2000). The sensitivity and specificity of self-reported symptoms in individuals with traumatic brain injury. Brain Injury, 14, 21–33. [PubMed: 10670659]

Herrmann N, Rapoport MJ, Rajaram RD, Chan F, Kiss A, Ma AK, … & Lanctôt KL (2009). Factor analysis of the Rivermead Post-Concussion Symptoms Questionnaire in mild-to-moderate traumatic brain injury patients. The Journal of Neuropsychiatry and Clinical Neurosciences, 21(2), 181–188. [PubMed: 19622689]

Hiploylee C, Dufort PA, Davis HS, Wennberg RA, Tartaglia MC, Mikulis D, … & Tator CH (2017). Longitudinal study of postconcussion syndrome: not everyone recovers. Journal of Neurotrauma, 34(8), 1511–1523. [PubMed: 27784191]

Hu LT and Bentler PM (1999), "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," Structural Equation Modeling, 6(1), 1–55.

MacCallum RC, Browne MW, and Sugawara H,M (1996), "Power Analysis and Determination of Sample Size for Covariance Structure Modeling," Psychological Methods, 1 (2), 130–49.

King NS (2014). A systematic review of age and gender factors in prolonged post-concussion symptoms after mild head injury. Brain injury, 28(13–14), 1639–1645. [PubMed: 25265040]

King NS, Crawford S, Wenden FJ, Moss NEG, & Wade DT (1995). The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms commonly experienced after head injury and its reliability. Journal of Neurology, 242(9), 587–592. [PubMed: 8551320]

King NS (1996). Emotional neuropsychological and organic factors: their use in the prediction of persisting postconcussion symptoms after moderate to mild head injury. Journal of Neurology, Neurosurgery, and Psychiatry, 61, 75–81.

King NS (1997). Mild head injury: Neuropathology sequelae measurement and recovery. British Journal of Clinical Psychology, 36, 161–184.

Lannsjö M, Borg J, Björklund G, Af Geijerstam JL, & Lundgren-Nilsson Å (2011). Internal construct validity of the rivermead post-concussion symptoms questionnaire. Journal of Rehabilitation Medicine, 43(11), 997–1002. [PubMed: 22031345]

Li C (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least square. Behavior Research Methods 48(3), 936–949. [PubMed: 26174714]

Lishman WA (1988) Physiogenesis and psychogenesis in the 'post-concussional syndrome'. British Journal of Psychiatry, 153, 460–469.

MacCallum RC, Browne MW, and Sugawara H,M (1996), "Power Analysis and Determination of Sample Size for Covariance Structure Modeling," Psychological Methods, 1(2), 130–49.

Mac Donald CL, Adam OR, Johnson AM, Nelson EC, Werner NJ, Rivet DJ, & Brody DL (2015). Acute post-traumatic stress symptoms and age predict outcome in military blast concussion. Brain, 138(5), 1314–1326. [PubMed: 25740219]

Mittenberg W, & Strauman S (2000). Diagnosis of mild head injury and the postconcussion syndrome. The Journal of Head Trauma Rehabilitation, 15(2), 783–791. [PubMed: 10739967]

Nelson LD, Kramer MD, Patrick CJ, & McCrea MA (2018). Modeling the structure of acute sport-related concussion symptoms: a bifactor approach. Journal of the International Neuropsychological Society, 1–12.

Potter S, Leigh E, Wade D, & Fleminger S (2006). The Rivermead post concussion symptoms questionnaire. Journal of Neurology, 253(12), 1603–1614. [PubMed: 17063314]

Preiss-Farzanegan SJ, Chapman B, Wong TM, Wu J, & Bazarian JJ (2009). The relationship between gender and postconcussion symptoms after sport-related mild traumatic brain injury. PM&R, 1(3), 245–253. [PubMed: 19627902]

Rabinowitz AR, Li X, McCauley SR, Wilde EA, Barnes A, Hanten G, … & Levin HS (2015). Prevalence and predictors of poor recovery from mild traumatic brain injury. Journal of Neurotrauma, 32(19), 1488–1496. [PubMed: 25970233]

Reise SP, Bonifay WE, & Haviland MG (2013). Scoring and modeling psychological measures in the presence of multidimensionality. Journal of Personality Assessment, 95(2), 129–140. [PubMed: 23030794]

Reise SP, Widaman KF, & Pugh RH (1993). Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. Psychological Bulletin, 114(3), 552. [PubMed: 8272470]

Rimel RW, Giordani B, Barth JT, Boll TJ, & Jane JA (1981). Disability caused by minor head injury. Neurosurgery, 9(3), 221–228. [PubMed: 7301062]

Savalei V (2010). Expected versus observed information in SEM with incomplete normal and nonnormal data. Psychological Methods, 15(4), 352. [PubMed: 20853954]

Smith-Seemiller L, Fow NR, Kant R, & Franzen MD (2003). Presence of post-concussion syndrome symptoms in patients with chronic pain vs mild traumatic brain injury. Brain Injury, 17(3), 199–206. [PubMed: 12623496]

Thompson C, Davies P, Herrmann L, Summers M, & Potter S (2016). Approaches to establishing validated cut-off scores on the Rivermead post-concussion symptoms questionnaire (RPQ). Brain Injury, 30(5–6), 770).

Voormolen DC, Cnossen MC, Polinder S, von Steinbuechel N, Vos PE, & Haagsma JA (2018). Divergent Classification Methods of Post-Concussion Syndrome after Mild Traumatic Brain Injury: Prevalence Rates, Risk Factors, and Functional Outcome. Journal of Neurotrauma, 35(11), 1233–1241. [PubMed: 29350085]

Yue JK, Vassar MJ, Lingsma HF, Cooper SR, Okonkwo DO, Valadka AB, … & Puccio AM (2013). Transforming research and clinical knowledge in traumatic brain injury pilot: multicenter implementation of the common data elements for traumatic brain injury. Journal of Neurotrauma, 30(22), 1831–1844. [PubMed: 23815563]

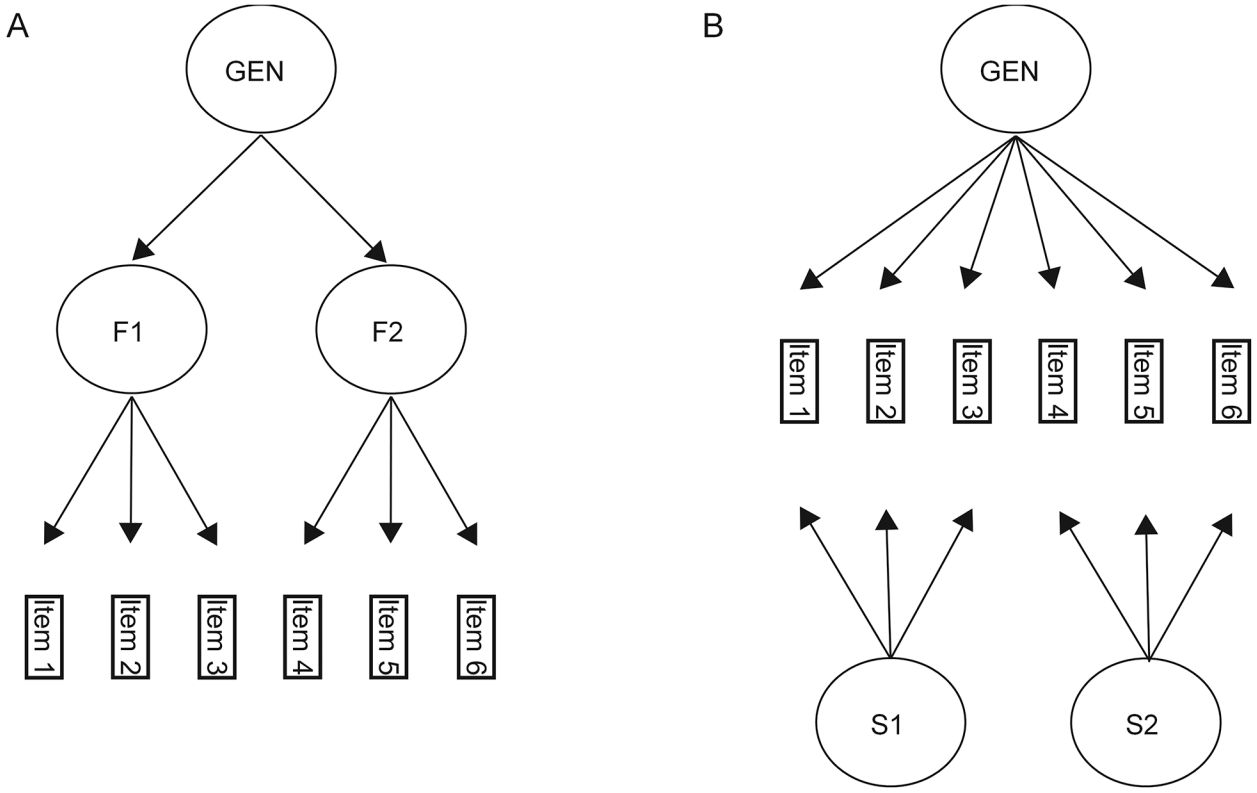

Figure 1.
Comparison between hypothetical second order factor model (A) and bifactor model (B). Used in Nelson et al. (2018).

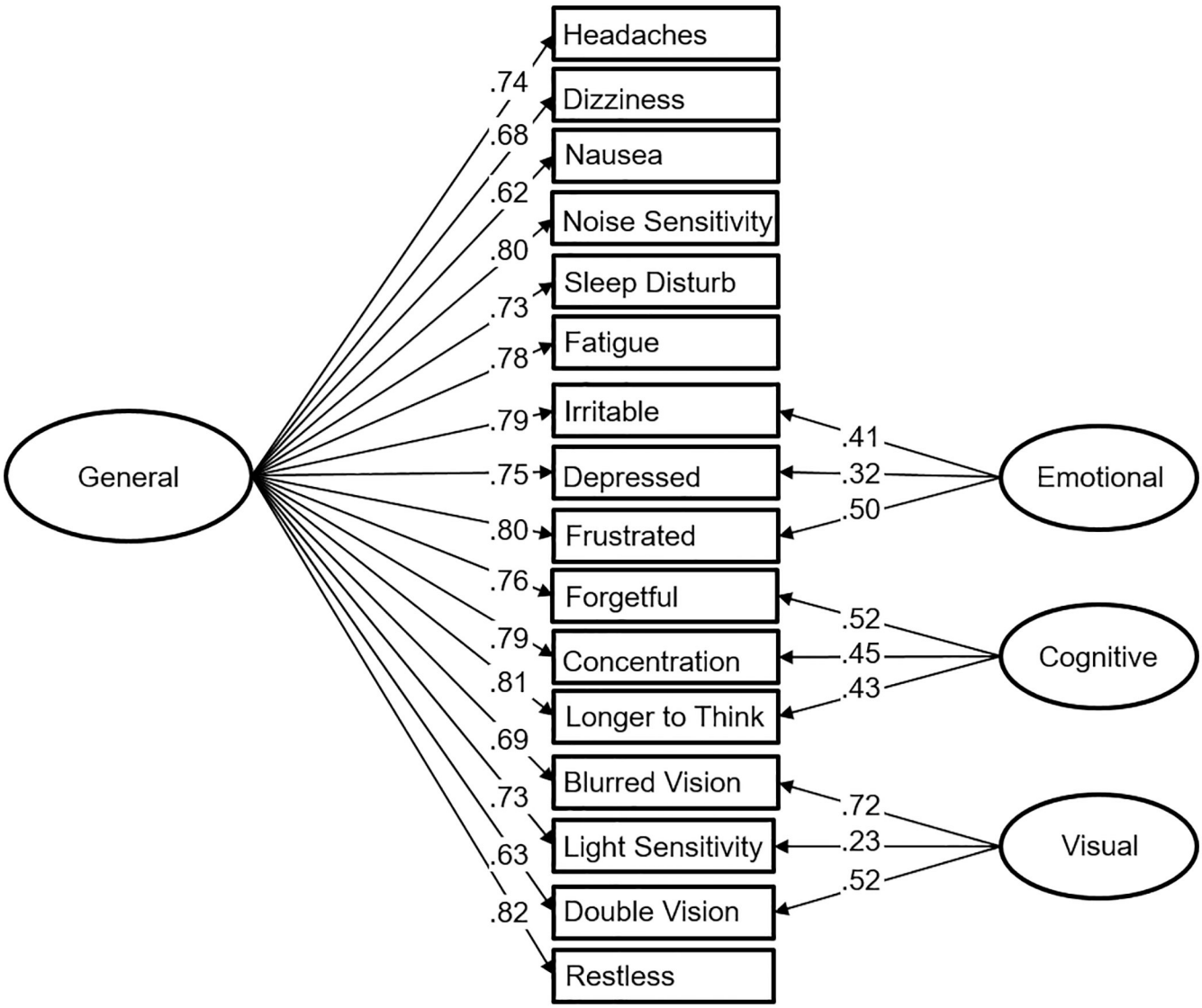

Figure 2.

Factor loadings of the strict measurement invariance four-factor bifactor model over time. Loadings differed slightly across T1 – T4 owing to standardization. These estimates reflect parameters from T1 2 weeks.

Table 1.

Sample Characteristics (N = 1,011)

Demographics/history	N (%), M (SD)	Injury-Related Variables	N (%), M (SD)
Female sex	355 (35.1%)	*Cause of Injury*	
Age (range 17–90)	39.74 (16.67)	Road traffic accident	712 (61.6%)
Ethnicity		Incidental fall	277 (21.0%)
Non-Hispanic	796 (78.7%)	Other nonintentional	60 (5.2%)
Hispanic	211 (20.9%)	Violence/assault	69 (5.2%)
Missing	4 (0.4%)	Mass violence	1 (0.1%)
Race		Other	34 (2.9%)
Caucasian American	769 (76.1%)	Missing	2 (0.2%)
African American	175 (17.3%)	*Loss of Consciousness*	
Native American	2 (0.2%)	Yes	760 (75.2%)
Asian American	34 (3.4%)	No	141 (13.9%)
Other/Multiracial	26 (2.6%)	Suspected/Unknown	108 (10.6%)
Missing	5 (0.5%)	Missing	2 (0.2%)
Years of education	13.55 (2.97)	*Post-traumatic Amnesia*	
Psychiatric History		Yes	683 (67.6%)
Yes	211 (20.9%)	No	203 (20.1%)
No	796 (78.7%)	Suspected/Unknown	124 (12.2%)
Missing	4 (0.4%)	Missing	2 (0.2%)
Prior TBI		*Positive Head CT*	
No	767 (66.4%)	Yes	300 (29.7%)
Yes	337 (29.2%)	No	696 (68.8%)
Missing	51 (4.4%)	Missing	15 (1.5%)

Note. N = sample size, M = mean, SD = standard deviation. GCS arrival score 13–15 criteria with at least one completed follow-up assessment for study inclusion. TBI = traumatic brain injury; PTA = post-traumatic amnesia; CT = computerized tomography. Sample was adult mTBI including six 17-year old patients.

Table 2.

Fit Statistics for Confirmatory Factor Models at Each Time Point

Model	χ^2	df	RMSEA	CFI	TLI
T1 2 Weeks					
Higher Order Models					
Correlated factor					
1-factor	992.4	104	.117	.939	.930
4-factor[a]	387.3	98	.069	.980	.976
4-factor[b]	373.1	98	.067	.981	.977
Second order					
4-factor (1SO, 3FO)[c]	353.9	102	.063	.983	.980
5-factor (1SO, 4FO)[a]	451.0	100	.075	.976	.971
5-factor (1SO, 4FO)[b]	383.7	100	.067	.981	.977
Bifactor Model					
4-factor (1G, 3S)	**301.3**	**95**	**.059**	**.986**	**.982**
T2 3 Months					
Higher Order Models					
Correlated factor					
1-factor	664.8	104	.095	.965	.960
4-factor[b]	280.5	98	.056	.989	.986
Second order					
4-factor (1SO, 3FO)[c]	338.3	102	.062	.985	.983
5-factor (1SO, 4FO)[b]	274.7	100	.054	.989	.987
Bifactor Model					
4-factor (1G, 3S)	**238.1**	**95**	**.050**	**.991**	**.989**
T3 6 Months					
Higher Order Models					
Correlated factor					
1-factor	522.1	104	.085	.976	.972
4-factor[a]	245.0	98	.052	.991	.989
4-factor[b]	214.9	98	.047	.993	.992
Second order					
4-factor (1SO, 3FO)[c]	278.0	102	.056	.990	.988
5-factor (1SO, 4FO)[a]	266.0	100	.055	.990	.988
5-factor (1SO, 4FO)[b]	226.6	100	.048	.993	.991
Bifactor Model					
4-factor (1G, 3S)	**207.6**	**95**	**.046**	**.993**	**.992**
T4 12 Months					

Model	χ^2	df	RMSEA	CFI	TLI
Higher Order Models					
Correlated factor					
1-factor	639.3	104	.100	.967	.962
4-factor[a]	410.1	98	.079	.981	.976
4-factor[b]	244.7	98	.054	.991	.989
Second order					
4-factor (1SO, 3FO)[c]	273.2	102	.057	.989	.987
5-factor (1SO, 4FO)[a]	404.1	100	.077	.981	.977
5-factor (1SO, 4FO)[b]	255.4	100	.055	.990	.988
Bifactor Model					
4-factor (1G, 3S)	**177.2**	**96**	**.041**	**.995**	**.994**

Note: χ^2 = chi-square statistic; *df* = degrees of freedom; RMSEA = root mean square error of approximation statistic; CFI = comparative fit index; TLI = Tucker-Lewis index; SO = second order factor; FO = first order factor; G = general factor; S = specific factor;

[a] derived from the time point's exploratory factor analysis;

[b] consistent with 3-month exploratory factor analysis;

[c] derived from consistent exploratory bifactor model at each time point

Table 3.

Omega Scores for Four-Factor Bifactor Strict Measurement Invariance Model across Time

Model Factor	ECV (New)	Omega/ OmegaS	OmegaH/ OmegaHS	Relative Omega
T1 2 weeks				
General factor	.82	.97	.93	.96
s1 Emotional	.22	.91	.20	.22
s2 Cognitive	.26	.94	.24	.26
s3 Visual	.37	.89	.30	.34
T2 3 months				
General factor	.85	.98	.95	.97
s1 Emotional	.17	.93	.15	.16
s2 Cognitive	.23	.95	.21	.23
s3 Visual	.34	.92	.28	.31
T3 6 months				
General factor	.87	.98	.95	.97
s1 Emotional	.14	.94	.13	.13
s2 Cognitive	.18	.95	.17	.18
s3 Visual	.31	.92	.26	.28
T4 12 months				
General factor	.86	.98	.95	.97
s1 Emotional	.12	.94	.11	.12
s2 Cognitive	.21	.96	.20	.20
s3 Visual	.35	.93	.30	.32

Note: ECV (New)= explained common variance; Omega/OmegaS = internal reliability of the multidimensional composite; Omega H/OmegaHS: proportion of variance in given scale attributed to individual differences in General factor

Table 4.

Fit Statistics for Strict Measurement Invariance of the Four-Factor Bifactor Model across Sex, Age, Psychiatric History, CT Scan Results, and Time

Group	χ^2	df	RMSEA	CFI	ΔCFI	TLI	ΔTLI	χ^2 diff p
Sex								
T1 2 weeks	378.6	259	.031	.994	.005	.994	.004	.111
T2 3 months	309.2	258	.021	.998	.	.998	.	.397
T3 6 months	330.1	259	.026	.997	.004	.997	.004	.456
T4 12 months	309.4	259	.023	.998	.003	.998	.003	.584
Age								
T1 2 weeks	557.2	423	.032	.994	.005	.995	.005	.195
T2 3 months	484.1	421	.023	.997	.003	.998	.003	.375
T3 6 months	648.6	423	.044	.992	−.002	.993	−.002	.000
T4 12 months	500.6	423	.027	.997	.003	.998	.003	.312
Race								
T1 2 weeks	362.3	259	.030	.995	.006	.995	.006	.637
T2 3 months	341.2	259	.028	.996	.001	.996	.001	.030
T3 6 months	361.8	259	.032	.996	.002	.996	.002	.020
T4 12 months	301.3	259	.021	.998	.003	.998	.003	.689
Psychiatric history								
T1 2 weeks	374.0	260	.031	.995	.005	.995	.004	.198
T2 3 months	353.1	259	.029	.996	.002	.996	.002	.019
T3 6 months	305.4	259	.021	.998	.003	.998	.003	.340
T4 12 months	308.1	259	.022	.998	.003	.998	.003	.540
CT Scan								
T1 2 weeks	403.5	259	.035	.993	.006	.994	.007	.185
T2 3 months	301.8	258	.020	.998	.003	.998	.003	.550
T3 6 months	366.9	259	.032	.996	.002	.996	.002	.015
T4 12 months	314.0	259	.024	.998	.003	.998	.003	.463
Time								
T1 – T4	2525.9	1925	.018	.992	.002	.992	.002	.023

Note. χ^2 = chi-square statistic; *df* = degrees of freedom; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; CT = computerized tomography. Age was stratified into terciles. CT scan results were coded as the presence or absence of acute intracranial findings. *df*s at T2 3 months for sex, age, and CT scan differ from other time points owing to collapsing nausea item responses for one of the groups.

Table 5.

Mean Differences in Strict Measurement Invariance Bifactor Model Factors by Sex, Race, Psychiatric History, CT Scan Results, and Time

Model Factor	**T1 2 weeks**			**T2 3 months**			**T3 6 months**			**T4 12 months**		
Sex	M	F	p	M	F	p	M	F	p	M	F	p
General factor	**.00**	**.56**	**<.001**	**.00**	**.45**	**<.001**	**.00**	**.41**	**<.001**	**.00**	**.33**	**<.001**
s1 Emotional	.00	−.10	.407	.00	−.08	.544	.00	.08	.584	.00	−.19	.214
s2 Cognitive	.00	.10	.381	.00	.04	.729	.00	−.04	.781	.00	.14	.268
s3 Visual	**.00**	**−.28**	**.024**	.00	.12	.442	.00	.18	.210	.00	.05	.733
Race	AA	CA	p	AA	CA	p	AA	CA	p	AA	CA	p
General factor	**.00**	**−.37**	**<.001**	**.00**	**−.54**	**<.001**	**.00**	**−.44**	**<.001**	**.00**	**−.42**	**<.001**
s1 Emotional	.00	−.23	.104	.00	−.14	.364	.00	−.01	.937	.00	−.11	.536
s2 Cognitive	.00	−.05	.706	.00	.06	.696	.00	.09	.593	.00	−.04	.779
s3 Visual	.00	−.11	.434	.00	−.10	.535	.00	−.29	.081	.00	−.14	.371
Psych history	Neg	Pos	p	Neg	Pos	p	Neg	Pos	p	Neg	Pos	p
General factor	**.00**	**.39**	**<.001**	**.00**	**.43**	**<.001**	**.00**	**.56**	**<.001**	**.00**	**.44**	**<.001**
s1 Emotional	.00	.18	.194	.00	.17	.224	.00	.06	.704	.00	.18	.272
s2 Cognitive	**.00**	**.26**	**.043**	.00	.14	.259	.00	.17	.244	.00	.23	.099
s3 Visual	.00	−.20	.151	.00	.10	.524	.00	−.26	.073	.00	−.18	.251
CT Scan	Neg	Pos	p	Neg	Pos	p	Neg	Pos	p	Neg	Pos	p
General factor	.00	.07	.371	.00	−.07	.463	.00	−.09	.317	.00	−.04	.700
s1 Emotional	**.00**	**−.37**	**.006**	.00	−.10	.489	.00	.25	.117	.00	.00	.995
s2 Cognitive	.00	−.12	.274	.00	.21	.081	**.00**	**.43**	**.003**	.00	.02	.903
s3 Visual	**.00**	**.30**	**.014**	**.00**	**.41**	**.013**	**.00**	**.44**	**.005**	**.00**	**.39**	**.010**
Time			p			p			p			p
General factor		**.00**	--		**−.46**	**<.001**		**−.60**	**<.001**		**−.64**	**<.001**
s1 Emotional		**.00**	--		**.75**	**<.001**		**.65**	**<.001**		**.66**	**<.001**
s2 Cognitive		**.00**	--		**.36**	**<.001**		**.74**	**<.001**		**.59**	**<.001**
s3 Visual		.00	--		.00	.987		.16	.147		−.02	.887

Note. AA = African American; CA = Caucasian; M = male; F = female; *p* = significance of χ^2 difference test; Psych = medical record psychiatric history; Neg = negative; Pos = positive; CT = computerized tomography.

Means of zero to two decimal places indicate fixed parameters in the first group and differences are represented in standard deviation units relative to these means. Estimates in bold emphasize significance at $p < .05$.

Table 6.

Mean Differences in Strict Measurement Invariance Bifactor Model Factors across Age Groups

Model Factor	**T1 2 weeks**					**T2 3 months**				
Age	Age1	Age2	p	Age3	p	Age1	Age2	P	Age3	p
General factor	**.00**	**.28**	**.002**	−.03	.736	**.00**	**.28**	**.004**	.15	.145
s1 Emotional	.00	−.23	.095	**−.31**	**.040**	**.00**	**−.44**	**.009**	**−.58**	**.001**
s2 Cognitive	**.00**	**−.32**	**.013**	.07	.603	.00	.17	.230	**.34**	**.021**
s3 Visual	.00	.25	.077	**.39**	**.011**	.00	−.13	.462	.17	.359
Model Factor	T3 6 months					T4 12 months				
Age	Age1	Age2	p	Age3	p	Age1	Age2	p	Age3	p
General factor	**.00**	**.23**	**.031**	.12	.257	**.00**	**.23**	**.036**	.15	.164
s1 Emotional	.00	−.18	.315	−.27	.150	.00	−.05	.789	−.21	.235
s2 Cognitive	.00	.10	.521	**.63**	**<.001**	.00	.00	.986	.27	.096
s3 Visual	.00	.21	.192	**.37**	**.038**	.00	.084	.638	**.34**	**.047**

Note. Age1 = 17 – 28; Age2 = 29 – 48; Age3 = 49 – 88. *p* = significance of χ^2 difference test for age group 2 or 3 versus 1. Means of zero to two decimal places indicate fixed parameters in the first group and differences are represented in standard deviation units relative to these means. Estimates in bold emphasize significance at $p<.05$.