Received: 24 April 2024 | Revised: 4 July 2024 | Accepted: 14 September 2024

DOI: 10.1002/mpr.70004

ORIGINAL ARTICLE

WILEY

The Holland Sleep Disorders Questionnaire: Factorial structure and measurement invariance in a psychiatric sample relative to the general population

Teus Mijnster[1,2] | Klaas J. Wardenaar[3,4] | Gretha J. Boersma[5] | Maaike M. van Veen[1,5] | Daniëlle Cath[3,4,6] | Gerard A. Kerkhof[7,8] | Marike Lancel[1,2,5]

[1]Centre of Expertise on Sleep and Psychiatry, GGZ Drenthe Mental Health Institute, Assen, The Netherlands

[2]Department of Clinical Psychology & Experimental Psychopathology, University of Groningen, Groningen, The Netherlands

[3]Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[4]Department of Psychiatry, University Medical Center Groningen, Rob Giel Onderzoekscentrum, Groningen, The Netherlands

[5]Forensic Psychiatric Hospital, GGZ Drenthe Mental Health Institute, Assen, The Netherlands

[6]Department of Specialist Trainings, GGZ Drenthe Mental Health Institute, Assen, The Netherlands

[7]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

[8]Sleep Disorders Center, Haaglanden Medical Center (HMC), The Hague, The Netherlands

Correspondence

Teus Mijnster, GGZ Drenthe Mental Health Institute, Dennenweg 9, Assen, LA 9404, The Netherlands.
Email: teus.mijnster@ggzdrenthe.nl

Funding information

ZonMw, Grant/Award Number: 626310014

Abstract

Objectives: Although common, sleep disorders often remain undiagnosed in psychiatric patients. A screening instrument, like the Holland Sleep Disorders Questionnaire (HSDQ) could improve this. Previous work indicated a 6-factor structure for the HSDQ, but this hasn't been investigated in psychiatric patients.

Methods: HSDQ data was collected in a psychiatric-outpatient sample (n = 1082) and general-population sample (n = 2089). Internal reliability of the HSDQ was investigated and Confirmatory Factor Analyses (CFA) were used to compare 1-, 6-, and second-order 6-factor models in both samples. Next, multigroup-CFA was used to investigate measurement invariance.

Results: Except for one subscale, internal reliability was acceptable in both samples. The 6-factor structure model fitted best in both samples and investigation of measurement invariance showed evidence for equality of the overall factor structure (configural invariance). Addition of equality constraints on factor loadings (metric invariance) and item thresholds (scalar invariance) showed good fit for all fit statistics, except for one. Exploratory analyses identified three items for metric and three different items for scalar invariance explaining this non-invariance.

Conclusion: The HSDQ has a 6-factor structure in psychiatric patients, which is comparable to the general population. However, due to the observed non-invariance, users should be cautious with comparing HSDQ scores between psychiatric and general populations.

KEYWORDS

construct validity, Holland Sleep Disorders Questionnaire, measurement invariance, mental health patients, sleep disorders

1 | INTRODUCTION

The international classification of sleep disorders (ICSD-3) identified six categories of sleep disorders (American Academy of Sleep Medicine, 2014). The distinguished categories are: insomnia characterized by chronic problems falling and/or maintaining sleep, parasomnia referring to undesirable behaviors/experiences during the sleep, circadian rhythm sleep-wake disorders (CRSWD) typified by misalignment of the timing of sleep, hypersomnia referring to excessive daytime sleepiness, sleep related movement disorders (SRMD) characterized by unwanted (urges to) limb movements during sleep, and sleep breathing disorders (SBD) with breathing problems during sleep (See Table 1). Within varying mental health disorder groups, the prevalence of nearly all types of sleep disorders is high (Hombali et al., 2019). For example, in people with a depressive disorder insomnia and the most frequent subtype of SBD, obstructive sleep apnea, are rather common, each affecting about 40% of patients (e.g. Stewart er al., 2006; Stubbs et al., 2016). Recurrent, distressing nightmares, a parasomnia, occur in so many persons with PTSD (50%–70%; Gieselman et al., 2019) that it is considered a diagnostic symptom. Attention Deficit Hyperactivity Syndrome (ADHD) is associated with the CRSWD delayed sleep phase syndrome (DSPS) in approximately 30% of cases (e.g. Spera et al., 2020). Importantly, sleep disorders have been related to higher severity of psychopathology. For instance, Spera et al. (2020) found that adults with ADHD and comorbid DSPS reported greater functional impairments than persons with ADHD without DSPS. Moreover, treatment of sleep disorders not only improves sleep, but may also facilitate recovery of co-occurring mental disorders. For instance, cognitive behavioral therapy for insomnia in persons with depression has been shown to improve both sleep and depression outcomes (Blom et al., 2017; Manber et al., 2011). Also, in persons with PTSD, various sleep treatments have been found to alleviate both sleep disturbances and daytime PTSD symptoms (Maher et al., 2021). Thus, improving sleep can be considered a crucial target for interventions across mental disorders. Unfortunately, despite their importance, sleep disorders have generally been under-diagnosed and consequentially under-treated in mental healthcare practice (Freeman et al., 2020).

A valid screening instrument for sleep disorders could aid mental healthcare professionals with the timely detection of comorbid sleep disorders in persons with a mental illness. A promising screening instrument in this regard is the Holland Sleep Disorders Questionnaire (HSDQ), a 32-item questionnaire that screens for sleep disturbances of the six main categories of sleep disorders according to the ICSD-3 (American Academy of Sleep Medicine, 2005; Kerkhof et al., 2013). The HSDQ is widely used in the Netherlands for both clinical and research purposes, is available in several languages, and has been well validated in a Dutch and an Iranian population (Kerkhof et al., 2013; Khazaie et al., 2022). An earlier study in a sample of the general population of the Netherlands demonstrated that the HSDQ accurately distinguishes between persons with and without a sleep disorder in 88% of cases (86% for control subjects and 90.5% for patients referred to a sleep center) (Kerkhof et al., 2013). Furthermore, it has good validity with the clinical sleep disorder diagnoses: for persons with a clinical sleep disorder (established with clinical interviews and additional diagnostics) the HSDQ adequately differentiated between the six sleep disorder categories in 84.5% of all cases (Kerkhof et al., 2013).

The HSDQ has been integrated into the Routine Outcome Monitoring of outpatient departments within GGZ Drenthe Mental Health Institute. Therefore, it is prudent to assess whether the factorial structure of the HSDQ and interpretation of its items are comparable with the general population, where the validity of the HSDQ was previously confirmed (Kerkhof et al., 2013). While various

TABLE 1 Brief description of the (groups of) sleep disorders.

Sleep disorder group	Item #	Brief description
Insomnia disorder	1–7–10–12–13–14–15–21	Problems with sleep initiation and/or maintenance at least 3 nights a week for a period of 3 months at minimum, associated with daytime distress.
Parasomnia	4–16–20–22–24–31	Undesirable behaviors or experiences that occur during non-rapid eye movement (NREM-) sleep, REM-sleep, or while transitioning between sleep and wake.
Circadian rhythm sleep disorders (CRSD)	5–10–13–26–27–30	Symptoms of insomnia and daytime impairment due to a mismatch between one's endogenous sleep-wake rhythm and the external light-dark cycle.
Hypersomnia	23–25–28–29–32	Excessive daytime sleepiness not explained by other sleep disorders.
Sleep related movement disorders (SRMD)[a]	2–6–8–9–11	Restless legs syndrome (RLS): Unpleasant sensations in the legs paired with a hard to resist urge to move the legs. Symptoms are worse at rest and during the evening/night and cause sleep disturbances and/or impairment in daytime functioning. Periodic limb movements during sleep (PLMD): Involves sleep disruptive repetitive movements of the limbs.
Sleep breathing disorders (SBD)	3–17–18–19	Frequent instances of partial and/or complete upper airway obstruction during sleep, paired with clinically relevant symptoms and/or diseases.

Note: For diagnostic criteria, see ICSD-2 (American Academy of Sleep Medicine, 2005). HSDQ scale terms are used for the sleep disorders. Parasomnia refers to the broad spectrum of rapid eye movement (REM) - and non-REM (NREM)-parasomnia in the ICSD-2, circadian rhythm sleep disorders equates to ICSD-2 circadian rhythm sleep-wake disorders, hypersomnia equates to hypersomnolences in the ICSD-2.

[a]Referred to as RLS/LMS in Kerkhof et al., 2013. SBD mostly refers to obstructive sleep apnea syndrome in the ICSD-2.

widely used and well-validated other sleep screening instruments do exist (Arrol et al., 2011; Bailes et al., 2008; Douglass et al., 1994; Koffel & Watson, 2010; Roth et al., 2002; Spoormaker et al., 2005), none of these are applied and validated in a population of individuals with mental health disorders as of yet. Therefore, the HSDQ could be the first self-reported instrument for assessing all groups of sleep disorders validated in a mental health disorder sample. Furthermore, the HSDQ is relatively short, which is an advantage given that many mental disorders, like schizophrenia, depression and anxiety disorders, are associated with attention deficits and other cognitive impairments (e.g., Castaneda et al., 2008; Kahn & Keefe, 2013; Roca et al., 2015; Rock et al., 2014).

The psychometric properties of the HSDQ as a multidimensional questionnaire were investigated by Kerkhof et al. (2013) using Principal Component Analysis in a population referred to a sleep center. They found six components, representative of the six categories of sleep disorders. As of date, it is unclear if this 6-factor structure is also optimal in the psychiatric population. Given how the HSDQ was designed to screen for six well-differentiated sleep disorder domains, it is likely that the general latent structure of the HSDQ will be largely the same in the psychiatric and general populations. However, the exact factor structure (e.g., magnitude of factor loadings, means and/or covariances) could be slightly different and/or some of the HSDQ's items might function differently (i.e., there could be measurement non-invariance). For instance, the Mental Health Continuum Short Form (MHC-SF) was administered in a general population sample and divided this sample into two groups classified as high distress (high score on the depression anxiety scale) and non-distressed (low score on the depression anxiety scale). The MHC-SF was found to be non-invariant (factor loadings) across these groups, suggesting differences in item interpretation based on the level of psychological distress (Iasiello et al., 2022). Conceivably, some items might be interpreted differently by people with and without a mental disorder. For instance, certain psychiatric disorders, such as depression, are characterized by a negative cognitive bias (Disner et al., 2017) and this could lead to perceptions of their sleep symptoms as more severe. If specific items of the HSDQ are indeed evaluated differently by psychiatric patients, caution in comparing scores between general population and psychiatric samples is warranted.

This study will be the first step in validating the HSDQ for use in psychiatric populations. It will do so by investigating internal consistency and evaluating the factorial structure of the HSDQ using confirmatory factor analysis (CFA) in a sample treated in an ambulatory psychiatric setting (individuals diagnosed with a mental health disorder such as affective, trauma-related and personality disorders) and a previously reported general Dutch population sample (Kerkhof, 2017). Using these two samples allows us to directly compare the internal consistency and factorial structure of the HSDQ between them. In the initial CFA, the original 6-factor solution found by Kerkhof et al. (2013) will be fitted and compared to both a 1-factor model and a 6-factor model with an overarching second-order factor. Next, measurement invariance of the best fitting model across the two samples will be investigated by fitting several multi-group CFA models with incrementally stricter equality constraints on the CFA model parameters. Finding measurement invariance would imply that the HSDQ assesses sleep disorders in psychiatric patients in a manner comparable to the general population, making direct comparisons of subscale scores possible.

2 | METHODS

2.1 | Participants and procedures

2.1.1 | General population

General population data were obtained from an online research panel database (ISO26362-certified) containing information on over 80,000 Dutch citizens. From this database, a final sample representative of the general Dutch population was obtained by propensity weighting, which corrected for demographic, socio-economic and cultural factors, Internet (non)-use and non-response (see Kerkhof, 2017). This resulted in a final sample of 2089 subjects (51.4%, $n = 1073$ female) with an age range of 18–70 years (Median = 50; IQR: 60 - 38).

2.1.2 | Psychiatric population

Psychiatric population data were collected as part of the "MOnitoring psychoPHARmacology" (MOPHAR) naturalistic cohort study conducted at outpatient departments of GGZ Drenthe Mental Health Institute (see Simoons et al., 2019 for MOPHAR design). Patients receiving mental health care at this facility filled out the HSDQ as part of their Routine Outcome Monitoring and provided consent for use of their data for scientific research. MOPHAR was approved by the local medical ethical committee. A total of 1090 (642 female; 58.9%) patients fulfilled the inclusion criteria (≥18 years and signed informed consent). Of this sample, 1082 completed the HSDQ and were included in the final sample (59.0%, $n = 638$ female) with an age range of 18–84 years (Median = 43.0; IQR: 54-31). Frequencies and percentages of the mental disorder diagnoses in the sample are given in Table 1.

2.2 | Measures

The HSDQ is a 32-item self-report questionnaire aimed at screening for potential sleep disorders. Each item describes one or more symptoms of a (group of) sleep disorder. Respondents indicate to what degree the symptom description applied to them during the past 3 months on a Likert scale from 1 ("Never") to 5 ("Almost always"). Descriptions of the individual items are listed in Supporting Information S1: Appendix 1. The HSDQ was designed to assess six groups of sleep disorders with separate scales (see Table 2). Additionally, the HSDQ is assumed to contain an overarching General

TABLE 2 Most prevalent mental disorders.

Psychiatric diagnosis	*n* (%)
Major depressive disorder	238 (22.0%)
Bipolar disorder	201 (18.6%)
Anxiety disorder	140 (12.9%)
Posttraumatic stress disorder	118 (10.9%)
Attention deficit hyperactivity disorder	84 (7.8%)
Autism spectrum disorder	49 (4.5%)
Substance use disorder	30 (2.8%)
Eating disorder	13 (1.2%)
Personality disorder	176 (16.3%)
Other disorders	69 (6.3%)
No diagnosis yet	269 (24.9%)

Note: Table is not cumulative since patients can have multiple psychiatric diagnoses.

Sleep Disturbance index (GSD) that includes all items and functions as an overall measure of disturbed sleep (Kerkhof, 2017). Internal consistency in a sleep clinic sample was satisfactory ($\alpha = 0.90$ for the GSD and $\alpha = 0.73$–0.81 for the six disorder-specific scales; Kerkhof et al., 2013).

2.2.1 | Missing data

The general-population sample had no missing data on the HSDQ by design. Only 8 out of 1090 (0.73%) participants in the psychiatric sample had any missing data on the HSDQ. Therefore, complete-case analyses were deemed acceptable to deal with missing data in this sample.

2.2.2 | Statistical analyses

Differences in age and sex between the two samples were investigated with a Mann-Whitney-U test and chi-square test respectively. Internal consistency of HSDQ-subscales was analyzed with Cronbach's alpha. Values of $\alpha > 0.90$ were considered excellent, between 0.90 and 0.80 good, between 0.80 and 0.70 acceptable, between 0.70 and 0.60 questionable, between 0.60 and 0.50 poor, and $\alpha < 0.50$ was considered unacceptable (Gliem & Gliem, 2003). Furthermore, McDonald's Omega (ω_t) was added as a measure of internal consistency following the procedure outlined in recent literature (Malkewitz et al., 2023).

To assess the factor structure of the HSDQ in the general and psychiatric samples, CFA models were fitted with Rosseel's (2012) R Lavaan package (R Core Team, 2019). Three candidate factor-models were fitted and compared. These analyses were done separately in both samples. In the 1-factor model, all HSDQ items were set to load on a single GSD factor. In the 6-factor model, each item was set to load on one of the originally identified factors: insomnia, parasomnia, Circadian rhythm sleep disorders (CRSD), hypersomnia, SRMD, and SBD. Congruent to the 6-factor model from Kerkhof et al. (2013), items 10 and 13 were set to cross-load on both the insomnia and CRSD factor in this model. In the second-order model, the six factors of the previous model were all set to load on an overarching, second-order GSD factor. Since the HSDQ items are scored on a five-point ordinal Likert scale, a diagonally weighted least squares estimator was used for model estimation. Here, four additional parameters (thresholds between adjacent response categories) are estimated for each HSDQ item in all models instead of one intercept per item, as would be the case for factor models with continuous items. In the 1- and 6-factor models, the loading of the first item on each factor was fixed to 1 to set the scale and identify the model. For the second order model, the variance of the overarching GSD factor was set to 1 for identification purposes. All reported factor loadings were standardized. Models were evaluated based on indices of model fit, where a Root Mean Square of Approximation (RMSEA) of 0.06 or lower, a Standardized Root Mean Square Residual (SRMR) of 0.08 or lower, and a Comparative Fit Index (CFI) of 0.95 or higher were taken to indicate good model fit (Hu & Bentler, 1998). To assess the influence and necessity of the cross-loading items in the 6-factor model, we refitted the 6-factor model with these items loading on only one factor.

The best-fitting model was established in both samples and the results showed the same model to have the best fit in both samples. Therefore, measurement invariance was assessed in a stepwise manner by estimating multi-group CFA (MG-CFA) models with incremental equality constraints on the model parameters across groups. Next, the fit of each model was compared to its less constrained counterpart. A drop in model-fit not exceeding its cut-off value for the more constrained model compared to the less constrained model was taken to be indicative of measurement invariance. First, *configural invariance* was tested to assess equality of factorial structure by fitting the best fitting model from the separate CFAs in a combined dataset containing both the psychiatric and general population samples. In this case, no equality constraints across samples were imposed. Subsequently, we assessed if factor loadings were invariant across samples (*metric invariance*) by constraining the factor loadings to be equal across samples. The fit of this constrained model was compared to the fit of the configural model. Next, equality of item thresholds (*scalar invariance*) was tested by fitting a model in which both the item-thresholds and factor loadings were constrained to be equal. The fit of the scalar model was then compared to that of the metric model. In addition to each model's absolute fit, assessed by the RMSEA, CFI and SRMR, we examined differences in fit of the metric versus the configural model and the scalar model versus the metric model. For these analyses, the χ^2-difference test was not used to compare the invariance models, since the χ^2-difference test is heavily influenced by sample size, with $n > 400$ usually leading to significant results (Kenny, 2015). Instead, a criterion of a maximum change in CFI was used (ΔCFI $\leq$0.01), which is less influenced by sample size (Cheung & Rensvold, 2002; Putnick & Bornstein, 2016). Furthermore, following Chen's (2007)

recommendation we looked at the differences in RMSEA between the models, with a ΔRMSEA ≥0.015 indicating a substantially worse fit, which is also less influenced by sample size (Fan & Sivo, 2007). When non-invariance was encountered in a particular step, we used modification indices to assess which items showed a large degree of non-invariance. If modification indices indicated that releasing equality constraints on items would have a substantial impact on the expected model fit (in terms of χ^2), equality constraints were released until measures of model fit met their cut-off values.

3 | RESULTS

3.1 | Demographics

The group of psychiatric patients was significantly younger (Mann Whitney U, $W = 13.90$, $p < 0.001$; Median = 50.0 years (IQR: 60 - 38) in the general versus Median = 43.0 years (IQR: 54 - 31) in the psychiatric sample) and had a higher percentage of women ($\chi^2(df = 1) = 16.27$, $p < 0.001$, with women making up 51.4% ($n = 1073$) in the general population versus 59.0% ($n = 638$) in the psychiatric sample). The distributions of HSDQ item-scores are reported for both samples in Supporting Information S1: Appendix 1. Data on prevalences of sleep disorders and their comorbidity from both samples have been published and discussed previously. For the general population, 10.3% ($n = 215$) were estimated to have one sleep disorder, while 12.2% ($n = 255$) were estimated to have 2 or more sleep disorders (Kerkhof, 2017). For the psychiatric population, these figures were 21.8% ($n = 236$) and 24.40% ($n = 264$) respectively (Mijnster et al., 2024).

3.2 | Internal consistency

Internal consistencies for the HSDQ-subscales are given in Table 3. Internal consistency of the GSD, the insomnia and CRSD scales was good to excellent in both samples. For parasomnia, internal consistency was good in the general population sample and acceptable in the psychiatric sample. The hypersomnia scale had acceptable internal consistency in both samples. Finally, internal consistency of SBD was questionable in the general population and poor in the psychiatric sample. The pattern of McDonald's ω_t values was similar to Cronbach's α in that SBD seemed to have the lowest internal consistency and Insomnia was the specific sleep disorder with the highest internal consistency.

3.3 | Confirmatory factor analyses

Table 4 gives the model fit statistics for all CFA models. All standardized factor loadings, factor variances/covariances and item residual variances are reported in Figures 1 and 2. Threshold values of the CFAs are given in Supporting Information S1: Appendix 2. A one-factor GSD model had a poor, unacceptable fit in both samples, with only the CFI indicating adequate fit in the general population sample. The original 6-factor model fitted adequately according to all fit measures in both samples. For this model, the SRMR and CFI showed a slightly better fit for the general-population sample. The RMSEA indicated a slightly better fit for the psychiatric population. In both samples, the overall fit for the second-order model was worse than that for the 6-factor model. This was most notable for the RMSEA, which did not meet the 0.06 cut-off criterion in either sample. Taken together, these analyses show that a 6-factor model fits best to the HSDQ data in both samples. Analyses without cross-loading items are included in Supporting Information S1: Appendix 3.

3.3.1 | Factor loadings for general population sample

Figure 1 depicts standardized factor loadings for the general population sample. The best fitting model (general 6-factor model) is discussed below. All items had significant loadings. For insomnia, the items that loaded substantially lower than the other items were those that cross-loaded with the CRSD scale (items 10 and 13). For the cross-loading items, factor loadings showed an inverse pattern for CSRD compared to insomnia: the items that loaded relatively low on insomnia had high loadings on CSRD. Item 27 of the CRSD scale and item 18 of the SBD scale had substantially lower standardized factor loadings than other items of their respective scales.

3.3.2 | Factor loadings for psychiatric population sample

Figure 2 depicts standardized factor loadings for first order and second order models for the psychiatric sample. The best fitting model (psychiatry 6-factors) is discussed below. All items had significant loadings. For insomnia, the items that cross-loaded with the CRSD scale (items 10 and 13) loaded substantially lower than the other items. Factor loadings for the cross-loading items in the first order models show relatively high loadings on CRSD, while having relatively low loadings on insomnia. On parasomnia, item 31 had a substantially lower factor loading than the other items. Item 27 of the CRSD scale, item 28 of the hypersomnia scale, and item 18 of the SBD scale had low loadings compared to the other items of their respective scales.

3.4 | Measurement invariance MG-CFA

Measurement invariance was investigated for the best fitting model, the 6-factor model. The results are shown in Table 5. The configural invariance model, in which the model configuration was the same, but all parameters were allowed to vary freely between the general population and psychiatric samples, showed adequate fit. Metric invariance was evaluated by comparing the configural model to a model where all factor loadings were constrained to equality

TABLE 3 Cronbach's α and McDonald's ω_t for HSDQ scales in general and psychiatric populations with 95% confidence intervals.

Scale	Cronbach's α general population	Cronbach's α psychiatric population	McDonald's ω_t general population	McDonald's ω_t psychiatric population
GSD	0.93 (0.93–0.94)	0.91 (0.90–0.92)	0.94 (0.93–0.94)	0.92 (0.91–0.92)
Insomnia	0.92 (0.92–0.93)	0.92 (0.91–0.93)	0.93 (0.92–0.93)	0.92 (0.92–0.93)
Parasomnia	0.83 (0.82–0.85)	0.76 (0.74–0.79)	0.84 (0.82–0.86)	0.83 (0.81–0.85)
CRSD	0.84 (0.83–0.85)	0.83 (0.81–0.85)	0.84 (0.84–0.86)	0.85 (0.83–0.86)
Hypersomnia	0.78 (0.76–0.80)	0.74 (0.71–0.77)	0.78 (0.75–0.79)	0.74 (0.70–0.77)
SRMD	0.82 (0.80–0.83)	0.78 (0.75–0.80)	0.82 (0.80–0.83)	0.77 (0.75–0.81)
SBD	0.65 (0.62–0.67)	0.58 (0.53–0.63)	0.61 (0.58–0.65)	0.56 (0.51–0.61)

Note: The 95% confidence intervals are reported between parentheses.

Abbreviations: CRSD, Circadian rhythm sleep disorder; GSD, General sleep disturbance index; SBD, Sleep breathing disorder; SRMD, Sleep related movement disorders.

TABLE 4 Model fit for all CFA fitted models in the general and psychiatric samples.

Group	Model	Estimated parameters	χ^2	*df*	*p*	CFI	RMSEA	SRMR
General population	1 factor	160	14,048.898	464	<0.0001	0.952[a]	0.118	0.111
	6 factors	177	3609.008	447	<0.0001	0.989[a]	0.058[a]	0.061[a]
	6 factors with GSD	168	5692.351	456	<0.0001	0.981[a]	0.074	0.077[a]
Psychiatric population	1 factor	160	8032.204	464	<0.0001	0.930	0.123	0.116
	6 factors	177	1882.448	447	<0.0001	0.987[a]	0.055[a]	0.062[a]
	6 factors with GSD	168	2806.219	456	<0.0001	0.978[a]	0.069	0.077[a]

Abbreviations: CFI, Comparative Fit Index; *df*, degrees of freedom; GSD, General Sleep Disturbance index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual.

[a]Denotes that a fit statistic met requirements established in the statistical analyses section (CFI ≥0.95, RMSEA ≤0.06, SRMR ≤0.08).

between the samples. The latter model did not have a substantially worse fit than the configural model based on ΔRMSEA and ΔCFI, and its CFI and SRMR values indicated adequate fit in absolute terms. However, the RMSEA of the model with constrained factor loadings was larger than 0.06, indicating suboptimal fit. Next, a scalar invariance model was fitted where the thresholds and loadings were both constrained to equality. Here, ΔCFI and ΔRMSEA were again very small and both the CFI and SRMS indicated adequate fit. However, the RMSEA again exceeded the cut-off score for adequate fit. Taken together, measures of relative fit (ΔCFI and ΔRMSEA) did not reflect substantial measurement invariance between the models. However, the RMSEA of both the scalar and metric invariance models indicated suboptimal fit, which suggests that the HSDQ's factor loadings and thresholds are unlikely to be entirely comparable between the psychiatric and general population samples.

3.5 | Partial invariance based on modification indices

Additional models were fitted to assess partial invariance (see Table 6). Modification indices identified three items with an expected χ^2-difference above 100 when the equality constraint for factor loadings would be released. These were item 28 (long sleep), item 31 (sleepwalking) and item 26 (abnormal sleep rhythm). Releasing the equality constraint for item 28 yielded an acceptable RMSEA and led to an improved SRMR (see Modified-metric model in Table 6). When equality constraints for the factor loadings of all three of these items were lifted, the RMSEA and SRMR improved further (Modified-metric2 in the table). The next step was to see if scalar invariance held for the Modified-metric2 model. As indicated by Modified-scalar in the table, the RMSEA once again exceeded its cutoff value of 0.060. When inspecting modification indices for this model, three specific thresholds would have a large influence on the χ^2, namely threshold 1 for item 27 (poor daytime sleep), threshold 1 for item 24 (confusional arousal) and threshold 1 for item 22 (involuntary self-harm). After releasing the equality constraints on these thresholds, partial invariance could be established (Modified-scalar2 in Table 6).

4 | DISCUSSION

This study built upon Kerkhof and colleagues' exploratory assessment of the HSDQ's multidimensional structure in a general population sample (Kerkhof et al., 2013). We aimed to investigate the internal consistency and factor structure of the HSDQ in a

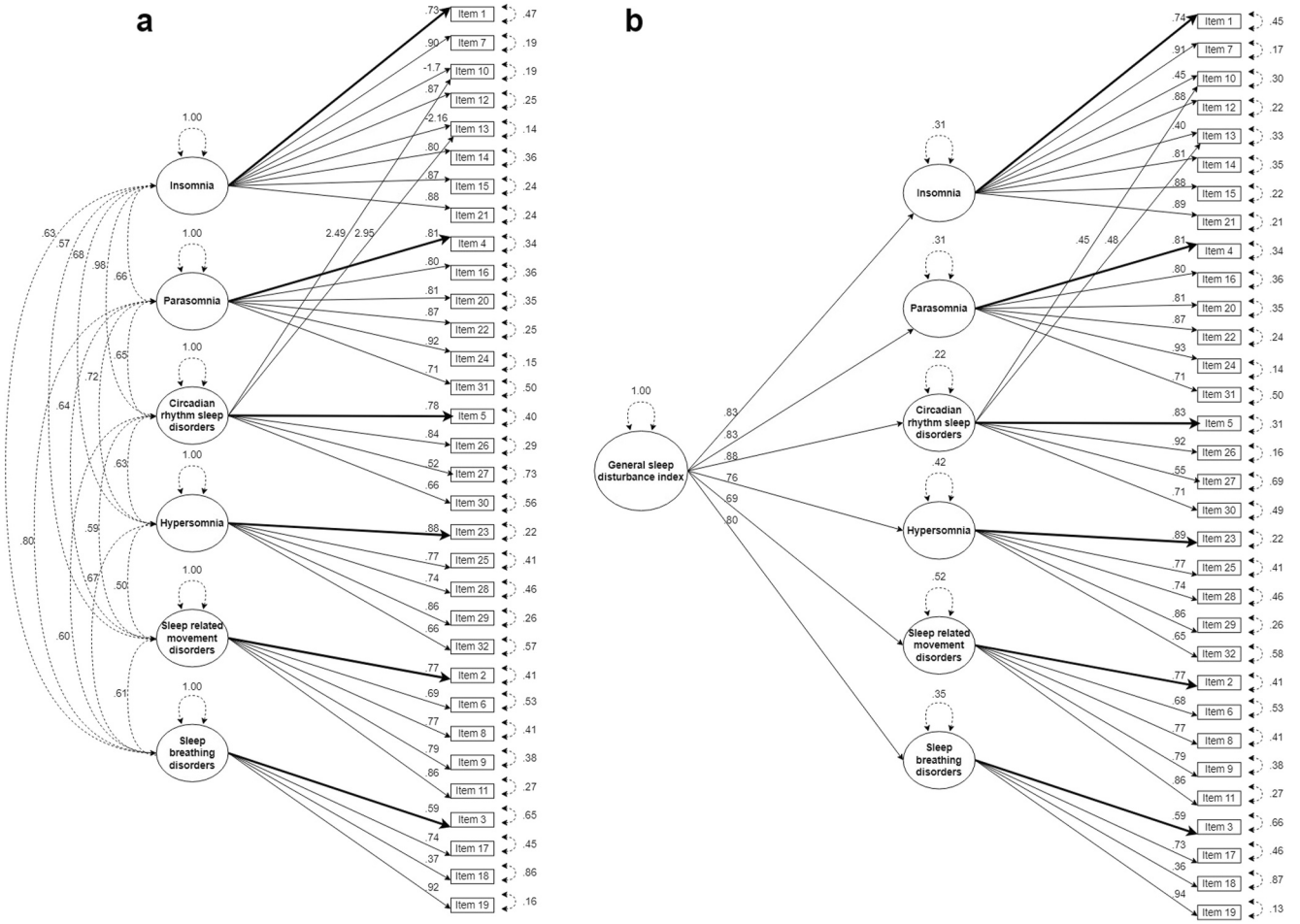

FIGURE 1 Results of the Confirmatory Factor Analyses for the general-population sample. Standardized factor loadings, item residual variances, factor variances and covariances are depicted. Bold lines indicate factor loadings set to one for identification purposes. (a) Reflects the six-factor model, (b) the second-order model with overarching general sleep disturbance index.

psychiatric sample and to determine if factor loadings and thresholds of HSDQ items are invariant between this psychiatric sample and a general population sample. Overall, internal consistency of the HSDQ scales was comparable and, apart from the SBD-subscale, ranged from acceptable to excellent. CFA showed that the original 6-factor model (insomnia, parasomnia, CRSD, SRMD, hypersomnia, and SBD) (Kerkhof et al., 2013), was the best-fitting factor structure in both samples. Importantly, this 6-factor model had a better fit than a 1-factor model and a second-order factor model with an overarching General Sleep Disturbance Index (GSD) factor. Furthermore, measurement invariance of factor loadings and thresholds across both samples was indicated by relative fit measures. However, model fit dropped slightly below the pre-defined cut-off for an acceptable fit on one of the absolute fit indices (i.e., RMSEA >0.06), indicating that some relevant differences in factor loadings and thresholds may exist between the two samples.

A notable finding was that the internal consistency for the SBD-subscale was insufficient in both samples. This could imply that the SBD-subscale does not form a stable factor and should not be used as such. However, these finding may also be due to a measurement issue, for example, it might be related to the absolute and relative (compared to other HSDQ scales) limited number of items within this subscale, since Cronbach's alpha increases when more items are added to a scale when inter-item correlation remains the same (Spiliotopoulou, 2009). Following the Spearman-Brown prophecy formula (Warrens, 2015), when increasing the items of the SBD scale from four to eight Cronbach's alpha would increase to 0.73 for the psychiatric sample and to 0.79 for the general population sample. Additionally, the HSDQ does not take crucial daytime SBD symptoms and relevant comorbidities into account, such as excessive daytime sleepiness, cardiovascular and/or affective disorders (McNicholas, 2008). Addition of items regarding these symptoms might elevate the internal consistency of the SBD-subscale. Further research is needed to understand the issues with and potentially improve the SBD-subscale.

Although the HSDQ is assumed to have a second-order overarching GSD-scale, addition of this factor to the six subscales worsened model fit in both samples. This indicates that psychometrically, the GSD is not supported as an overall representation of sleep symptoms encompassing all other latent factors. In theory

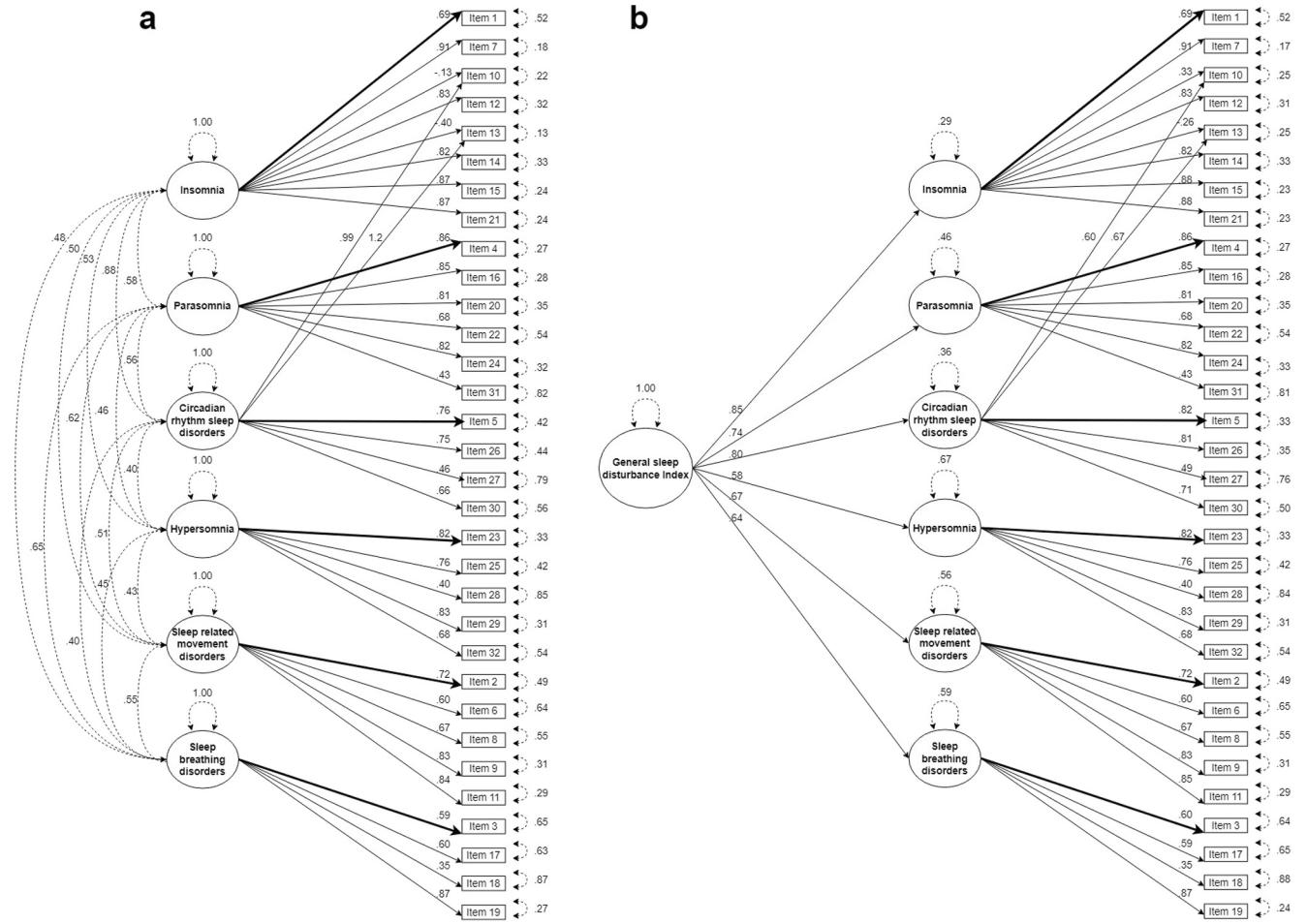

FIGURE 2 Results of the Confirmatory Factor Analyses for the psychiatric population. Standardized factor loadings, item residual variances, factor variances and covariances are depicted. Bold lines indicate factor loadings set to one for identification purposes. (a) Reflects the six-factor model, (b) the second-order model with overarching general sleep disturbance index.

TABLE 5 Measurement invariance of the HSDQ across psychiatric and general populations.

Invariance model	Estimated parameters	χ^2	*df*	*p*	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA
Configural	354	5491.456	894	<0.0001	0.988	0.057	0.061		
Metric	326	6382.303	922	<0.0001	0.986	0.061[a]	0.068	−0.002	0.004
Scalar	236	7359.983	1012	<0.0001	0.984	0.063[a]	0.065	−0.002	0.002

Abbreviations: CFI, Comparative Fit Index; *df*, degrees of freedom; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual.

[a]Denotes that a fit statistic did not meet the criteria established in the statistical analyses section (CFI ≥0.95, RMSEA ≤0.06, SRMR ≤0.08, ΔCFI ≤0.01, ΔRMSEA ≥0.015).

one might expect the GSD to be a transdiagnostic construct overarching specific sleep disorders, due to some symptom overlap between sleep disorders. However, the GSD may be inappropriate as a second-order latent factor because certain sleep disorders, such as hypersomnia and parasomnia, have more disorder specific symptoms. Therefore, one factor overarching all six subscales might be less appropriate. Like Kerkhof et al. (2013), we allowed item 10 and item 13 to load on both insomnia and CRSD in our CFA analyses. In both samples, extreme factor loadings (either direction) were found for these items on the insomnia and CRSD scales. This may be explained by that correlations between the latent factors insomnia and CRSD were estimated alongside the cross-loading items. When fitting models, in which these items only load on either insomnia or CRSD, high factor loadings were found (see Supporting Information S1: Appendix 3). Still, overall model fit was better when items 10 and 13 were allowed to cross-load. This may be expected, since CRSD and insomnia partially overlap in symptomatology.

TABLE 6 Measurement invariance of the HSDQ across the psychiatric and general population with partial invariance models.

Invariance model	Estimated parameters	χ^2	*df*	*p*	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA
Configural	354	5491.456	894	<0.0001	0.988	0.057	0.061		
Metric	326	6382.303	922	<0.0001	0.986	0.061[b]	0.068	−0.002	0.004
Modified-metric[a]	327	6123.601	921	<0.0001	0.987	0.060	0.066	−0.001	0.003
Modified-metric2	329	5892.937	919	<0.0001	0.987	0.058	0.064	−0.001	0.001
Scalar	236	7359.983	1012	<0.0001	0.984	0.063[b]	0.065	−0.002	0.002
Modified-scalar[a]	239	7133.101	1009	<0.0001	0.984	0.062[b]	0.063	−0.003	0.004
Modified-scalar2[a]	242	6787.029	1006	<0.0001	0.985	0.060	0.063	−0.001	0.002

Abbreviations: CFI, Comparative Fit Index; *df*, degrees of freedom; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual.

[a]For the Modified metric model, the equality constraint on the loading for item 28 was released. For the Modified metric 2 model, equality constraints on the loadings for items 26, 28 and 31 were released. The Modified scalar model was run with the same equality constraints on factor loadings lifted as in the Modified metric 2 model. The Modified scalar model 2 build upon the Modified scalar model and had additional released equality constraints for the thresholds of item 27 (threshold 1), item 24 (threshold 1) and item 22 (threshold 1).

[b] Denotes that a fit statistic did not meet the criteria established in the statistical analyses section (CFI ≥0.95, RMSEA ≤0.06, SRMR ≤0.08, ΔCFI ≤0.01, ΔRMSEA ≥0.015).

Finding that the original 6-factor model had the best fit in both samples allowed us to assess measurement invariance of the HSDQ. Results showed that the factor loadings and thresholds were not completely invariant across the two samples. For the MG-CFA with metric invariance, the RMSEA only slightly exceeded its cutoff for acceptable fit. This was also the case for the scalar invariance model. Still, other fit indices indicated acceptable fit, and the difference in CFI between the configural versus metric and scalar versus metric models were below their respective cut-off values. Together, these results indicate that some differences in item-functioning may be present between the samples.

In order to understand these differences, we inspected measurement invariance with additional models where equality constraints were released based on modification indices to establish partial invariance. Modification indices indicated three equality constraints on factor loadings that led to a relevant improvement in fit when released, with the largest expected change for item 28 ("I usually sleep more than 10 h a night, have difficulty waking up in the morning, and nap during the day"). Releasing the equality constraint of equal factor loadings for item 28 led to an acceptable fit that further improved when releasing these constraints for items 26 ("I sleep poorly because I don't manage to fall asleep at a normal hour and wake up at a normal hour in the morning") and 31 ("I suffer from sleepwalking") as well. This indicates that these items do not have the same strength of association to their latent factors in the two samples. There could be several reasons for this. Item 28 could be considered a triple barreled question since it aims to identify three different symptoms in one item. Item 26 is a double-barreled question that incorporates poor sleep because of an inability to fall asleep at a normal time and/or an inability to wake up at a normal time. Respondents might focus on different aspects of the double/triple barreled items, and factors such as educational level might influence their response, in addition to true differences in the measured construct (Menold, 2020; Taple et al., 2022). Potentially, deviating interpretations of (parts) of these items or differing response strategies lead to differences in how strong these items are related to their factors between the two populations, making these items non-invariant. Additionally, item 26 uses the word "normal" in its wording to describe the sleep-wake rhythm, of which the interpretation might differ based on factors such as being employed, work schedule and chronotype. These differences in interpretation could also lead the item to show different associations to their latent factors between the samples. Finally, a parasomnia item (item 31) had relatively low factor loadings in the psychiatric sample CFAs. Given that psychiatric patients more often (and, often exclusively) scored high on the parasomnia item concerning nightmares, it could be that sleepwalking was less strongly related to overall parasomnia in this sample than in the general population sample. This is supported by the fact that nightmares are more common in those with psychiatric disorders, especially in posttraumatic stress disorders, where nightmares are considered a key symptom (Sheaves et al., 2022).

To meet the RMSEA criterion for scalar invariance, the constraints on threshold 1 of item 22 ("I have injured myself during sleep and had no recollection of the event afterward."), item 24 ("Quite often I partially wake up and find myself thrashing my arms. I usually don't recall this later on.") and item 27 ("When I have to stay awake during the night, my daytime sleep is poor.") had to be released. Since items 22 and 24 are parasomnia items, it could be that a similar explanation for the factor loading of 31 mentioned above is applicable. Explanations for differences in threshold 1 for item 27 between the two samples still elude us.

Regarding invariance of the HSDQ, at least partial metric and scalar invariance could be established. We would therefore argue that the HSDQ could be applied in psychiatric populations. However, we recommend that those who use the HSDQ in psychiatric populations consider potential differences in interpretation of items, especially those implicated by the modification indices in this study (factor

loadings for items 26, 28 and 31 and thresholds for item 27 (threshold 1), item 24 (threshold 1) and item 22 (threshold 1)). This means that direct comparisons of HSDQ scale scores between psychiatric patients and subjects from the general population warrant some caution.

A few limitations and avenues for future research are of note. Firstly, convergent validity and test-retest reliability were not investigated. Future research should also include instruments with similar purposes, such as the Sleep-50 (Spoormaker et al., 2005), to assess convergent validity and include multiple repeated measurements to enable evaluation of test-retest reliability. Secondly, the results from the psychiatric sample were based on data of three outpatient clinics in one mental healthcare institution and lacked patients with schizophrenia spectrum disorders. This may limit the generalizability of the results to other psychiatric populations. Thirdly, the psychiatric sample differed significantly from the general population sample in the distribution of age and sex. These and other unobserved sample differences could also have contributed to the observed differences in item functioning between the samples.

Future research should investigate the accuracy of the HSDQ in screening psychiatric patients for specific sleep disorders. Here, Receiver Operating Characteristics Curves should be inspected to identify the optimal cut-off score per subscale, as these may differ between psychiatric populations. Also, the HSDQ screens for ICSD-2 sleep disorders and while criteria for sleep disorders that can be assessed with questionnaires seem mostly unchanged between the ICSD-2 and ICSD-3 (Sateia, 2014), a revised HSDQ could consider the ICSD-3. Furthermore, additional assessments of measurement invariance within different specific psychiatric diagnosis groups would be advised. The present study investigated the factorial structure and measurement invariance of the HSDQ in psychiatric patients. Our findings that the validity of the HSDQ's original 6-factor structure as found earlier (Kerkhof et al., 2013) in the general population extends to the psychiatric population, and that most of the HSDQ items are invariant between general population and psychiatric samples, draws us closer toward validating the HSDQ for assessing sleep disorders in psychiatric patients.

AUTHOR CONTRIBUTIONS

Teus Mijnster: Conceptualization; formal analysis; investigation; methodology; project administration; software; validation; visualization; writing – review & editing; writing - original draft. **Klaas J. Wardenaar**: Conceptualization; formal analysis; methodology; software; supervision; validation; writing – review & editing; writing – original draft; visualization. **Gretha J. Boersma**: Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; writing – review & editing; writing – original draft; validation. **Maaike M. van Veen**: Writing - review & editing; conceptualization; funding acquisition. **Daniëlle Cath**: Writing – review & editing; data curation; investigation; resources. **Gerard A. Kerkhof**: Writing – review & editing; data curation; conceptualization; investigation; methodology; resources. **Marike Lancel**: Writing – review & editing; conceptualization; formal analysis; investigation; funding acquisition; methodology; writing – original draft; validation; project administration; resources; supervision

ACKNOWLEDGMENT

This work is supported by ZonMW, grant number 626310014. The MOPHAR database is funded by the Espria innovation fund.

CONFLICT OF INTEREST STATEMENT

None of the authors have conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

Materials and data for this study are not available, analysis code is available upon reasonable request.

ETHICS STATEMENT

MOPHAR was approved by the local medical ethical committee (RTOP Leeuwarden, study #298).

PATIENT CONSENT

Informed consent was obtained from all individual participants involved in the study.

ORCID

Teus Mijnster https://orcid.org/0000-0002-8936-2256

REFERENCES

American Academy of Sleep Medicine. (2005). The international classification of sleep disorders: Diagnostic and coding manual (2nd ed.).

American Academy of Sleep Medicine. (2014). The international classification of sleep disorders (3rd ed.). Westchester, Illinois, USA.

Arroll, B., Fernando III, A., Falloon, K., Warman, G., & Goodyear-Smith, F. (2011). Development, validation (diagnostic accuracy) and audit of the Auckland sleep questionnaire: A new tool for diagnosing causes of sleep disorders in primary care. *Journal of Primary Health Care*, *3*(2), 107–113. https://doi.org/10.1071/HC11107

Bailes, S., Baltzan, M., Rizzo, D., Fichten, C. S., Amsel, R., & Libman, E. (2008). A diagnostic symptom profile for sleep disorder in primary care patients. *Journal of Psychosomatic Research*, *64*(4), 427–433. https://doi.org/10.1016/j.jpsychores.2007.10.011

Blom, K., Jernelöv, S., Rück, C., Lindefors, N., & Kaldo, V. (2017). Three-year follow-up comparing cognitive behavioral therapy for depression to cognitive behavioral therapy for insomnia, for patients with both diagnoses. *Sleep*, *40*(8). https://doi.org/10.1093/sleep/zsx108

Castaneda, A. E., Tuulio-Henriksson, A., Marttunen, M., Suvisaari, J., & Lönnqvist, J. (2008). A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults. *Journal of Affective Disorders*, *106*(1–2), 1–27. https://doi.org/10.1016/j.jad.2007.06.006

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Disner, S. G., Shumake, J. D., & Beevers, C. G. (2017). Self-referential schemas and attentional bias predict severity and naturalistic

course of depression symptoms. *Cognition & Emotion*, *31*(4), 632–644. https://doi.org/10.1080/02699931.2016.1146123

Douglass, A. B., Bomstein, R., Nino-Murcia, G., Keenan, S., Miles, L., Zarcone Jr, V. P., Guillemninault, C., & Dement, W. C. (1994). The sleep disorders questionnaire I: Creation and multivariate structure of SDQ. *Sleep*, *17*(2), 160–167

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, *42*(3), 509–529. https://doi.org/10.1080/00273170701382864

Freeman, D., Sheaves, B., Waite, F., Harvey, A. G., & Harrison, P. J. (2020). Sleep disturbance and psychiatric disorders. *The Lancet Psychiatry*, *7*(7), 628–637. https://doi.org/10.1016/S2215-0366(20)30136-X

Gieselmann, A., Ait Aoudia, M., Carr, M., Germain, A., Gorzka, R., Holzinger, B., Kleim, B., Krakow, B., Kunze, A. E., Lancee, J., Nadorff, M. R., Nielsen, T., Riemann, D., Sandahl, H., Schlarb, A. A., Schmid, C., Schredl, M., Spoormaker, V. I., Steil, R., & Pietrowsky, R. (2019). Aetiology and treatment of nightmare disorder: State of the art and future perspectives. *Journal of Sleep Research*, *28*(4), e12820. https://doi.org/10.1111/jsr.12820

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. In *Midwest research-to-practice Conference in adult*. Continuing, and Community Education.

Hombali, A., Seow, E., Yuan, Q., Chang, S. H. S., Satghare, P., Kumar, S., Verma, S. K., Mok, Y. M., Chong, S. A., & Subramaniam, M. (2019). Prevalence and correlates of sleep disorder symptoms in psychiatric disorders. *Psychiatry Research*, *279*, 116–122. https://doi.org/10.1016/j.psychres.2018.07.009

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Iasiello, M., Muir-Cochrane, E., van Agteren, J., & Fassnacht, D. B. (2022). The Effect of psychological distress on measurement invariance in measures of mental Wellbeing. *International Journal of Environmental Research and Public Health*, *19*(16), 10072. https://doi.org/10.3390/ijerph191610072

Kahn, R. S., & Keefe, R. S. (2013). Schizophrenia is a cognitive illness: Time for a change in focus. *JAMA Psychiatry*, *70*(10), 1107–1112. https://doi.org/10.1001/jamapsychiatry.2013.155

Kenny, D. A. (2015). Measuring model fit. https://davidakenny.net/cm/fit.htm

Kerkhof, G. A. (2017). Epidemiology of sleep and sleep disorders in The Netherlands. *Sleep Medicine*, *30*, 229–239. https://doi.org/10.1016/j.sleep.2016.09.015

Kerkhof, G. A., Geuke, M. E., Brouwer, A., Rijsman, R. M., Schimsheimer, R. J., & Van Kasteel, V. (2013). Holland sleep disorders questionnaire: A new sleep disorders questionnaire based on the international classification of sleep disorders-2. *Journal of Sleep Research*, *22*(1), 104–107. https://doi.org/10.1111/j.1365-2869.2012.01041.x

Khazaie, H., Jalali, A., Mohammadi, R., Chehri, A., & Khazaie, A. (2022). Assessment of the psychometric properties of the holland sleep disorders questionnaire in the iranian population. *Sleep Disorders*, *2022*, 1–11. https://doi.org/10.1155/2022/1367067

Koffel, E., & Watson, D. (2010). Development and initial validation of the Iowa sleep disturbances Inventory. *Assessment*, *17*(4), 423–439. https://doi.org/10.1177/1073191110362864

Maher, A. R., Apaydin, E. A., Hilton, L., Chen, C., Troxel, W., Hall, O., Azhar, G., Larkin, J., Motala, A., & Hempel, S. (2021). Sleep management in posttraumatic stress disorder: A systematic review and meta-analysis. *Sleep Medicine*, *87*, 203–219. https://doi.org/10.1016/j.sleep.2021.08.016

Malkewitz, C. P., Schwall, P., Meesters, C., & Hardt, J. (2023). Estimating reliability: A comparison of Cronbach's α, McDonald's ωt and the greatest lower bound. *Social Sciences & Humanities Open*, *7*(1), 100368. https://doi.org/10.1016/j.ssaho.2022.100368

Manber, R., Bernert, R. A., Suh, S., Nowakowski, S., Siebern, A. T., & Ong, J. C. (2011). CBT for insomnia in patients with high and low depressive symptom severity: Adherence and clinical outcomes. *Journal of Clinical Sleep Medicine*, *7*(6), 645–652. https://doi.org/10.5664/jcsm.1472

McNicholas, W. T. (2008). Diagnosis of obstructive sleep apnea in adults. *Proceedings of the American Thoracic Society*, *5*(2), 154–160.

Menold, N. (2020). Double barreled questions: An analysis of the similarity of elements and effects on measurement quality. *Journal of Official Statistics*, *36*(4), 855–886. https://doi.org/10.2478/jos-2020-0041

Mijnster, T., Boersma, G. J., van Veen, M. M., Lieburg, E., Cath, D., Pijnenborg, G. H., De Jong, P. J., & Lancel, M. (2024). Sleep disorders in a naturalistic cohort of Dutch psychiatric outpatients: Prevalence rates and associations with psychopathology symptom severity and well-being. *Journal of Sleep Research*, *33*(2), e14009. https://doi.org/10.1111/jsr

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Roca, M., López-Navarro, E., Monzón, S., Vives, M., García-Toro, M., García-Campayo, J., & Harrison & Gili, M. (2015). Cognitive episodes. *European Neuropsychopharmacology*, *25*(11), 1991–1998. https://doi.org/10.1016/j.euroneuro.2015.07.020

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029–2040. https://doi.org/10.1017/S0033291713002535

Rosseel, Y. (2012). lavaan: An R package for structural Equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Roth, T., Zammit, G., Kushida, C., Doghramji, K., Mathias, S. D., Wong, J. M., & Buysse, D. J. (2002). A new questionnaire to detect sleep disorders. *Sleep Medicine*, *3*(2), 99–108. https://doi.org/10.1016/S1389-9457(01)00131-9

Sateia, M. J. (2014). International classification of sleep disorders. *Chest*, *146*(5), 1387–1394. https://doi.org/10.1378/chest.14-0970

Sheaves, B., Rek, S., & Freeman, D. (2022). Nightmares and psychiatric symptoms: A systematic review of longitudinal, experimental, and clinical trial studies. *Clinical Psychology Review*, *102241*, 102241. https://doi.org/10.1016/j.cpr.2022.102241

Simoons, M., Ruhé, H. G., van Roon, E. N., Schoevers, R. A., Bruggeman, R., Cath, D. C., Muis, D., Arends, J., Doornbos, B., & Mulder, H. (2019). Design and methods of the 'monitoring outcomes of psychiatric pharmacotherapy'(MOPHAR) monitoring program–a study protocol. *BMC Health Services Research*, *19*(1), 1–10. https://doi.org/10.1186/s12913-019-3951-2

Spera, V., Maiello, M., Pallucchini, A., Novi, M., Elefante, C., De Dominicis, F., Palagini, L., Biederman, J., & Perugi, G. (2020). Adult attention-deficit hyperactivity disorder and clinical correlates of delayed sleep phase disorder. *Psychiatry Research*, *291*, 113162. https://doi.org/10.1016/j.psychres.2020.113162

Spiliotopoulou, G. (2009). Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy. *Australian Occupational Therapy Journal*, *56*(3), 150–155. https://doi.org/10.1111/j.1440-1630.2009.00785.x

Spoormaker, V. I., Verbeek, I., van den Bout, J., & Klip, E. C. (2005). Initial validation of the SLEEP-50 questionnaire. *Behavioral Sleep Medicine*, *3*(4), 227–246. https://doi.org/10.1207/s15402010bsm0304_4

Stewart, R., Besset, A., Bebbington, P., Brugha, T., Lindesay, J., Jenkins, R., Singleton, N., & Meltzer, H. (2006). Insomnia comorbidity and impact

and hypnotic use by age group in a national survey population aged 16 to 74 years. *Sleep*, *29*(11), 1391–1397. https://doi.org/10.1093/sleep/29.11.1391

Stubbs, B., Vancampfort, D., Veronese, N., Solmi, M., Gaughran, F., Manu, P., Rosenbaum, S., De Hert, M., & Fornaro, M. (2016). The prevalence and predictors of obstructive sleep apnea in major depressive disorder, bipolar disorder and schizophrenia: A systematic review and meta-analysis. *Journal of Affective Disorders*, *197*, 259–267. https://doi.org/10.1016/j.jad.2016.02.060

Taple, B. J., Chapman, R., Schalet, B. D., Brower, R., & Griffith, J. W. (2022). The impact of education on depression assessment: Differential item functioning analysis. *Assessment*, *29*(2), 272–284. https://doi.org/10.1177/107319112097

Warrens, M. J. (2015). Some relationships between Cronbach's alpha and the Spearman-Brown formula. *Journal of Classification*, *32*(1), 127–137. https://doi.org/10.1007/s00357-015-9168-0

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mijnster, T., Wardenaar, K. J., Boersma, G. J., van Veen, M. M., Cath, D., Kerkhof, G. A., & Lancel, M. (2024). The Holland Sleep Disorders Questionnaire: Factorial structure and measurement invariance in a psychiatric sample relative to the general population. *International Journal of Methods in Psychiatric Research*, e70004. https://doi.org/10.1002/mpr.70004