Sleep Health: Journal of the National Sleep Foundation 9 (2023) 925–932

Contents lists available at ScienceDirect

Sleep Health: Journal of the National Sleep Foundation

journal homepage: www.sleephealthjournal.org

Multidimensionality of the PROMIS sleep disturbance 8b short form in working adult populations

Rebecca M. Brossoit, PhD [a,*], Hannah P. Stark, MA [a], Tori L. Crain, PhD [b], Todd E. Bodner, PhD [b], Leslie B. Hammer, PhD [c], Cynthia D. Mohr, PhD [b], Steven A. Shea, PhD [c]

[a] Louisiana State University, Department of Psychology, Baton Rouge, Louisiana, USA
[b] Portland State University, Department of Psychology, Portland, Oregon, USA
[c] Oregon Health & Science University, Oregon Institute of Occupational Health Sciences, Portland, Oregon, USA

ARTICLE INFO

Article history:
Received 23 November 2022
Received in revised form 25 July 2023
Accepted 4 August 2023

Keywords:
Sleep disturbance
Insomnia symptoms
Dissatisfaction with sleep
Psychometrics

ABSTRACT

Objectives: The Patient-Reported Outcomes Measurement Information System sleep disturbance measures were developed using item response theory assumptions of unidimensionality and local independence. Given that sleep health is multidimensional, we evaluate the factor structure of the Patient-Reported Outcomes Measurement Information System sleep disturbance 8b short form to examine whether it reflects a unidimensional or multidimensional construct.
Methods: Six full-time working adult samples were collected from civilian and military populations. Exploratory and confirmatory factor analyses were conducted. Single-factor and two-factor models were performed to evaluate the dimensionality of sleep disturbance using the 8b short form. Sleep duration and subjective health were examined as correlates of the sleep disturbance dimensions.
Results: Across six working adult samples, single-factor models consistently demonstrated poor fit, whereas the two-factor models, with insomnia symptoms (ie, trouble sleeping) and dissatisfaction with sleep (ie, subjective quality of sleep) dimensions demonstrated sufficient fit that was significantly better than the single-factor models. Across each sample, dissatisfaction with sleep was more strongly correlated with sleep duration and subjective health than insomnia symptoms, providing additional evidence for distinguishability between the two sleep disturbance factors.
Conclusions: In working adult populations, the Patient-Reported Outcomes Measurement Information System sleep disturbance 8b short form is best modeled as two distinguishable factors capturing insomnia symptoms and dissatisfaction with sleep, rather than as a unidimensional sleep disturbance construct.

The Patient-Reported Outcomes Measurement Information System (PROMIS)

The National Institutes of Health (NIH) developed the Patient-Reported Outcomes Measurement Information System (PROMIS) to collect high-quality self-reported data for assessing symptoms relevant to physical (eg, pain), mental (eg, emotional distress), and social (eg, social role participation) well-being, with the intention of effectively evaluating health-related treatments and interventions.[1] The PROMIS measures have been used extensively in past clinical and research work; there are approximately 70 health domains captured across all PROMIS measures, and the measures have been translated into over 40 languages.[1] Thus, this rigorous NIH initiative has had a significant impact on medical and public health research and practice, work that has consequently led to improvements in individuals' health, well-being, and quality of life.

Development of PROMIS sleep measures

Recognizing sleep quality as a critical facet of health and well-being, one of the PROMIS initiatives was to create item banks and scales to accurately measure self-reported sleep. As such, the PROMIS sleep-wake project conducted rigorous studies to achieve this goal, relying on systematic literature reviews, advice from subject matter experts, and extensive psychometric testing (including both classical test theory and IRT techniques). Using these approaches, Buysse et al developed the original PROMIS sleep item banks, which were intended to capture different aspects of sleep, and demonstrated initial validity evidence for two unidimensional

* Corresponding author: Rebecca M. Brossoit, PhD, Louisiana State University, Department of Psychology, Audubon Hall, Baton Rouge, LA 70803, USA.
E-mail address: rbrossoit@lsu.edu (R.M. Brossoit).

https://doi.org/10.1016/j.sleh.2023.08.001
2352-7218/

Table 1
Comparison of PROMIS sleep disturbance measures

		Short form 8a items	Short form 8b items
Form 4a items	Form 6a items		
X	X	1. My sleep quality was…	1. My sleep was restless
X	X	2. My sleep was refreshing	2. I was satisfied with my sleep
X	X	3. **I had a problem with my sleep**	3. My sleep was refreshing
X	X	4. I had difficulty falling asleep	4. I had difficulty falling asleep
	X	5. My sleep was restless	5. **I had trouble staying asleep**
	X	6. **I tried hard to get to sleep**	6. **I had trouble sleeping**
		7. **I worried about not being able to fall asleep**	7. **I got enough sleep**
		8. I was satisfied with my sleep	8. My sleep quality was...

Note. Bold items indicate different items between the 8a and 8b sleep disturbance short forms. Positively worded items that reflect favorable sleep (eg, my sleep was refreshing) are reverse-coded.

sleep constructs – sleep disturbance (ie, trouble sleeping, poor sleep quality) and sleep-related impairment (ie, daytime fatigue, cognitive and behavioral issues).[2]

In this paper, we focus on the PROMIS *sleep disturbance* measure, particularly the 8b short form.[3] The PROMIS measures of sleep disturbance have been used to understand symptoms associated with diseases (eg, cancer, multiple sclerosis) and injuries[4,5] and to evaluate intervention effects on sleep.[6–10] Therefore, it is important to continue gathering validity evidence for these measures. Since Buysse et al's initial work,[2] several different versions of the PROMIS sleep disturbance measure have been created (drawing from different combinations of items) and used in research.

From Buysse et al's item bank,[2] two 8-item short forms were developed and validated: short form 8a[11] and short form 8b.[3] Nested within short form 8a are short forms 6a and 4a (see Table 1). Some researchers have created additional customized sleep disturbance measures that combine items from the established forms.[12] Moreover, the 8a short form items are included in larger PROMIS profiles that assess multiple health domains (ie, physical function, anxiety, depression, fatigue, sleep disturbance, pain interference, and ability to participate in social roles and activities); form 8a is included in the PROMIS-57 Profile, form 6a is included in the PROMIS-43 Profile, and form 4a is included in PROMIS-29 Profile. Additionally, there are child and parent proxy PROMIS measures to assess sleep disturbance. Although there are numerous versions of the PROMIS sleep disturbance measure, we focus specifically on Yu et al's sleep disturbance 8b short form[2] for adult use in the present study, given that it is widely used, yet very little is known about its dimensionality.

Assumptions underlying PROMIS sleep disturbance measures

The PROMIS sleep disturbance scales were developed using an item-response theory (IRT) approach, as used across the development of all the PROMIS measures. One of the primary assumptions of IRT is unidimensionality, meaning that all items in a measure reflect a single underlying latent dimension rather than multiple dimensions.[13] A related assumption of the IRT approach is local independence, or that each item should capture a unique aspect of the underlying construct.[13] Consequently, although different aspects of sleep (ie, sleep disturbance, sleep-related impairment) were identified in the initial PROMIS validation studies, dimensionality *within* these sleep constructs was overlooked, and sleep disturbance has been deemed a unidimensional construct.[2,3,11] Yet, researchers using the different PROMIS sleep disturbance scales have found mixed results regarding dimensionality, with some finding support for unidimensionality,[14,15] others *not* finding evidence for unidimensionality,[12,16–18] and most not reporting factor analyses. Of note is that many different versions of the sleep disturbance scales were used throughout this prior work, so inconsistencies regarding dimensionality may be due to the various measurement approaches used.

Although the PROMIS sleep disturbance 8a measure appears to be the most popular short form, we identified that 724 articles have cited Yu et al as of March 2023 (using the Google Scholar search engine), which highlights that the 8b short form has also been extensively used in past research. Yet, a systematic literature search[a] yielded only two published articles that examined the dimensionality of the PROMIS sleep disturbance 8b measure: Jensen et al and Yang et al.[12,19] Jensen et al's work with cancer patients did not find evidence that sleep disturbance was a unidimensional construct and instead proposed another custom PROMIS sleep disturbance measure – 6b.[12] Yang et al's study with chronic fatigue patients only found evidence for unidimensionality when a bi-factor model was used.[19] Confirmatory factor analyses (CFA) and bi-factor models are both factor analysis approaches. The primary difference between a two-factor CFA (which we use in the present study), and a bi-factor model is that a bi-factor model is a latent variable modeling approach that includes a higher-order latent construct ("general factor") with lower-order factors, whereas a two-factor CFA provides information about the hypothesized factor structure of a construct but does not model a higher-order general factor. As such, the bi-factor model approach may be impractical given the large sample sizes required for latent variable modeling. Taken together, very little work has explored the dimensionality of the PROMIS sleep disturbance 8b measure (with no work to our knowledge in nonclinical working adult samples), presumably due to the unidimensionality assumption, which we assess in the present study.

The present study

Given the early conceptualization and evidence of the PROMIS sleep disturbance measures as being unidimensional, most researchers have assumed unidimensionality. We recognize that the PROMIS sleep disturbance measures were not designed to capture the multidimensional construct of sleep health, though we challenge this assumption by exploring whether there are additional distinguishable dimensions within the PROMIS sleep disturbance 8b measure. As such, our study makes three primary contributions.

First, we align measurement approaches with theory. Despite the many advantages of the IRT approach (including the ability to construct customized scales using different combinations of items), the assumption of unidimensionality is misaligned with sleep experts' theoretical understanding of sleep as a multifaceted construct.

[a] We conducted a systematic literature review to investigate the extent of factor analytic work conducted specifically on Yu et al's sleep disturbance 8b short form that was explored in the present study. In addition to Google Scholar, the electronic databases PsycINFO, PsycArticles, Psychology and Behavioral Sciences Collection, CINAHL, CINAHL Plus with full text, Health Source, MEDLINE, and Web of Science were searched using the following search terms: PROMIS Sleep Disturbance, factor*, factor analysis, psychometric, measurement, measurement properties, dimension*, and valid*.

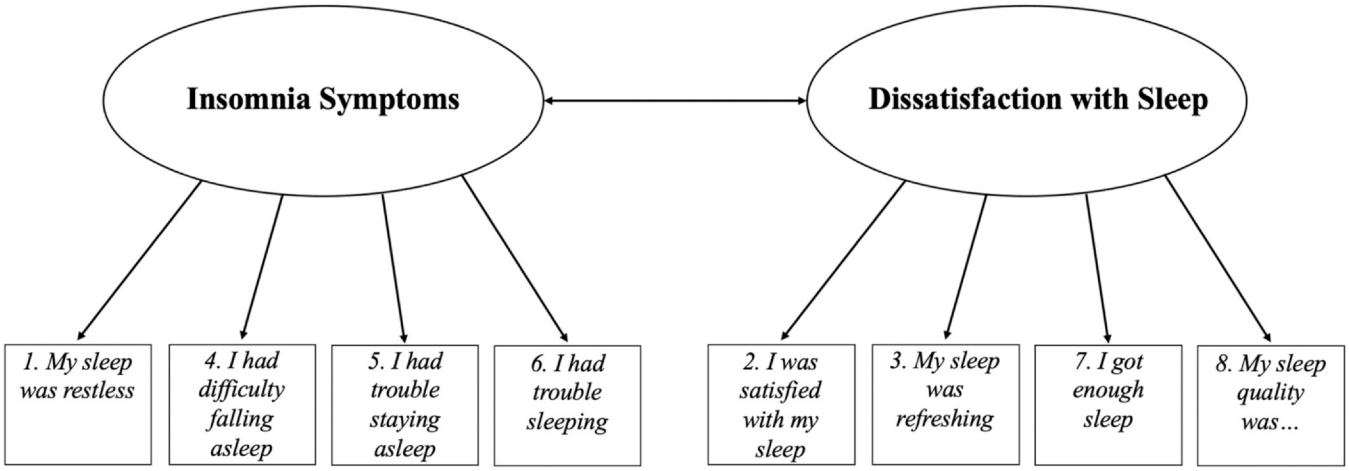

Fig. 1. The two-dimensional factor structure of the PROMIS sleep disturbance 8b short form. Response options for all items are on five-point Likert scales. Response options for items 1-4 range from "Not at All" to "Very Much" response options for items 5-7 range from "Never" to "Always" response options for item 8 range from "Very Poor" to "Very Good"

Illustratively, Buysse proposed a model of sleep health with five sleep dimensions (ie, duration, timing, satisfaction, efficiency, and alertness) and called for researchers to refine and validate these dimensions.[20] Buysse suggested that capturing these different dimensions with specificity is important, as prior research has found varying relationships between sleep and health outcomes, depending on the dimension in question.[20] The PROMIS sleep disturbance measures are meant to capture a single construct, rather than multidimensional facets of sleep disturbance. Yet, items within the PROMIS 8b scale appear to represent elements of both sleep satisfaction and sleep efficiency. Sleep satisfaction is the overall subjective assessment of sleep quality as being good or poor (eg, "I was satisfied with my sleep"), which is the core criterion of insomnia disorder,[21] whereas sleep efficiency is the extent to which someone is able to fall asleep with ease and maintain sleep over the course of the night (eg, "I had trouble staying asleep"), which reflects specific symptoms of insomnia.[21] In the current study, we explore whether these two theoretical dimensions are in fact represented empirically, discuss the implications of this multidimensionality for sleep research, and provide guidance for future use of the PROMIS 8b scale.

Fig. 1 depicts our hypothesized two-factor model structure. Of note is both insomnia symptoms and dissatisfaction with sleep are conceptually related to insomnia disorder as defined in the DSM-5.[21] Similarly, Ohayon's review on the epidemiology of insomnia demonstrates the relevance of, but also the conceptual distinction between, insomnia symptoms and dissatisfaction with sleep as features of insomnia.[22] Therefore, in addition to drawing from Buysse's sleep health conceptualization, we use the label of *insomnia symptoms* to capture the specific indicators in the diagnostic criteria for insomnia disorder (ie, trouble falling and staying asleep[21]). Dissatisfaction with the quality or duration of sleep is central to the definition of insomnia disorder[21] and we use the label *dissatisfaction with sleep* to capture the subjective assessment of how one slept.

Our second contribution is the evaluation of possible contradictions to the IRT assumption of unidimensionality. Because the PROMIS sleep disturbance measure was developed under the IRT assumption that all items in an item bank measure the same underlying latent construct, it is argued that any number or combination of items from an item bank may be used to measure the target domain without methodological concern. Although this flexibility is typically considered advantageous,[3,11] there may be drawbacks to this approach if the assumptions underlying it (eg, unidimensionality) are refuted. Our exploration can serve as a model toward understanding how other measures that were validated using IRT principles may have unrealized multidimensional factor structures that limit our understanding of the overall construct of interest. Although we recognize the utility in measurement development using IRT principles, continued investigation into dimensionality using multiple approaches may be advantageous for a more holistic understanding of health measures, especially when there is a theoretical rationale for multidimensionality.

Third, previous research using PROMIS sleep disturbance measures has largely been conducted with specific clinical samples, whereas we take a broader perspective and explore sleep disturbance within six working adult samples, considering the interconnections between work and sleep. Sleep quality is critical to employee health, well-being, safety, and job performance, and the work environment (eg, workload, work hours, schedules, social support) can influence employee sleep, as well.[23-27] The bidirectional associations between work and sleep make workplaces particularly advantageous settings to implement interventions targeting sleep health. Accordingly, it is important to accurately measure self-reported sleep disturbances in working adult populations. Overall, this study has implications for research and practice in occupational sleep medicine and public health, as well as disciplines that draw from these areas relevant to working adults' sleep (eg, industrial-organizational psychology, occupational health psychology, management[28,29]).

Methods

Participants

We used archival data from six samples of working adults in both civilian and military (ie, National Guard; NG) samples, all of whom provided informed consent. Institutional review board approval was obtained for each of the data collections. Participants worked full-time (M = 38.6 hours worked per week across samples) in primarily regular daytime schedules (70% across samples). We relied on data from MTurk workers (N = 564), working college students (N = 239), Army NG Service Members (N = 306), Army NG supervisors (N = 111), Air NG Service Members (N = 398), and Air NG supervisors (N = 104). The MTurk workers and working college students represented various occupations, whereas the military samples included individuals working for the National Guard. The MTurk workers were from geographically dispersed areas throughout the United States (U.S.), the working students were in the Western region of the U.S., and the NG samples were in the Pacific Northwest region of the U.S. MTurk workers were primarily white women in their late-30s. Working college students were predominantly white women in their early-

20s. Across the four NG samples, participants were predominantly white men in their mid-30s to early-40s.

Measures

Participants completed the PROMIS sleep disturbance 8b short form[3] (Cronbach's α range: 0.89-0.93), and T-scores were computed using the PROMIS HealthMeasures system, as recommended in the PROMIS sleep disturbance scoring manual. Sleep duration and subjective health were also measured. Sleep duration was computed using two items from the Pittsburgh Sleep Quality Index[30] that capture the typical time participants went to bed and the typical time they woke up. Subjective health was assessed with a single item across samples. In the MTurk worker and working student samples, subjective health was measured with the item: "Would you say your physical health is...?" with response options ranging from 1 (poor) to 5 (excellent). Similarly, in the four NG samples, subjective health was measured with the item: "In general, would you say your health is…" with response options ranging from 1 (poor) to 5 (excellent).

Analytic strategy

First, CFAs were performed on the eight-item sleep disturbance 8b short form using Mplus Version 8. Mplus employs the full information maximum likelihood approach to make model estimations for missing data, though there was very little item-level missingness on the PROMIS sleep disturbance measure (ie, less than 2% of responses were missing across samples). Model fit indices were evaluated using Hu and Bentler's and Yu's recommendations.[31,32] Model fit indices, factor loadings, residual covariances, and factor correlations were assessed.[33] Next, to further explore the distinguishability of the PROMIS sleep disturbance 8b factors, correlations between the two sleep disturbance factors – insomnia symptoms and dissatisfaction with sleep – and measures of sleep duration and subjective health were examined. Finally, as an additional assessment of dimensionality, supplemental exploratory factor analyses (EFAs) with oblique rotation were performed across the six samples.

Results

Confirmatory factor analyses

Across the six samples, the single-factor CFAs consistently demonstrated poor model fit (see Table 2). After exploring high residual covariances (indicative of violations to the unidimensionality assumption), alongside the eight sleep disturbance items, Buysse's model of sleep health, and the DSM-5 description of insomnia disorder, two-factor CFAs were performed – one factor with items specific to insomnia symptoms and one factor with items specific to dissatisfaction with sleep (see Fig. 1). Across each sample, the two-factor CFAs demonstrated significantly improved model fit indices compared to the single-factor models (see Table 2)[b,c] providing

Table 2
Comparison of single-factor and two-factor confirmatory factor analyses for the PROMIS sleep disturbance 8b short form

	Single-factor model						Two-factor model							
Sample	χ^2	CFI	TLI	RMSEA	SRMR	Standardized factor loadings	χ^2	CFI	TLI	RMSEA	SRMR	Standardized factor loadings	Change in χ^2	*p*-value	Factor correlations
MTurk workers	621.67	0.84	0.78	0.23	0.07	0.64-0.89	148.03	0.97	0.95	0.11	0.04	0.72-0.94	473.64	< .001	0.77
Working students	238.73	0.84	0.77	0.21	0.07	0.64-0.88	81.79	0.95	0.93	0.12	0.04	0.69-0.91	156.94	< .001	0.76
Army NG service members	335.39	0.80	0.72	0.23	0.08	0.52-0.84	81.37	0.96	0.94	0.10	0.04	0.58-0.91	254.01	< .001	0.70
Army NG supervisors	170.83	0.77	0.68	0.26	0.09	0.58-0.88	55.63	0.94	0.92	0.13	0.06	0.62-0.92	115.20	< .001	0.67
Air NG service members	422.35	0.78	0.69	0.23	0.09	0.60-0.83	148.62	0.93	0.90	0.13	0.07	0.66-0.90	273.73	< .001	0.65
Air NG supervisors	112.21	0.80	0.73	0.21	0.08	0.43-0.82	52.04	0.93	0.90	0.13	0.05	0.47-0.90	60.16	< .001	0.73

χ^2, chi-square; CFI, comparative fit index; NG, National Guard; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker-Lewis index. Sample sizes: MTurk workers (N = 564), Working students (N = 239), Army NG service members (N = 306), Army NG supervisors (N = 111), Air NG service members (N = 398), Air NG supervisors (N = 104).

[b] In the four NG samples, there were two follow-up data collections conducted 4-months and 9-months following the baseline data that we report in this paper. The pattern of CFA results is consistent across the three time points for these samples but are omitted for simplicity and because measurement invariance was not central to our research questions for this paper.

[c] Given that Item 1 ("my sleep was restless") could theoretically fit with the dissatisfaction factor, we ran two-factor CFAs with item 1 moved from the insomnia symptoms factor to the dissatisfaction with sleep factor. Across each of the six samples, the model fit indices are better in the originally proposed two-factor model (with item 1 modeled in the insomnia symptoms factor), compared to when item 1 is modeled in the dissatisfaction with sleep measure.

Table 3
Insomnia symptoms, dissatisfaction with sleep, sleep duration, and subjective health descriptives

Sample	Insomnia symptoms		Dissatisfaction with sleep		Sleep duration		Subjective health	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
MTurk workers	50.86	9.29	50.58	9.05	7.69	1.18	3.38	0.92
Working students	52.25	9.07	50.39	8.58	8.18	1.34	3.13	0.97
Army NG service members	54.07	8.09	55.16	7.68	7.41	1.04	3.25	0.84
Army NG supervisors	53.23	7.85	54.60	7.52	7.60	0.95	3.42	0.82
Air NG service members	51.47	8.11	53.04	7.24	7.19	0.98	3.51	0.81
Air NG supervisors	51.89	6.89	51.82	7.04	7.14	0.96	3.73	0.77

M, mean; NG, National Guard; *SD*, standard deviation. Insomnia symptoms and dissatisfaction with sleep values are in T-scores (which have a population mean of 50 and standard deviation of 10). Sleep duration values are in hours. Subjective health is on a 1-5 scale, in which higher scores reflect greater perceived health.

Table 4
Comparative correlations among the PROMIS sleep disturbance 8b factors, sleep duration, and subjective health

	Sleep duration					
	MTurk workers	Working students	Army NG service members	Air NG service members	Army NG supervisors	Air NG supervisors
Insomnia symptoms	-0.07	-0.15*	-0.04	0.00	0.12	0.02
Dissatisfaction with sleep	-0.20**	-0.33**	-0.19**	-0.13*	-0.04	-0.07
	Subjective health					
	MTurk workers	Working students	Army NG service members	Air NG service members	Army NG supervisors	Air NG supervisors
Insomnia symptoms	-0.36**	-0.30**	-0.20**	-0.22**	-0.36**	0.15
Dissatisfaction with sleep	-0.43**	-0.37**	-0.40**	-0.35**	-0.41**	-0.42**

NG, National Guard. * = $p < .05$, ** = $p < .001$.

evidence for multidimensionality rather than unidimensionality. Although the insomnia symptoms and dissatisfaction with sleep factors were positively correlated (average $r = 0.71$; r range: 0.65-0.77), a correlation of 0.85 or greater is a typical threshold used to indicate that constructs lack discriminant validity.[34,35] Based on the CFA results, separate factors were created for insomnia symptoms (eg, "I had difficulty falling asleep") and dissatisfaction with sleep (eg, "I was satisfied with my sleep" reverse-coded), as depicted in Fig. 1. Average insomnia symptoms T-scores ranged from 50.86 to 54.07 and average dissatisfaction with sleep T-scores ranged from 50.39 to 55.16 (see Table 3).

Correlations with sleep duration and subjective health

Among the six samples, participants reported adequate sleep durations, with averages ranging from 7.14 to 8.18 hours[27] (see Table 3).[d] Overall, participants reported good subjective health, with averages ranging from 3.13 to 3.73 (See Table 3). Notably, the dissatisfaction with sleep factor was more strongly correlated with both sleep duration and subjective health compared to insomnia symptoms (See Table 4). This pattern was consistent across each of the samples. These findings further demonstrate that insomnia symptoms and dissatisfaction with sleep are distinguishable factors.

Supplemental exploratory factor analyses

Unlike CFAs, the hypothesized number of factors are not specified in an EFA. The EFA results provided additional support for the hypothesized two-factor structure of the PROMIS sleep disturbance 8b measure. Specifically, across the six samples, the two-factor model had excellent fit, nearly all eigenvalues for the second factor were greater than one, and the factor loadings aligned with the hypothesized two-factor structure[e] (see Tables 5 and 6). Like the CFA results, the EFAs demonstrated positive correlations between the insomnia symptoms and dissatisfaction with sleep factors (average $r = 0.64$; r range: 0.55-0.73).

Discussion

In six full-time working adult samples, the PROMIS sleep disturbance 8b short form is best modeled as having two distinguishable factors: an insomnia symptoms factor and a dissatisfaction with sleep factor. Drawing from Buysse's model of sleep health and the DSM-5 definition of insomnia disorder, the *insomnia symptoms* factor captures trouble sleeping, falling asleep, staying asleep, and having restless sleep, whereas the *dissatisfaction with sleep* factor reflects the subjective assessment of one's sleep quality and satisfaction with sleep, including the extent to which enough sleep was obtained and whether sleep was refreshing (see Fig. 1). Across six samples, and in both EFAs and CFAs, we do not find support for a unidimensional sleep disturbance construct when using the PROMIS 8b measure. Therefore, our findings are counter to assumptions of unidimensionality and local independence under the IRT approach, though align with Buysse's model of sleep health; the sleep efficiency dimension (ie, the ease of falling and returning to sleep) overlaps with the insomnia symptoms factor; the satisfaction dimension (ie, subjective perception of sleep as good vs. poor) overlaps with the dissatisfaction with sleep factor.[20]

The two-factor structure of sleep disturbance we found corresponds with Kim et al's work using a Korean version of the PROMIS-29 Profile (which includes the 4a short-form items) in a clinical sample of patients with lower extremity problems.[16] More broadly, our findings align with previous research that also found PROMIS sleep disturbance to violate the unidimensionality assumption, though this past work has relied on markedly different populations from working adults (eg, Dutch children

[d] Outliers were determined as sleep duration values shorter than 4 hours and longer than 10.75 hours and were removed. Even when outliers are included, the average sleep durations fall within the healthy range. MTurk workers and working students had the largest proportion of sleep duration outliers (8% and 7%, respectively). In the four NG samples, less than 2% of the participants had sleep duration outliers.

[e] An exception was found in the Air NG Service Member sample, in which item 8 loaded similarly on both factors.

Table 5
Two-factor model results from exploratory factor analyses for the PROMIS sleep disturbance 8b short form

	Two-factor model							
Sample	χ^2	CFI	TLI	RMSEA	SRMR	Factor correlations	Factor loadings for insomnia symptoms items	Factor loadings for dissatisfaction with sleep items
MTurk workers	64.65	0.99	0.97	0.08	0.02	0.73	0.57-0.99	0.73-0.95
Working students	41.16	0.98	0.95	0.10	0.02	0.69	0.59-0.94	0.66-0.88
Army NG service members	40.19	0.98	0.96	0.08	0.02	0.63	0.56-0.93	0.64-0.92
Army NG supervisors	20.78	0.99	0.98	0.07	0.02	0.63	0.55-0.99	0.61-0.94
Air NG service members	25.87	0.99	0.99	0.05	0.02	0.55	0.70-0.91	0.46-0.94
Air NG supervisors	37.13	0.95	0.89	0.14	0.04	0.63	0.32-1.00	0.56-0.88

χ^2, chi-square; CFI, comparative fit index; NG, National Guard; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker-Lewis index. Single-factor model results are analogous to those presented in Table 2. Items 1, 4, 5, and 6 reflect insomnia symptoms items. Items 2, 3, 7, and 8 reflect dissatisfaction with sleep items. Sample sizes: MTurk workers (N = 564), Working students (N = 239), Army NG service members (N = 306), Army NG supervisors (N = 111), Air NG service members (N = 398), Air NG supervisors (N = 104).

Table 6
Eigenvalues from exploratory factor analyses for the PROMIS sleep disturbance 8b short form

	Eigenvalues							
Sample	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
MTurk workers	5.44	0.93	0.49	0.33	0.29	0.21	0.16	0.15
Working students	5.06	1.03	0.55	0.41	0.32	0.23	0.23	0.18
Army NG service members	4.80	1.10	0.64	0.45	0.33	0.30	0.21	0.17
Army NG supervisors	5.04	1.18	0.61	0.33	0.31	0.22	0.16	0.15
Air NG service members	4.62	1.27	0.57	0.45	0.34	0.32	0.23	0.21
Air NG supervisors	4.60	1.03	0.87	0.44	0.38	0.31	0.21	0.17

NG, National Guard.

and adolescents[17,18]). Other researchers have similarly noted high inter-item correlations, though some did not test two-dimensional models,[12,34] chose to employ a bi-factor model,[19] or correlated items' error terms[36] to achieve model fit. On the other hand, our results are discrepant from work that has found PROMIS sleep disturbance measures to be unidimensional in clinical populations (eg, patients with chronic hepatitis C[14]), combined clinical and community populations,[2,3] and older adult populations.[15] Overall, researchers typically do not report information about the dimensionality of the PROMIS sleep disturbance measure, and of those who do, there is considerable variability in the specific measure used, participant demographics (eg, age, clinical health status), sample sizes, and analytic approaches. Thus, the dimensionality of the PROMIS sleep disturbance measure may vary by these factors.

Researchers have relied on several different sleep disturbance short forms with various combinations of items. Despite intending to reflect a single underlying construct, the items used within and across separate PROMIS measures may tap into different aspects of sleep disturbance. Indeed, our results suggest that it may be inappropriate to model sleep disturbance as a unidimensional measure when using the 8b form with working adults. Illustratively, *across* sleep disturbance forms, the *8a* item "I tried hard to get to sleep" may capture an individual's prioritization or motivation to sleep that is not reflected in the 8b items. Additionally, the 8a item "I worried about not being able to fall asleep" assesses concern about not falling asleep, which is not mirrored in the 8b items (see Table 1). Although we cannot speak to the dimensionality of form 8a, considering that forms 8a and 8b differ by three items, they reflect slightly different aspects of sleep disturbance (eg, only the 8a form captures motivations and worries related to sleep). It cannot be assumed that the different forms are assessing sleep disturbance equivalently, so the correspondence between the PROMIS sleep disturbance 8a and 8b short forms could be explored in future research. Ultimately, using numerous sleep disturbance measures throughout the research literature can limit the content and construct validity of sleep disturbance and may confound the conclusions that are drawn from this work.

Implications

Our results have implications for research and practice. Despite the strong and positive correlations between the insomnia symptoms and dissatisfaction with sleep factors in the PROMIS sleep disturbance 8b measure, they may be differentially related to variables of interest, as demonstrated in our assessment of their correlations with sleep duration and subjective health (see Table 4). Researchers who combine these unique aspects of sleep disturbance in a single measure may fail to detect true effects in their studies and risk misrepresenting sleep disturbance as a unidimensional construct. In both research and practice, understanding whether a treatment method or intervention improves symptoms of insomnia as well as subjective experiences of sleep quality provides important information for how to refine treatment and intervention approaches. Examining sleep disturbance unidimensionally may lead to an inability to detect change when it occurs in one dimension of sleep disturbance but not the other (eg, a treatment may increase perceptions of sleep quality but not reduce symptoms of insomnia, which could otherwise be missed if the dimensions are combined). From a clinical perspective, the diagnostic criteria for insomnia disorder includes both subjective dissatisfaction with sleep and specific associated symptoms of insomnia, which are currently assessed simultaneously in the PROMIS 8b measure. Assessing these features of insomnia disorder as separate dimensions may allow clinicians to identify a patient's needs more precisely and/or better evaluate the effectiveness of a treatment method. Our results suggest that there are advantages to dividing the PROMIS 8b items into separate insomnia symptoms and dissatisfaction with sleep factors (as illustrated in Fig. 1) and recommend that researchers and practitioners consider evaluating each as a distinguishable dimension of sleep disturbance.

Limitations and future research directions

The two-factor models demonstrated significantly improved model fit relative to the single-factor models, but the model fit remained

imperfect in the two-factor CFA model. In particular, the RMSEA values did not meet recommendations for good fit. One possible explanation is that item 8 ("My sleep quality was…" modeled in the dissatisfaction with sleep factor) had high and positive residual covariance values with items in the insomnia symptoms factor, particularly item 1 ("My sleep was restless"). This pattern suggests that item 8 is more correlated than would be expected with the insomnia symptoms items. Similarly, in the EFA, item 8 loaded on to both factors in the Air NG Service Member sample. Nevertheless, we retained item 8 in the dissatisfaction with sleep factor for theoretical reasons, given that it reflects an individual's overall impression of the quality of their sleep.[20]

Although we find consistent results across six samples, some characteristics of the samples may limit the generalizability of the results. Our focus on working adults presumably generalizes more to the general population compared to specific clinical samples (eg, a sample exclusively of individuals with a diagnosed sleep disorder), with the working sample collected from MTurk likely being the most generalizable to the general population, given that participants worked in a variety of jobs and across the U.S.[37] However, the employment status of the participants in our samples may limit generalizability to unemployed populations, given the negative association between work time and sleep time.[38] Participants across the samples had characteristics that may have protected their sleep, including their relatively young age, race (ie, predominantly white), standard work hours, and daytime work schedules.[39–42] On the other hand, the National Guard participants were unique due to their role in the military, which poses the potential for combat exposure and related vulnerability to post-traumatic stress disorder[43] and corresponding sleep issues.[21] Overall, the present study was limited by a lack of alternative samples from working populations that may be particularly susceptible to disrupted sleep, such as older employees, employees from marginalized racial or ethnic groups, employees in industries affected by shiftwork (eg, healthcare) or other work arrangements that can disrupt circadian rhythms (eg, mining), and employees with diagnosed psychological disorders or sleep disorders. We encourage researchers to continue examining the dimensionality of sleep disturbance in more diverse working populations.

Finally, we focus on Yu et al's 8b short-form version of the PROMIS sleep disturbance measure, so future research could build on Jensen et al's work (in which several different existing and custom versions of the PROMIS sleep disturbance measures were explored[12]) to examine the dimensionality of sleep disturbance across different measures. For example, researchers could investigate whether the insomnia symptoms and dissatisfaction with sleep dimensions are evident in the sleep disturbance 8a version, particularly given the overlap between the two short forms (ie, five of the eight items are the same). The three unique items in the PROMIS sleep disturbance 8a measure (ie, 8a items 3, 6, and 7 in Table 1) seem to reflect additional insomnia symptoms more than dissatisfaction with sleep, which could be empirically tested in future work. Researchers could also examine multidimensional models in the PROMIS sleep-related impairment measure, as well. More broadly, it may be worthwhile to investigate the factor structure and potential multidimensionality of other PROMIS measures that were developed using IRT methods, to ensure these health-related measures are being modeled appropriately in future research and practice.

Conclusion

Drawing from Buysse's model of sleep health, the DSM-5, and factor-analytic results from six samples of full-time working adults, we argue that the commonly used PROMIS sleep disturbance 8b short form[3] is best modeled multidimensionally – with distinguishable factors capturing insomnia symptoms and dissatisfaction with sleep – rather than unidimensionally. Future research should continue exploring the dimensionality of sleep disturbance and improving the measurement and utility of this construct.

Funding

The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office. This work was supported by Office of the Assistant Secretary of Defense for Health Affairs, through the Psychological Health and Traumatic Brain Injury Research Program - Comprehensive Universal Prevention/Health Promotion Interventions Award, under Award No. W81XWH-16-1-0720. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the Department of Defense. This work was also supported by Grants #T42OH009229 and #T03OH008435, funded by the Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, and Grant #R35HL155681 from the National Institutes of Health, and Oregon Institute of Occupational Health Sciences via funds from the Division of Consumer and Business Services of the State of Oregon (ORS 656.630). The contents are solely the responsibility of the authors and do not necessarily represent the official views of Oregon Health & Science University, State of Oregon, NIH, NIOSH, CDC, or HHS.

Author contributions

Rebecca Brossoit: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Resources, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition, Project administration. Hannah P. Stark: Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. Tori L. Crain: Methodology, Investigation, Writing – review & editing, Funding acquisition. Todd E. Bodner: Supervision, Methodology, Investigation, Writing – review & editing, Funding acquisition. Leslie B. Hammer: Supervision, Investigation, Writing – review & editing, Funding acquisition. Cynthia D. Mohr: Supervision, Investigation, Writing – review & editing, Funding acquisition. Steven A. Shea: Supervision, Writing – review & editing.

Declaration of conflict of interest

Drs Brossoit, Crain, Hammer, Bodner, & Mohr report grants from the Department of Defense, during the conduct of the study. Drs Brossoit and Crain report grants from the Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, during the conduct of the study. Dr Shea and Hannah P. Stark have nothing to disclose.

References

1. National Institutes of Health. *Patient-Reported Outcomes Measurement Information System (PROMIS).* National Institutes of Health Office of Strategic Coordination – The Common Fund; 2019. Available from: ⟨https://commonfund.nih.gov/promis/index⟩.
2. Buysse DJ, Yu L, Moul DE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep.* 2010;33(6):781–792. https://doi.org/10.1093/sleep/33.6.781
3. Yu L, Buysse DJ, Germain A, et al. Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. *Behav Sleep Med.* 2012;10(1):6–24. https://doi.org/10.1080/15402002.2012.636266
4. Fogelberg DJ, Vitiello MV, Hoffman JM, Bamer AM, Amtmann D. Comparison of self-report sleep measures for individuals with multiple sclerosis and spinal cord injury. *Arch Phys Med Rehabil.* 2015;96(3):478–483. https://doi.org/10.1016/j.apmr.2014.10.001
5. Flynn KE, Shelby RA, Mitchell SA, et al. Sleep-wake functioning along the cancer continuum: focus group results from the patient-reported outcomes measurement information system (PROMIS®). *Psycho-Oncology.* 2010;19(10):1086–1093. https://doi.org/10.1002/pon.1664
6. Lillehei AS, Halcón LL, Savik K, Reis R. Effect of inhaled lavender and sleep hygiene on self-reported sleep issues: a randomized controlled trial. *J Altern Complement Med.* 2015;21(7):430–438. https://doi.org/10.1089/acm.2014.0327

7. Loewenthal J, Dyer NL, Lipsyc-Sharf M, et al. Evaluation of a yoga-based mind-body intervention for resident physicians: a randomized clinical trial. *Glob Adv Health Med.* 2021;10:1–12. https://doi.org/10.1177/21649561211001038
8. Sarfan LD, Gasperetti CE, Gumport NB, Harvey AG. Outcomes from the trans-diagnostic sleep and circadian intervention (TranS-C) for midlife and older adults with serious mental illness and sleep and circadian dysfunction. *Behav Ther.* 2022;53(4):585–599. https://doi.org/10.1016/j.beth.2022.02.001
9. Sohl SJ, Danhauer SC, Birdee GS, et al. A brief yoga intervention implemented during chemotherapy: a randomized controlled pilot study. *Complement Ther Med.* 2016;25:139–142. https://doi.org/10.1016/j.beth.2022.02.001
10. Spadola CE, Rottapel RE, Zhou ES, et al. A sleep hygiene and yoga intervention conducted in affordable housing communities: pilot study results and lessons for a future trial. *Complement Ther Clin Pract.* 2020;39:101–121. https://doi.org/10.1016/j.ctcp.2020.101121
11. Cella D, Riley W, Stone A, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol.* 2010;63(11):1179–1194. https://doi.org/10.1016/j.jclinepi.2010.04.011
12. Jensen RE, King-Kallimanis BL, Sexton E, et al. Measurement properties of PROMIS sleep disturbance short forms in a large, ethnically diverse cancer cohort. *Psychol Test Assess Model.* 2016;58(2):353. https://doi.org/10.1080/15402002.2012.636266
13. Nguyen TH, Han HR, Kim MT, Chan KS. An introduction to item response theory for patient-reported outcome measurement. *Patient Patient Cent Outcomes Res.* 2014;7(1):23–35. https://doi.org/10.1007/s40271-013-0041-0
14. Evon DM, Amador J, Stewart P, et al. Psychometric properties of the PROMIS short form measures in a US cohort of 961 patients with chronic hepatitis C prescribed direct acting antiviral therapy. *Aliment Pharmacol Ther.* 2018;47(7):1001–1011. https://doi.org/10.1111/apt.14531
15. Full KM, Malhotra A, Crist K, Moran K, Kerr J. Assessing psychometric properties of the PROMIS sleep disturbance scale in older adults in independent-living and continuing care retirement communities. *Sleep Health.* 2019;5(1):18–22. https://doi.org/10.1016/j.sleh.2018.09.003
16. Kim Y, Kang D, Kang E, et al. Psychometric validation of the Korean version of PROMIS 29 Profile V2. 1 among patients with lower extremity problems. *BMC Sports Sci Med Rehabil.* 2021;13(1):1–9. https://doi.org/10.1186/s13102-021-00374-1
17. Peersmann SH, Luijten MA, Haverman L, Terwee CB, Grootenhuis MA, van Litsenburg RR. Psychometric properties and CAT performance of the PROMIS pediatric sleep disturbance, sleep-related impairment, and fatigue item banks in Dutch children and adolescents. *Psychol Assess.* 2022;34(9):860–869. https://doi.org/10.1037/pas0001150
18. van Kooten JA, van Litsenburg RR, Yoder WR, Kaspers GJ, Terwee CB. Validation of the PROMIS sleep disturbance and sleep-related impairment item banks in Dutch adolescents. *Qual Life Res.* 2018;27(7):1911–1920. https://doi.org/10.1007/s11136-018-1856-x
19. Yang M, Keller S, Lin JMS. Assessing sleep and pain among adults with myalgic encephalomyelitis/chronic fatigue syndrome: psychometric evaluation of the PROMIS® sleep and pain short forms. *Qual Life Res.* 2022;31:1–17. https://doi.org/10.1007/s11136-022-03199-8
20. Buysse DJ. Sleep health: can we define it? Does it matter? *Sleep.* 2014;37(1):9–17. https://doi.org/10.5665/sleep.3298
21. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. 2013. ⟨https://doi.org/10.1176/appi.books.9780890425596⟩.
22. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. *Sleep Med Rev.* 2002;6(2):97–111. https://doi.org/10.1053/smrv.2002.018
23. Barnes CM, Watson NF. Why healthy sleep is good for business. *Sleep Med Rev.* 2019;47:112–118. https://doi.org/10.1016/j.smrv.2019.07.005
24. Buxton OM, Shea SA. Bidirectional relationships between sleep and work. *Sleep Health.* 2020;6(3):259–261. https://doi.org/10.1016/j.sleh.2020.04.012
25. Jackson CL, Lee S, Crain TL, Buxton OM. Bidirectional relationships between work and sleep. In: Duncan DT, Kawachi I, Redline S, eds. *The Social Epidemiology of Sleep.* Oxford, United Kingdom: Oxford University Press; 2019:351–371.
26. Litwiller B, Snyder LA, Taylor WD, Steele LM. The relationship between sleep and work: a meta-analysis. *J Appl Psychol.* 2017;102(4):682–699. https://doi.org/10.1037/apl0000169
27. Watson NF, Badr MS, Belenky G, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. *Sleep.* 2015;38:843–844. https://doi.org/10.5665/sleep.4716
28. Barnes CM. Working in our sleep: sleep and self-regulation in organizations. *Organ Psychol Rev.* 2012;2(3):234–257. https://doi.org/10.1177/2041386612450181
29. Crain TL, Brossoit RM, Fisher GG. Work, nonwork, and sleep (WNS): a review and conceptual framework. *J Bus Psychol.* 2018;33(6):675–697. https://doi.org/10.1007/s10869-017-9521-x
30. Buysse DJ, Reynolds III CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 1989;28(2):193–213. https://doi.org/10.1016/0165-1781(89)90047-4
31. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis; 1999:328-351.
32. Yu CY. *Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes.* Los Angeles: University of California; 2002.
33. Raykov T, Marcoulides GA. *Introduction to Psychometric Theory.* Routledge; 2011.
34. Kenny DA. Multiple latent variable models: confirmatory factor analysis; 2012. http://davidakenny.net/cm/mfactor.htm.
35. Shaffer JA, DeGeest D, Li A. Tackling the problem of construct proliferation: a guide to assessing the discriminant validity of conceptually related constructs. *Organ Res Methods.* 2016;19(1):80–110. https://doi.org/10.1177/1094428115598239
36. Green K, Eddy A, Flowers J, Christopher M. A confirmatory factor analysis of the PROMIS sleep disturbance among law enforcement officers. *J Police Crim Psychol.* 2022;37(1):141–145. https://doi.org/10.1007/s11896-021-09471-w
37. Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality data? In: Kazdin AE, ed. *Methodological Issues and Strategies in Clinical Research.* American Psychological Association; 2016:133–139. https://doi.org/10.1037/14805-009
38. Basner M, Fomberstein KM, Razavi FM, et al. American time use survey: sleep time and its relationship to waking activities. *Sleep.* 2007;30(9):1085–1095.
39. Neikrug AB, Ancoli-Israel S. Sleep disorders in the older adult–a mini-review. *Gerontology.* 2010;56(2):181–189.
40. Grandner MA, Williams NJ, Knutson KL, Roberts D, Jean-Louis G. Sleep disparity, race/ethnicity, and socioeconomic position. *Sleep Med.* 2016;18:7–18.
41. Åkerstedt T. Shift work and disturbed sleep/wakefulness. *Occup Med.* 2003;53(2):89–94.
42. Åkerstedt T, Fredlund P, Gillberg M, Jansson B. Work load and work hours in relation to disturbed sleep and fatigue in a large representative sample. *J Psychosomatic Res.* 2002;53(1):585–588.
43. Cameron KL, Sturdivant RX, Baker SP. Trends in the incidence of physician-diagnosed posttraumatic stress disorder among active-duty US military personnel between 1999 and 2008. *Mil Med Res.* 2019;6(1):1–13. https://doi.org/10.1186/s40779-019-0198-5