

Published in final edited form as:

Sleep Med. 2012 June ; 13(6): 752–758. doi:10.1016/j.sleep.2012.02.014.

Associations between Pittsburgh Sleep Quality Index Factors and Health Outcomes in Women with Posttraumatic Stress Disorder

Melynda D. Casement^{*,a}, Kelly M. Harrington^a, Mark W. Miller^a, and Patricia A. Resick^{**,a}

National Center for PTSD at VA Boston Healthcare System and Boston University Medical School

Abstract

Objective—The Pittsburgh Sleep Quality Index (PSQI) is a widely used measure of subjective sleep disturbance in clinical populations, including individuals with posttraumatic stress disorder (PTSD). Although the severity of sleep disturbance is generally represented by a global symptom score, recent factor analytic studies suggest that the PSQI is better characterized by a two- or three-factor model than a one-factor model. This study examined the replicability of two- and three-factor models of the PSQI, as well as the relationship between PSQI factors and health outcomes, in a female sample with PTSD.

Methods—The PSQI was administered to 319 women with PTSD related to sexual or physical assault. Confirmatory factor analyses tested the relative fit of one-, two-, and three-factor solutions. Bivariate correlations were performed to examine the shared variance between PSQI sleep factors and measures of PTSD, depression, anger, and physical symptoms.

Results—Confirmatory factor analyses supported a 3-factor model with Sleep Efficiency, Perceived Sleep Quality, and Daily Disturbances as separate indices of sleep quality. The severity of symptoms represented by the PSQI factors was positively associated with the severity of PTSD, depression, and physical symptoms. However, these health outcomes correlated as much or more with the global PSQI score as with PSQI factor scores.

Conclusions—These results support the multidimensional structure of the PSQI. Yet, the global PSQI score has as much or more explanatory power as individual PSQI factors in predicting health outcomes.

Keywords

factor analysis; Pittsburgh Sleep Quality Index; sleep; PTSD; depression; anger; physical symptoms

^{*}Correspondence during manuscript review should be addressed to Melynda Casement, National Center for PTSD, VA Boston Healthcare System, 150 S. Huntington Ave. (116B-4), Boston, MA 02130-4893; Telephone: 857-364-2311; Fax: 857-364-6524; Melynda.Casement@va.gov. ^{**}Correspondence following manuscript publication should be addressed to Patricia Resick, National Center for PTSD, VA Boston Healthcare System, 150 S. Huntington Ave. (116B-3), Boston, MA 02130; Telephone: 857-364-4129; Fax: 857-364-4515; Patricia.Resick@va.gov.

^aMailing address for all authors: VA Boston Healthcare System, 150 S. Huntington Ave. (116B), Boston, MA 02130., in order of authorship: Melynda.Casement@va.gov; Kelly.Harrington@va.gov; Mark.Miller5@va.gov; Patricia.Resick@va.gov

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

1. Introduction

The Pittsburgh Sleep Quality Index (PSQI) [1] is considered an essential measure of sleep and insomnia symptoms in treatment research, and it is a recommended assessment tool for both epidemiological studies and studies that address the mechanisms of sleep disorders [2]. It was developed, in part, to provide a user-friendly and clinically-oriented measure of sleep quality. The 18-items on the PSQI index seven clinically-formulated sleep domains: sleep duration, sleep disturbance, sleep latency, daytime disturbance, habitual sleep efficiency, sleep quality, and use of sleep medications. Scores on each of the seven PSQI subscales are used to calculate a global score that ranges from 0 to 21. Global scores above 5 distinguish poor sleepers from good sleepers with high sensitivity (90% to 99%) and specificity (84% to 87%) [1,3].

Despite broad use of the PSQI in clinical and research settings, relatively few studies examine the factor structure of the instrument [4–7]. The few that have indicate that the PSQI is better represented by two- or three-factor models than a single factor [4–7]. For example, Cole et al. [4] examined the factor structure of the PSQI using a combination of exploratory factor analysis and confirmatory factor analysis in a large sample of older adults. Results suggested that the seven PSQI subscales are best represented by three latent factors: Sleep Efficiency, Perceived Sleep Quality, and Daily Disturbances. The Sleep Efficiency factor included the sleep duration and habitual sleep efficiency subscales of the PSQI; the Perceived Sleep Quality factor included the sleep quality, sleep latency, and sleep medication subscales of the PSQI; and the Daily Disturbances factor included the sleep disturbance and daytime dysfunction subscales of the PSQI. Cole et al. [4] also found marginal fit for a two-factor model that combined Perceived Sleep Quality and Daily Disturbances into a single factor. These two- and three-factor models were later replicated by Magee et al. [6] using exploratory and confirmatory factor analysis in a sample of Australian adults. However, Magee et al. [6] argued for the two-factor model rather than the three-factor model due to evidence of overlap between the Perceived Sleep Quality and Daily Disturbances factors. The multidimensionality of the PSQI has also been supported by principal component analysis [5,7]. For example, principal component analysis performed by Kotronoulas et al. [5] supported a two-factor model of the PSQI in which Factor 1 included all of the PSQI component scores with the exception of sleep medication and daytime dysfunction. Each of these studies suggests that a multidimensional PSQI scoring system may be preferable to the traditional PSQI global score. However, the best representation of the PSQI factor structure remains unclear. Structural models of the PSQI need to be replicated and extended to a range of patient populations before they can be broadly applied. It is also important to consider whether PSQI factor scores provide incremental value beyond the PSQI global score. Clinicians and researchers who use the PSQI as a screening instrument for sleep disorders and a tool for case conceptualization may well wonder whether using a multifactorial scoring system will improve the PSQI's convergent or discriminant validity. However, the relationship between PSQI factor scores and other mental and physical health outcomes has not been established.

The primary aim of the present study was to compare the relative fit of alternative two- and three-factor PSQI structural models in data from a sample of individuals with posttraumatic stress disorder (PTSD). Understanding the performance of the measure in a sample with PTSD is important because of the high prevalence and wide variety of sleep problems in this patient population. Of individuals diagnosed with PTSD in the general population, 44% report difficulty initiating sleep, 91% report difficulty maintaining sleep, and 52% report nightmares [8]. This represents a two- to five-fold increase in the rate of these sleep complaints relative to healthy individuals [8]. Moreover, the concurrence of sleep disruption

and PTSD predicts poor clinical outcomes [9,10]. Sleep disruptions in PTSD need to be accurately characterized in order for basic research and clinical interventions to progress, and the PSQI is often the primary measure of sleep in descriptive studies of PTSD [11,12] and PTSD intervention trials [13,14].

The secondary aim of the present study was to evaluate the convergent and discriminant validity of the best-fitting PSQI factor model. We accomplished this by comparing patterns of association between observed PSQI factor scores and observed scores on measures of PTSD, depression, anger, and physical symptoms. Given prior evidence that poor sleep quality assessed with the PSQI is associated with increased negative affect (e.g., anxiety, depression, anger) [15–17] and physical health concerns [16,18–20], we anticipated a moderate to high degree of association between PSQI factor scores and external health outcomes. We also expected to find evidence of discriminant associations between PSQI factor scores and these outcomes. Specifically, based on evidence that sleep quality is more predictive of negative affect and physical health than sleep quantity [16], we predicted that mental and physical health outcomes would have a stronger correlation with the Perceived Sleep Quality factor score than the Sleep Efficiency factor score.

Our third and final aim was to evaluate the incremental value of the PSQI factor scores in relation to the global PSQI score. To accomplish this, we compared the patterns of association between observed PSQI factor scores and observed scores on mental and physical health indices to the pattern of association between the PSQI global score and the same external symptom measures. We anticipated that some external symptoms would have stronger correlations with certain PSQI factor scores than with the PSQI global score. More specifically, because the Daily Disturbances factor includes a PSQI subscale with several questions about physical health complaints, we expected scores on our external measure of physical health complaints to have a stronger correlation with the Daily Disturbances factor score than the PSQI global score.

2. Methods

2.1. Participants

Participants were 319 females enrolled in one of two clinical trials of trauma-focused cognitive behavioral therapy that were conducted sequentially in the same location [21,22]. All participants were victims of sexual assault (86%) or physical assault (14%) and had a current diagnosis of PTSD. Participants were at least three months post-trauma at the time of assessment. About half of the participants (44%) met criteria for current major depressive disorder at the time of initial assessment in addition to the diagnosis of PTSD. Exclusionary criteria included history of substance dependence with active use within the last six months, current substance abuse, medication instability, current abuse or stalking, current psychosis, current suicidal intent or parasuicidal behavior, and illiteracy. Participants were diverse in terms of age (range = 18 to 74, $M = 33.68$, $SD = 11.27$), years of education (range = 0–24, $M = 14.10$, $SD = 2.58$), race (30% African American, 1% Asian American, 67% Caucasian, 1% Native American, and 2% other or mixed race), marital status (49% single, 22% married or cohabitating, 26% separated or divorced, 2% widowed), and annual household income (55% < \$20,000, 30% \$20,000–\$50,000, 15% > \$50,000). This research was approved by local institutional review boards and followed appropriate procedures for informed consent.

2.2. Measures

The Pittsburgh Sleep Quality Index (PSQI) [1] was used to assess sleep disturbance over the past month. The PSQI is an 18-item self-report questionnaire. The items produce seven component scores which range from 0 (no difficulty) to 3 (severe difficulty): sleep duration,

sleep disturbance, sleep latency, daytime disturbance, habitual sleep efficiency, sleep quality, and use of sleep medications. The sum of these component scores yields a measure of global sleep quality which ranges from 0 to 21. Reliability measures indicate that the PSQI generally has acceptable internal consistency ($\alpha = .80$ to $.85$) [1,3,23] and test-retest reliability ($r = .85$ to $.87$) [1,3]. The PSQI also has acceptable concurrent validity; scores on the PSQI are highly correlated with scores on other subjective measures of sleep quality ($r > .69$) [23]. The PSQI and all other self-report symptom questionnaires were administered as part of a baseline assessment battery prior to treatment initiation.

The Clinician Administered PTSD Scale (CAPS) [24] was used to confirm PTSD diagnosis and measure PTSD symptom severity during the month prior to assessment. The CAPS is a 30-item, semi-structured, clinician-administered assessment of PTSD that is consistent with the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [25]. The frequency and intensity of each of the 17 DSM-IV PTSD symptoms are rated on separate 0–4 point scales. The severity of PTSD symptoms is indexed by the sum of the frequency and intensity scores across items 1 through 17. Diagnosis of PTSD requires at least one re-experiencing symptom, three avoidance symptoms, and two hyperarousal symptoms each with a frequency score of 1 or greater and an intensity score of 2 or greater (i.e., 1–2 scoring rule) [26]. The 1–2 scoring rule on the CAPS distinguishes individuals with a clinician-administered diagnosis of PTSD from those without a PTSD diagnosis with high sensitivity (74%) and specificity (84%) [27]. Reliability measures indicate that the CAPS has acceptable internal consistency ($\alpha = .94$) and test-retest reliability ($r = .90$ to $.98$) [24]. The CAPS also has acceptable concurrent validity; scores on the CAPS are highly correlated with scores on other PTSD assessments ($r = .73$ to $.83$) [28]. The CAPS was administered during study screening.

The Beck Depression Inventory (BDI) was used to measure the severity of depressive symptoms. The BDI-I [29] was administered to participants in the first clinical trial ($n = 169$) and the BDI-II [30] was administered to participants in the second clinical trial ($n = 147$). The BDI-II reflects changes to the diagnostic criteria for depression in the fourth revision of the Diagnostic and Statistical Manual of Mental Disorders [25]. Specifically, BDI items assessing body image, hypochondria, and difficulty working were replaced, and most of the other items were reworded. Both the BDI-I and the BDI-II are 21-item self-report questionnaires. Scores on both measures range from 0 to 63. Reliability measures indicate that the BDI-I and BDI-II have acceptable internal consistency (BDI-I $\alpha = .86$; BDI-II $\alpha = .91$) and test-retest reliability (BDI-I $r = .48$ to $.86$, BDI-II $r = .93$) [30,31]. The BDI also has acceptable concurrent validity; scores on the BDI are highly correlated with scores on other subjective measures of depression (BDI-I $r = .55$ to $.96$, BDI-II $r = .71$) [30,31]. The coefficient of correlation between scores on the BDI-I and BDI-II is $.93$ [32]. Therefore, we chose to examine BDI scores across the combined study sample (using BDI-I scores from the sample in the first clinical trial and BDI-II scores from the sample in the second clinical trial) rather than performing separate analyses for the BDI-I and BDI-II.

The Trait Anger scale of the State-Trait Anger Expression Inventory (STAXI) [33] was used to assess the propensity to experience and express anger. The Trait Anger scale includes 10-items, and composite scores range from 10 to 40. Scores of 22 or higher indicate high propensity to express anger (75th percentile or above in females). Reliability measures indicate that the Trait Anger scale has high internal consistency ($\alpha = .82$) [34]. The Trait Anger scale also has acceptable concurrent validity; scores on the Trait Anger scale are moderately correlated with scores on other subjective measures of hostility ($r = .27$ to $.73$) [34].

The Pennebaker Inventory of Limbic Languidness (PILL) was used to measure the severity of physical symptoms [35]. The PILL is a 54-item self-report Likert scale questionnaire that assesses the frequency of a broad range of common physical sensations (e.g., racing heart, headache, soreness). Scores range from 0 to 216. Reliability measures indicate that the PILL has acceptable internal consistency ($\alpha = .91$) and test-retest reliability ($r = .83$) [35]. The PILL also has acceptable concurrent validity; scores on the PILL are moderately correlated with scores on other subjective measures of physical symptoms ($r = .48$) [35].

2.3. Statistical analyses

2.3.1. Confirmatory factor analysis—Replicating the series of confirmatory factor analyses performed by Cole et al. [4] and Magee et al. [6], we specified one-factor, two-factor, and three-factor models to compare competing representations of the structure of subscales on the PSQI. The use of subscale scores rather than item-level data is consistent with previous factor analyses of this measure [4–7]. In the one-factor model all seven PSQI subscales were loaded onto a single latent variable. In the two-factor model the PSQI subscales for sleep duration and habitual sleep efficiency were loaded onto the latent variable Sleep Efficiency; the PSQI subscales for sleep disturbance, sleep latency, daytime disturbance, sleep quality, and use of sleep medications were loaded onto the latent variable Perceived Sleep Quality. The latent variable Perceived Sleep Quality was subdivided into two latent factors in the three-factor model: Perceived Sleep Quality and Daily Disturbances. Thus, in the three-factor model, PSQI subscales for sleep duration and habitual sleep efficiency were loaded onto the latent variable Sleep Efficiency; the PSQI subscales for sleep quality, sleep latency, and sleep medications were loaded onto the latent variable Perceived Sleep Quality; and the PSQI subscales for sleep disturbances and daytime dysfunction were loaded onto the latent variable Daily Disturbances. *T*-tests were performed to evaluate the difference in the magnitude of associations between pairs of correlations in the three-factor model [36]. For example, we examined whether the strength of the correlation between Sleep Efficiency and Perceived Sleep Quality was statistically different from the strength of the correlation between Sleep Efficiency and Daily Disturbances.

We also tested the fit of an alternative two-factor model presented by Kotronoulas et al. [5]. In the Kotronoulas et al. two-factor model, the PSQI subscales for subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, and sleep disturbances loaded onto Factor 1; the PSQI subscales for use of sleep medications and daytime dysfunction loaded onto Factor 2 [5]. We considered testing the fit of an alternative three-factor model presented by Aloba et al. [7]. However, this model was ultimately omitted from confirmatory analyses because its factors were not well differentiated; several of the PSQI subscales loading highly ($> .40$) onto more than one factor in the principal component analysis [7]. This limited interpretation of the factors and our ability to compare Aloba et al.'s three-factor model to the other models tested here.

Confirmatory factor analyses were performed using Mplus statistical modeling software, version 5.1 [37]. The PSQI subscale scores were treated as ordinal measures. Confirmatory factor analyses with the PSQI subscale scores used the mean and variance-adjusted weighted least squares (WLSMV) estimator. A built-in four-step procedure was used to model partial missing data for 21 (7%) of the study participants. Measurement error for all indicators was assumed to be uncorrelated, and correlations were permitted between latent variables. All *p*-values reported for *t*-statistics are two-tailed.

Goodness of fit was evaluated using absolute fit (χ^2), root mean square error of approximation (RMSEA), weighted root mean residual (WRMR), comparative fit index (CFI), and Tucker-Lewis Index (TLI). The following criteria defined adequate model fit: ratio of χ^2 to *df* < 3, RMSEA < .06, WRMR < .90, CFI > .95, and TLI > .96 [38]. In

addition, nested chi-square difference tests were performed to statistically compare the fit of the one-, two-, and three-factor solutions for the PSQI.

2.3.2. Correlation analyses—To address study aims two and three, bivariate correlation analyses were performed to assess the amount of shared variance between observed scores on the PSQI and observed scores on the CAPS, BDI, STAXI Trait Anger scale, and PILL. Observed scores for the PSQI were derived from the best-fitting factor model suggested by the confirmatory factor analysis results. Specifically, the sum of the PSQI subscale scores for sleep duration and habitual sleep efficiency composed the Sleep Efficiency observed score; the sum of PSQI subscale scores for sleep quality, sleep latency, and sleep medications composed the Perceived Sleep Quality observed score; and the sum of PSQI subscale scores for sleep disturbances and daytime dysfunction composed the Daily Disturbances observed score. The sum of all of the PSQI subscale scores composed the standard PSQI global score. Total scores were used for the CAPS, BDI, STAXI Trait Anger scale, and PILL.

Fisher's *z*-tests for independent samples were used to determine whether some PSQI factor scores accounted for more shared variance in mental and physical health outcomes than other PSQI factor scores. Fisher's *z*-tests were also used to determine whether PSQI factor scores accounted for more shared variance in mental and physical health outcomes than the PSQI global score. All *p*-values reported for *z*-statistics are two-tailed.

3. Results

3.1. Sample characteristics

A *t*-test indicated that PSQI scores did not differ significantly between samples in the two intervention trials included in study analyses, $t(288) = 0.89$, $p = .38$. Therefore, data were pooled across samples for all analyses. As expected, most participants (85%) reported poor sleep quality (PSQI scores > 5). Table 1 presents mean scores on the PSQI, CAPS, BDI, STAXI Trait Anger scale, and PILL for the combined sample. Regular use of sleep medications (scores of 2 or 3 on the PSQI sleep medication subscale) was reported by 27% of study participants.

3.2. Confirmatory factor analysis

Table 2 presents goodness-of-fit statistics and chi-square difference tests for the one-, two-, and three-factor models of the PSQI. Goodness-of-fit statistics indicated that the one-factor model and Kotronoulas et al.'s [5] two-factor model provided poor fit to the PSQI data. Goodness-of-fit statistics were acceptable for two out of five goodness-of-fit indices (i.e., WRMR, CFI) for the two-factor model advocated by Magee et al. [6]. In contrast, all goodness-of-fit statistics with the exception of RMSEA (marginal at .07) indicated that the three-factor model provided acceptable fit to the PSQI data. Furthermore, chi-square difference tests indicated that the three-factor model provided a significantly better fit to the PSQI data than either of the two-factor models. Standardized parameter estimates for the best-fitting three-factor model are presented in Figure 1. All freely estimated standardized parameters were statistically significant ($p < .05$).

T-tests comparing the magnitude of the correlations between latent PSQI factors indicated that the association between Sleep Efficiency and Perceived Sleep Quality ($r = .77$) was greater in magnitude than the association between Sleep Efficiency and Daily Disturbances ($r = .46$), $t(316) = 16.22$, $p < .001$. In addition, the association between Daily Disturbances and Perceived Sleep Quality ($r = .82$) was greater in magnitude than the association between Daily Disturbances and Sleep Efficiency ($r = .46$), $t(316) = 18.45$, $p < .001$. The magnitude

of the association between Perceived Sleep Quality and Sleep Efficiency ($r = .77$) did not differ significantly from the magnitude of the association between Perceived Sleep Quality and Daily Disturbances ($r = .82$), $t(316) = 1.94$, $p = .06$.

3.3. Correlation analyses

Table 3 presents correlation coefficients between assessment instruments. The first set of Fisher's z -tests indicated that some PSQI factor scores accounted for more shared variance in mental and physical health outcomes than other PSQI factor scores. Specifically, the magnitude of the correlation between the Daily Disturbance and BDI scores ($r = .47$) was significantly larger than the magnitude of the correlation between the Perceived Sleep Quality and BDI scores ($r = .34$), $z = 1.95$, $p = .05$, and the Sleep Efficiency and BDI scores ($r = .22$), $z = 3.44$, $p = .001$. Similarly, the magnitude of the correlation between the Daily Disturbance and PILL scores ($r = .53$) was significantly larger than the magnitude of the correlation between the Perceived Sleep Quality and PILL scores ($r = .34$), $z = 3.00$, $p = .003$, and the Sleep Efficiency and PILL scores ($r = .20$), $z = 4.80$, $p < .001$. Notably, all of the PSQI factor scores accounted for more shared variance in the CAPS, BDI, and PILL than the STAXI ($p < .05$ for all contrasts). Furthermore, the Sleep Efficiency score accounted for more shared variance in the CAPS ($r = .36$) than the PILL ($r = .20$), $z = 2.09$, $p = .04$, while the Daily Disturbances score accounted for more shared variance in the PILL ($r = .53$) than the CAPS ($r = .37$), $z = 2.50$, $p = .01$. All other contrasts failed to reach the threshold for statistical significance ($p > .05$).

The second set of Fisher's z -tests indicated that the PSQI global score accounted for as much or more variance in external measures of health as the PSQI factor scores. Specifically, the magnitude of the correlations between the global PSQI score and scores on the CAPS, BDI, STAXI, and PILL were generally statistically equivalent to the magnitudes of the correlations between the PSQI factors scores and scores on the same mental and physical health indices ($p > .05$). There were two exceptions. Specifically, the magnitude of the correlation between the global PSQI and BDI scores ($r = .41$) was significantly larger than the magnitude of the correlation between the Sleep Efficiency and BDI scores ($r = .22$), $z = 2.52$, $p = .01$. Likewise, the magnitude of the correlation between the global PSQI and PILL scores ($r = .43$) was significantly larger than the magnitude of the correlation between the Sleep Efficiency and PILL scores ($r = .20$); $z = 3.05$, $p = .002$. Finally, as with the PSQI factor scores, the PSQI global score accounted for more shared variance in the CAPS, BDI, and PILL than the STAXI ($p < .05$ for all contrasts). All other contrasts failed to reach the threshold for statistical significance ($p > .05$).

4. Discussion

The present study compared the relative fit of one-, two-, and three-factor measurement models of the PSQI. It also examined the convergent and discriminant validity of each PSQI factor score and the PSQI global score in relation to measures of PTSD, depression, anger, and physical symptoms. The results provide additional evidence that sleep characteristics assessed by the PSQI are best-represented by three latent factors: Sleep Efficiency, Perceived Sleep Quality, and Daily Disturbances [4]. This replicates and extends the three-factor model from Cole et al.'s [4] sample of older adults with generally good sleep quality to a treatment-seeking female sample with PTSD and generally poor sleep quality. In addition, the two-factor model advocated by Magee et al. [6] provided marginal fit to the PSQI data in this sample. In contrast, the PSQI was not well represented by a one-factor model or the two-factor model proposed by Kotronoulas et al. [5].

While the three-factor model provided adequate fit to sleep characteristics assessed by the PSQI, much of the variance in the sleep medication indicator could not be accounted for by

the latent variable Perceived Sleep Quality. Nor could Daily Disturbances account for most of the variance in daytime dysfunction. This suggests that sleep medication and daytime dysfunction may not be the best indicators of their purported constructs, Perceived Sleep Quality and Daily Disturbances, respectively. The sleep medication subscale of the PSQI may be particularly limited as an indicator of latent sleep constructs. Though this subscale score has a moderate correlation with the global PSQI score ($r = .62$) [1], it was a poor indicator of latent sleep constructs in previous factor analyses of the PSQI [4,6] and two factor analyses that pooled sleep items across multiple self-report instruments [39,40]. Sleep medication use may be a poor indicator of latent sleep constructs due to the variety of applications for sleep medications (e.g., problems with daytime functioning, long sleep latency, interrupted sleep). Furthermore, many individuals with significant sleep problems avoid sleep medications due to side effects or the availability of behavioral treatment options.

These results also indicate that the Sleep Efficiency and Daily Disturbances factors are both strongly correlated with the Perceived Sleep Quality factor. The strong degree of association between PSQI factors is consistent with factor analytic results from other study samples [4,6]. The lack of factor independence, coupled with inconsistent factor analytic results between studies [4–7], suggests that the PSQI may be limited as a measure of latent sleep constructs. This may be due, in part, to the low number of PSQI indicators (i.e., two to three) per factor. Notably, two-factor solutions were derived in factor analyses that pooled sleep items across multiple self-report instruments [39–40]. The PROMIS (Patient Reported Outcomes Information System) sleep inventory validated by Buysse et al. had 16 to 27 observed indicators per latent factor [39]. The index used by Koffel and Watson had six to 13 indicators per latent factor [40]. Furthermore, these conglomerate sleep indices were both well-characterized by a factor representing sleep disturbance (i.e., Sleep Disturbance in Buysse et al.'s model, Insomnia in Koffel and Watson's model), and by a factor representing sleep-related daytime impairment (i.e., Sleep Related Impairment in Buysse et al.'s model; Lassitude in Koffel and Watson's model). Measures of sleep that have more than seven observed indicators may be better equipped than the PSQI to accurately and consistently represent latent sleep constructs.

Correlation analyses indicated that individuals with poor sleep quality tended to have more severe PTSD symptoms, depressive symptoms, and physical health concerns. This was consistent with our expectations. In contrast, scores on the PSQI were not significantly associated with Trait Anger on the STAXI. Furthermore, correlations between PSQI scores and scores on the CAPS, BDI, and PILL were significantly larger in magnitude than correlations between PSQI scores and the STAXI. Trait Anger on the STAXI has been associated with sleep disruption in previous studies with nonclinical samples [15,17]. The lack of significant association between Trait Anger and sleep quality in the present study may reflect differences between combined-gender nonclinical samples with relatively good sleep quality and treatment-seeking female samples with PTSD and poor sleep quality.

There was limited support for the hypothesis that some PSQI factor scores would account for more variance in mental and physical health outcomes than other PSQI factor scores. Specifically, z -tests indicated that Daily Disturbances accounted for more shared variance in BDI and PILL scores than what could be accounted for by either Sleep Efficiency or Perceived Sleep Quality. This may reflect overlap between symptoms assessed by the Daily Disturbance factor (i.e., trouble performing daytime activities and interruptions in sleep due to physical or psychological symptoms) and those assessed by the BDI and PILL. However, in contrast to our hypothesis, Perceived Sleep Quality did not account for more shared variance in external health measures than Sleep Efficiency.

We did not find evidence to support the hypothesis that PSQI factor scores would provide superior discriminant validity compared to the global PSQI score. In fact, the magnitude of the correlations between the global PSQI score and other health indices was always as large or larger than the magnitude of the correlations between the PSQI factor scores and other health indices. In short, the global PSQI score was at least as good as any of the three individual PSQI factor scores at predicting health outcomes. This result suggests that a three-factor PSQI scoring system may have little incremental value to researchers and clinicians who currently use the PSQI global score to inform case conceptualization and predict coincident health outcomes.

There are a number of limitations to this study. In particular, we cannot rule out the possibility that PSQI factors may offer incremental improvement in discriminant validity for measures of mental and physical health other than those included in the present study. Furthermore, while we believe it was appropriate to test measurement models of the PSQI in a sample with PTSD because PTSD is associated with severe and varied forms of sleep disturbance, the results presented here may not generalize to other populations. Finally, we could not perform structural equation modeling to examine the PSQI factors in relation to other health indices because the model would be underidentified.

As noted above, the PSQI is one of the most commonly used assessments of sleep quality, and it is considered essential in studies of treatment outcome [2]. Therefore, changes to the scoring structure of the PSQI would have broad implications. While the results of the present factor analysis suggest that a three-factor model of the PSQI provides relatively better fit than one- or two-factor solutions, the utility of the three-factor model is dubious. The correlational analyses presented here indicate that the PSQI global score is as good or better as any of the PSQI factor scores at accounting for shared variance with measures of PTSD, depression, and physical symptoms. Thus, a three-factor scoring model may not add predictive value over the PSQI global score. This, as well as practical barriers to scoring and interpreting three individual PSQI factor scores, do little to recommend the three-factor PSQI scoring system proposed by Cole et al. [4] for general use. Clinicians and researchers using the PSQI may wish to continue using the global score rather than the three-factor scoring system if their primary aim is to identify clinically significant sleep disturbance or predict the severity of negative affect and health symptoms in a psychiatric sample. Nonetheless, future studies that examine the relationship between PSQI factor scores and clinical outcomes in other samples are warranted and may demonstrate the value of a three-factor PSQI scoring system. In addition, if a three-factor model accurately represents latent sleep constructs, it should be possible to identify factors representing Sleep Efficiency, Perceived Sleep Quality, and Daily Dysfunction with other measures. Ultimately, emerging assessments of sleep quality, such as the PROMIS sleep inventory [39], may provide an alternative to the PSQI that is both clinically and psychometrically valid, and more straightforward to score and interpret.

Acknowledgments

This research was supported in part by National Institute of Mental Health Grant R01-MH51509, awarded to Patricia A. Resick at the University of Missouri- St. Louis. Support for the first and second authors was provided by National Institute of Mental Health Institutional Training Grant T32-MH019836, awarded to Terence M. Keane at the National Center for PTSD, Boston, MA. We thank all of the contributors to this work, including the research participants, assessors, therapists, research assistants, fidelity raters, and data managers.

References

1. Buysse DJ, Reynolds CF, Monk TH, Berman SR. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Res.* 1989; 28:193–213. [PubMed: 2748771]
2. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep.* 2006; 29:1155–1173. [PubMed: 17040003]
3. Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J Psychosom Res.* 2002; 53:737–740. [PubMed: 12217446]
4. Cole JC, Motivala SJ, Buysse DJ, Oxman MN, Levin MJ, Irwin MR. Validation of a 3-factor scoring model for the Pittsburgh sleep quality index in older adults. *Sleep.* 2006; 29:112–116. [PubMed: 16453989]
5. Kotronoulas GC, Papadopoulou CN, Papapetrou A, Patiraki E. Psychometric evaluation and feasibility of the Greek Pittsburgh Sleep Quality Index (GR-PSQI) in patients with cancer receiving chemotherapy. *Support Care Cancer.* 2011; 19:1831–1840. [PubMed: 20972588]
6. Magee CA, Caputi P, Iverson DC, Huang X-F. An investigation of the dimensionality of the Pittsburgh Sleep Quality Index in Australian adults. *Sleep Biol Rhythms.* 2008; 6:222–227.
7. Aloba OO, Adewuya AO, Ola BA, Mapayi BM. Validity of the Pittsburgh Sleep Quality Index (PSQI) among Nigerian university students. *Sleep Med.* 2007; 8:266–270. [PubMed: 17368977]
8. Ohayon MM, Shapiro CM. Sleep disturbances and psychiatric disorders associated with posttraumatic stress disorder in the general population. *Compr Psychiatry.* 2000; 41:469–478. [PubMed: 11086154]
9. Belleville G, Guay S, Marchand A. Persistence of sleep disturbances following cognitive-behavior therapy for posttraumatic stress disorder. *J Psychosom Res.* 2011; 70:318–327. [PubMed: 21414451]
10. Belleville G, Guay S, Marchand A. Impact of sleep disturbances on PTSD symptoms and perceived health. *J Nerv Ment Dis.* 2009; 197:126–132. [PubMed: 19214048]
11. Calhoun PS, Wiley M, Dennis MF, Means MK, Edinger JD, Beckham JC. Objective evidence of sleep disturbance in women with posttraumatic stress disorder. *J Trauma Stress.* 2007; 20:1009–1018. [PubMed: 18157880]
12. Germain A, Buysse DJ, Shear MK, Fayyad R, Austin C. Clinical correlates of poor sleep quality in posttraumatic stress disorder. *J Trauma Stress.* 2004; 17:477–484. [PubMed: 15730066]
13. Cook JM, Harb GC, Gehrman PR, Cary MS, Gamble GM, Forbes D, et al. Imagery rehearsal for posttraumatic nightmares: A randomized controlled trial. *J Trauma Stress.* 2010; 23:553–563. [PubMed: 20839311]
14. Krakow B, Hollifield M, Johnston L, Koss M, Schrader R, Warner TD, et al. Imagery rehearsal therapy for chronic nightmares in sexual assault survivors with posttraumatic stress disorder: A randomized controlled trial. *JAMA.* 2001; 286:537–545. [PubMed: 11476655]
15. Stewart JC, Rand KL, Hawkins MAW, Stines JA. Associations of the shared and unique aspects of positive and negative emotional factors with sleep quality. *Pers Individ Differ.* 2011; 50:609–614.
16. Pilcher JJ, Ginter DR, Sadowsky B. Sleep quality versus sleep quantity: relationships between sleep and measures of health, well-being and sleepiness in college students. *J Psychosom Res.* 1997; 42:583–596. [PubMed: 9226606]
17. Shin C, Kim J, Yi H, Lee H, Lee J, Shin K. Relationship between trait-anger and sleep disturbances in middle-aged men and women. *J Psychosom Res.* 2005; 58:183–189. [PubMed: 15820847]
18. Galovski TE, Monson C, Bruce SE, Resick PA. Does cognitive-behavioral therapy for PTSD improve perceived health and sleep impairment? *J Trauma Stress.* 2009; 22:197–204. [PubMed: 19466746]
19. Magee CA, Caputi P, Iverson DC. Relationships between self-rated health, quality of life and sleep duration in middle aged and elderly Australians. *Sleep Med.* 2011; 12:346–350. [PubMed: 21388876]

20. Shankar A, Charumathi S, Kalidindi S. Sleep duration and self-rated health: the national health interview survey 2008. *Sleep*. 2011; 34:1173–1177. [PubMed: 21886354]
21. Resick PA, Galovski TE, O'Brien Uhlmansiek M, Scher CD, Clum GA, Young-Xu Y. A randomized clinical trial to dismantle components of cognitive processing therapy for posttraumatic stress disorder in female victims of interpersonal violence. *J Consult Clin Psychol*. 2008; 76:243–258. [PubMed: 18377121]
22. Resick PA, Nishith P, Weaver TL, Astin MC, Feuer CA. A comparison of cognitive-processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *J Consult Clin Psychol*. 2002; 70:867–879. [PubMed: 12182270]
23. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh Sleep Quality Index. *J Psychosom Res*. 1998; 45:5–13. [PubMed: 9720850]
24. Blake DD, Weathers FW, Nagy LM, Kaloupek DG. The development of a Clinician-Administered PTSD Scale. *J Trauma Stress*. 1995; 8:75–90. [PubMed: 7712061]
25. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (4th ed., text revision)*. Washington, DC: Author; 2000.
26. Weathers FW, Ruscio AM, Keane TM. Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychol Assessment*. 1999; 11:124–133.
27. Hovens JE, Van der Ploeg HM, Klaarenbeek MTA, Bramsen I. The assessment of posttraumatic stress disorder: With the Clinician Administered PTSD Scale: Dutch results. *J Clin Psychol*. 1994; 50:325–340. [PubMed: 8071438]
28. Foa EB, Tolin DF. Comparison of the PTSD Symptom Scale-Interview Version and the Clinician-Administered PTSD Scale. *J Trauma Stress*. 2000; 13:181–191. [PubMed: 10838669]
29. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961; 4:561–571. [PubMed: 13688369]
30. Beck, AT. *Manual for Beck Depression Inventory II (BDI-II)*. San Antonio, TX: Psychology Corporation; 1996.
31. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clin Psychol Rev*. 1988; 8:77–100.
32. Dozois DJA, Dobson KS, Ahnberg JL. A psychometric evaluation of the Beck Depression Inventory–II. *Psychol Assessment*. 1998; 10:83–89.
33. Spielberger, CD.; Sydeman, SJ. *State-Trait Anxiety Inventory and State-Trait Anger Expression Inventory*. In: Maruish, ME., editor. *The Use of Psychological Testing for Treatment Planning and Outcome Assessment*. Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc; 1994. p. 292-321.
34. Spielberger, CD. *State-Trait Anger Expression Inventory Research Edition Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.; 1988.
35. Pennebaker, JW. *The Psychology of Physical Symptoms*. New York: Springer-Verlag; 1982.
36. Cohen, J.; Cohen, P. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum; 1983. Bivariate correlation and regression; p. 25-78.
37. Muthén, LK.; Muthén, BO. *Mplus User's Guide*. Fifth Edition. Los Angeles, CA: Muthén & Muthén; 1998.
38. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *J Educ Res*. 2006; 99:323–337.
39. Buysse DJ, Yu L, Moul DE, Germain A, Stover A, Dodds NE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep*. 2010; 33:781–792. [PubMed: 20550019]
40. Koffel E, Watson D. The two-factor structure of sleep complaints and its relation to depression and anxiety. *J Abnorm Psychol*. 2009; 118:183–194. [PubMed: 19222324]

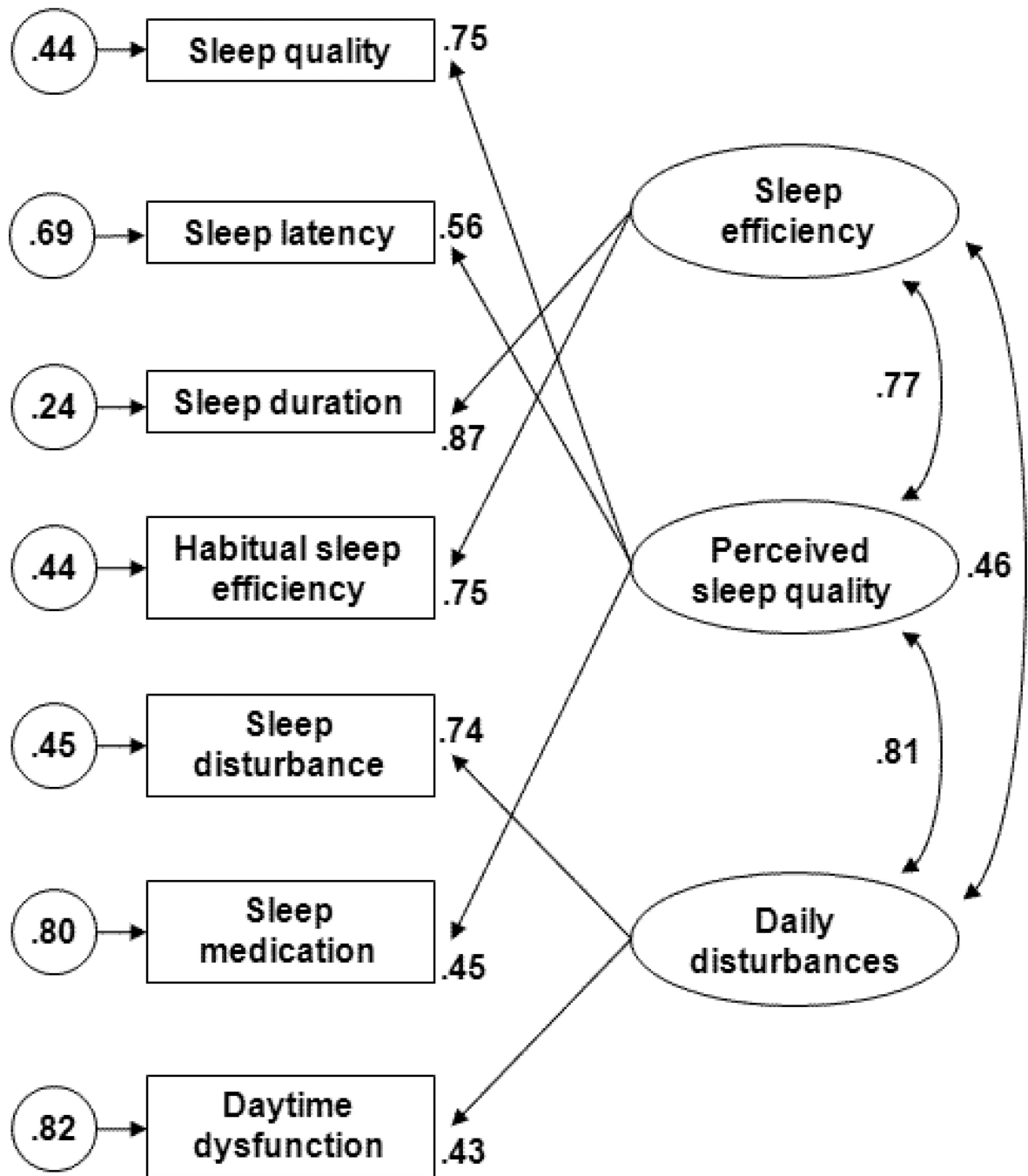


Figure 1. Factor structure of sleep characteristics assessed with the Pittsburgh Sleep Quality Index. Ovals represent latent variables, rectangles represent observed indicators, and circles represent the variance that is unique to each indicator.

Table 1

Mean Scores on Assessment Instruments

Measure	Mean	SD	N
PSQI Sleep Efficiency	2.83	2.05	301
PSQI Perceived Sleep Quality	4.75	2.12	306
PSQI Daily Disturbances	3.68	1.13	303
PSQI Global Score	11.26	4.08	290
CAPS	72.90	18.88	314
BDI	24.90	10.36	316
STAXI Trait Anger	20.25	5.96	312
PILL	130.90	33.70	312

BDI, Beck Depression Inventory; CAPS, Clinician Administered PTSD Scale; PILL, Pennebaker Inventory of Limbic Languidness; PSQI, Pittsburgh Sleep Quality Index; STAXI, State-Trait Anger Expression Index.

Table 2

Goodness-of-fit Indices for PSQI Measurement Models

Model	χ^2 (df)	RMSEA	WRMR	CFI	TLI	Model Comparison	χ^2 diff
1. 1 factor	58.02 (11)	0.12	1.06	0.90	0.89	-	-
2. Kotronoulas et al. 2 factor	58.13 (10)	0.12	1.06	0.90	0.88	2 vs. 1	0.32 (<i>p</i> = .57)
3. Magee et al. 2 factor	32.47 (11)	0.08	0.76	0.96	0.95	3 vs. 1	22.49 (<i>p</i> < .001)
4. Cole et al. 3 factor	22.79 (9)	0.07	0.63	0.97	0.96	4 vs. 2 4 vs. 3	33.48 (<i>p</i> < .001) 8.90 (<i>p</i> = .01)

CFI, comparative fit index; RMSEA, root mean square error of approximation; TLI, Tucker-Lewis Index; WRMR, weighted root mean residual. All absolute measures of model fit (χ^2) were statistically significant (*p* < .05).

Table 3

Bivariate Correlations Between Assessment Instruments

	1	2	3	4	5	6	7	8
1. PSQI Sleep Efficiency	-							
2. PSQI Perceived Sleep Quality	.44*	-						
3. PSQI Daily Disturbances	.25*	.39*	-					
4. PSQI Global Score	.79*	.84*	.60*	-				
5. CAPS	.36*	.40*	.37*	.48*	-			
6. BDI	.22*	.34*	.47*	.41*	.50*	-		
7. STAXI Trait Anger	.03	.01	.15	.05	.22*	.30*	-	
8. PILL	.20*	.34*	.53*	.43*	.45*	.48*	.23*	-

Values represent Pearson's product-moment correlation coefficients. BDI, Beck Depression Inventory; CAPS, Clinician Administered PTSD Scale; PILL, Pennebaker Inventory of Limbic Languidness; PSQI, Pittsburgh Sleep Quality Index; STAXI, State-Trait Anger Expression Index.

* $p < .05$ with Bonferroni correction for multiple comparisons.