Title

Development and validation of the Hypersomnia Severity Index (HSI): A measure to assess hypersomnia severity and impairment in psychiatric disorders.

Authors

Kaplan, Katherine
Plante, David
Cook, Jesse
et al.

Publication Date

2019-11-01

DOI

10.1016/j.psychres.2019.112547

Peer reviewed

HHS Public Access

Author manuscript

Psychiatry Res. Author manuscript; available in PMC 2020 November 01.

Published in final edited form as:

Psychiatry Res. 2019 November ; 281: 112547. doi:10.1016/j.psychres.2019.112547.

Development and Validation of the Hypersomnia Severity Index (HSI): A Measure to Assess Hypersomnia Severity and Impairment in Psychiatric Disorders

Katherine A. Kaplan[1], **David T. Plante**[2], **Jesse D. Cook**[2], **Allison G. Harvey**[3]

[1]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine

[2]Department of Psychiatry, University of Wisconsin School of Medicine and Public Health

[3]Department of Psychology, University of California, Berkeley

Abstract

Hypersomnia is common in psychiatric disorders, yet there are few self-report measures that adequately characterize this sleep disturbance. The objective of this study was to validate the Hypersomnia Severity Index (HSI), a tool designed to measure severity, distress and impairment of hypersomnia in psychiatric populations. Psychometric properties were evaluated in an undergraduate Scale Development sample (N=381) and two psychiatric Scale Validation samples: euthymic bipolar participants with a range of sleep complaints (N=89), and unmedicated unipolar depressed participants (N=21) meeting operational criteria for hypersomnolence disorder. Exploratory factor analysis and confirmatory factor analysis in the Scale Development and Validation samples, respectively, suggested a two-factor structure representing Hypersomnia Symptoms and Distress/Impairment best fit the data. Convergent validity was established by significant associations with the Epworth Sleepiness Scale (ESS), Pittsburgh Sleep Quality Index (PSQI), and the Sheehan Disability Scale in both samples. Construct validity was further supported by significant correlations between the Scale Validation sample and two weeks of diary- and actigraphy-determined total sleep time and time in bed. A cutoff score of 10 maximally discriminated between those with hypersomnia and those without. The HSI shows promise as a measure of hypersomnia that is commonly seen in psychiatric disorders, and may be of use to both researchers and clinicians.

Keywords

Sleepiness; long sleep; hypersomnolence; self-report; mood disorders; psychometric validation; assessment

Corresponding Author Katherine A. Kaplan Ph.D., Address: 401 Quarry Road, Stanford CA 94305 USA, Telephone: 650-200-0131, katekate@stanford.edu.

Conflict of Interest

KAK has served as a consultant for Balance Therapeutics. DTP has served as a consultant for Teva Pharmaceuticals Australia and Jazz Pharmaceuticals. All other authors report no competing interests.

1. INTRODUCTION

Hypersomnia, generally defined via excessive total sleep time and/or sleepiness, is common across psychiatric disorders. The Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) notes hypersomnia as a diagnostic criterion across mood disorders – namely, Bipolar I and II disorders (BD), MDD (MDD), persistent depressive disorder and premenstrual dysphoric disorder – with prevalence estimates ranging from 30 – 50% across major depressive and bipolar disorders (Grigolon et al., 2018; Kaplan and Harvey, 2009). Moreover, the burden of hypersomnia in these conditions is significant. Hypersomnia has been linked to poor illness course (Worthington et al., 1995; Zimmerman et al., 2005), suicide risk (Goldstein et al., 2008) and illness relapse (Cho et al., 2008; Kaplan et al., 2015).

The definition of hypersomnia has evolved over the past two decades, differing across diagnostic nomenclatures and within subsequent editions (see Plante, 2015 for review). Currently, the DSM-5 defines Hypersomnolence Disorder via a prolonged main sleep period, frequent naps, or difficulty awakening after abrupt awakenings, along with distress or impairment. Hypersomnia as a diagnostic criterion of depressive disorders, however, is described in more general terms as long nighttime sleep or frequent daytime sleep. The atypical specifier of MDD further changes the requirements of hypersomnia to include a total daily sleep time of 10 hours or two hours greater than the euthymic sleep duration (American Psychiatric Association, 2013). The second edition of the International Classification of Sleep Disorders (ICSD-2; American Academy of Sleep Medicine, 2005) required only self-reported excessive daytime sleepiness or excessive sleep, while the updated ICSD-3 (American Academy of Sleep Medicine, 2014) requires irrepressible need to sleep or daytime lapses into sleep in order to meet criteria for hypersomnia but no longer emphasizes prolonged total sleep time in its primary criteria. Clearly, discrepancies among diagnostic nosologies (e.g., the inclusion of a long main sleep period to define hypersomnia) has implications for hypersomnia evaluation.

Paralleling the difficulty in defining hypersomnia, measuring hypersomnia in clinical practice can be a challenge. A variety of available self-report instruments capture individual features of hypersomnia. For example, the Epworth Sleepiness Scale (Johns, 1991) and the Sleep Inertia Questionnaire (Kanady and Harvey, 2015) assess excessive daytime sleepiness and excessive morning sleepiness, respectively, but would not adequately capture hypersomnia manifested by long sleep duration. Likewise, instruments such as the Functional Outcomes of Sleepiness Questionnaire (Weaver et al., 1997) may be useful to assess impairment related to sleepiness or tiredness but would not capture distress. There is no single self-report measure available to assess primary features of psychiatric hypersomnia (excessive sleep and/or sleepiness) together with associated distress and/or impairment. Hence, the overall aim of the present research was to develop and evaluate the psychometric properties of a brief self-report instrument for hypersomnia, the Hypersomnia Severity Index (HSI). We focused our initial validation on psychiatric hypersomnia given its prevalence relative to other hypersomnia disorders, its associated morbidity and the absence of an instrument currently available to assess it.

The present research has two aims, addressed using two separate samples. Aim 1 sought to explore reliability, validity and factor structure of the HSI in a large Scale Development sample. To address this aim, we sampled undergraduate students from the University of California, Berkeley. Aim 2 sought to further explore reliability, validity and confirm factor structure of the HSI in a Scale Validation sample, as well as to establish the construct validity of the HSI against prospective sleep measurement. To address this aim, we sampled individuals with psychiatric disorders (MDD and BD) and collected sleep diary and actigraphy data to validate the HSI as a measure sensitive to increased sleep.

2. METHODS

2.1 Scale Validation Overview

The HSI was validated in five steps. First, an initial set of items was selected using definitions of psychiatric hypersomnia in current diagnostic nosologies and subjected to review by a panel of sleep experts. Second, items were administered to a large Scale Development sample of undergraduates to explore reliability, validity and explore underlying factor structure using exploratory factor analysis (EFA). Third, a separate Scale Validation sample of individuals with mood disorders was collected to confirm factor structure using confirmatory factor analysis (CFA). Fourth, convergent and construct validity of the HSI was established against available instruments and prospective sleep monitoring, respectively. Finally, criterion validity and a criterion cutoff score were established by comparing across groups. The Committee for the Protection of Human Subjects at the University of California, Berkeley and the University of Wisconsin-Madison Health Sciences Institutional Review Board approved the procedures described in this study.

2.2 Item Selection

Definitions of hypersomnia from available diagnostic nomenclatures, along with operational definitions used in research on psychiatric hypersomnia, were reviewed. As data collection commenced in 2010, diagnostic systems corresponded to the DSM-IV and ICSD-2. Likert-type scales were created to form a composite of the definitions in these diagnostic nomenclatures. Questions assessing functional impairment were added that were identical to those in the Insomnia Severity Index (Morin et al., 2011), a validated and widely-used self-report instrument to measure insomnia severity and related distress/impairment. All items were subsequently evaluated and refined by a panel of nine sleep experts. This resulted in removal of redundancies and clarification of ambiguities.

2.3 Samples

To explore reliability, validity and factor structure, the HSI was first administered to a Scale Development sample of undergraduate students (N=381). The racial and ethnic breakdown of this sample, described in Table 1, reflects the diversity of University of California student population. All undergraduates completed the HSI and related validation measures (described below) online for research participation credit. While individuals in this group were not selected based on the presence of depressive or hypersomnia symptomatology, a sizeable portion of the student sample endorsed depressive symptoms in the moderate to severe range (29% scoring ≥ 11 on the QIDS-SR).

To confirm reliability, validity and factor structure, the HSI and related measures were administered to two separate groups, which together comprised the Scale Validation sample (N=110). The first psychiatric group consisted of individuals with bipolar spectrum disorders (I=80, II=3, NOS=3) and a range of sleep complaints who were currently interepisode. Individuals in this group were recruited for larger parent studies involving sleep and BD (Gershon et al., 2012; Harvey et al., 2015). Individuals were required to meet DSM-IV criteria for a diagnosis of BD as determined by the Structured Clinical Interview for the DSM-IV (SCID; First et al., 1997) and confirmed interepisode via SCID and established cutoff scores on the Inventory of Depressive Symptomatology, Clinician Version (IDS-C; Rush et al., 1996) and Young Mania Rating Scale (YMRS; Young et al., 1978). Individuals were excluded if they met criteria for current substance or alcohol abuse or dependence; narcolepsy, sleep apnea, restless leg syndrome or periodic limb movement disorder based on the Duke Structured Interview for Sleep Disorders (DSISD; Edinger et al., 2004); severe head trauma, stroke, neurological disease, or severe medical illness. Participants were not excluded on the basis of comorbidities or pharmacological treatments, given that comorbidity and polytherapy are common features of BD.

The second psychiatric group comprising the scale validation sample consisted of individuals with unipolar MDD (N=21) meeting operational criteria for hypersomnolence disorder (“MDD+HS”). This group was prospectively collected as part of a larger study examining sleep disturbance in MDD. All participants in this group were unmedicated and diagnosed with unipolar MDD via the SCID. Sleep and medical disorders were ruled out via semi-structured history and physical exam. MDD+HS additionally met operationalized criteria for hypersomnolence disorder proposed by Ohayon and colleagues (Ohayon et al., 2012), which were later adopted in the DSM-5 with only minor changes.

2.4 Instruments

Three retrospective self-report instruments and two forms of prospective sleep measurement were used to establish the validity of the HSI.

The Epworth Sleepiness Scale (ESS; Johns, 1991) is an 8-item self-report measure of excessive daytime sleepiness. Items assess propensity for falling asleep in common daytime situations, yielding a composite score of sleepiness severity with scores > 10 representing excessive sleepiness. The ESS has shown good internal consistency and high test-retest reliability (Johns, 1992). Internal consistency in our samples was good (Scale Development α = 0.80, Scale Validation α = 0.76).

The Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989) is a 19-item self-report measure assessing prior-month subjective sleep quality, yielding a global score and seven component scores. The PSQI has been shown to have good internal consistency and test-retest reliability (Carpenter and Andrykowski, 1998). Two items from the PSQI were chosen as indicators of hypersomnia. The first was an item assessing self-reported sleep duration in the past month (Question 4), which has been validated against actigraphy and sleep diary in various samples (Backhaus et al., 2002; Grandner et al., 2006) and has been used to estimate habitual sleep duration in previous research (King et al., 1997; Knutson et al., 2006). The second indicator was the Daytime Dysfunction subscale, derived from two questions about

excessive sleepiness and daytime impairment. This subscale has been validated against other measures of daytime impairment (e.g. Buysse et al., 2008). Internal consistency was not calculated for the PSQI subscale given the reduced number of response items.

Functional impairment was assessed using the Sheehan Disability Scale (Sheehan et al., 1996). The Sheehan Disability Scale is a brief three-item questionnaire asking for 0–10 ratings on sleep-related impairment in work, social life, and family life, yielding a total impairment score (0, not impaired, to 30, highly impaired). Psychometrics are well established (Sheehan et al., 1996) and internal consistency in our sample was high (Scale Development α = 0.83, Scale Validation α = 0.97).

To measure sleep prospectively, all participants in the Scale Validation sample kept sleep diaries and wore wrist actigraphy for approximately two weeks (15.5 ± 5.1 diary days, 13.9 ± 6.7 actigraphy days). The sleep diary is considered the gold standard subjective measure of sleep (Buysse et al., 2006; Carney et al., 2012). Participants completed the log upon waking, and compliance was confirmed in a subset of participants by calls to a voicemail. Total sleep time was calculated by subtracting all time spent awake from all time spent in bed over a 24 hour period including naps. Time in bed was scored by summing all intended sleep periods. Naps were included in total sleep time and time in bed calculations given that individuals with hypersomnia may have longer sleep (Plante et al., 2017) and/or bedrest (Billiard et al., 1994) durations.

A subset of Scale Validation participants (N=77) were also equipped with an actigraph (Actiwatch AW-64 in the bipolar sample and Actiwatch-2 in the unipolar sample; Philips Respironics Inc., Bend OR) to obtain an objective estimate of sleep. Actigraphs are small wrist watch-like devices that provide an empirical estimate of the sleep/wake cycle via movement. Actigraphy has been previously validated in bipolar disorder (Kaplan et al., 2012) and depression comorbid with insomnia (McCall and McCall, 2012). Analyses were completed using the medium sensitivity setting in Actiware. Mirroring the variables extracted from sleep diaries, total daily sleep time and daily time in bed were extracted from actigraph output.

2.5 Statistical Analyses

To address Aim 1 in the Scale Development sample, reliability, validity, and exploratory factor analysis (EFA) were evaluated using SPSS version 24. The number of factors retained in EFA (principal axis factoring considering both varimax and promax rotations) was determined using break in scree plot. Items with factor loadings below 0.40 were evaluated for exclusion.

To address Aim 2 in the Scale Validation sample, confirmatory factor analysis (CFA) was evaluated using “lavaan” R package version 0.5–18. Missing data, minimal for all measures (<3%), were first multiply imputed using the Expectation Maximization algorithm in the “Amelia” R package version 1.7.3 before proceeding. Model fit was evaluated using established standards including chi-square to degrees of freedom ratio (χ^2/df) ≤ 3, comparative fit indices (CFI) and Tucker-Lewis indices (TLI) > 0.85, and the root mean

square error of approximation (RMSEA) < 0.08 and standardized root mean square residual (SRMR) of <0.08 (Hair et al., 1998; Hu and Bentler, 1995).

For both aims, reliability was determined using Chronbach alpha coefficients. Convergent validity was established by comparing the HSI against PSQI (both the Daytime Dysfunction subscale and Sleep Duration item), Epworth Sleepiness Scale and Sheehan Disability Scale. Construct validity was demonstrated by comparing HSI scores against sleep diary and actigraphy total sleep time and time in bed. Criterion validity was established by comparing across groups, and a criterion cutoff score was determined by calculating the score that maximized the sum of sensitivity and specificity.

3. RESULTS

3.1 Demographics and Scale Characteristics.

Sociodemographic data are presented in Table 1. The Scale Validation samples tended to be younger, and both Scale Validation samples contained more female participants ($p<0.05$, χ^2), two characteristics more commonly associated with psychiatric hypersomnia (Kaplan and Harvey, 2009). Consistent with inclusion criteria, both the Undergraduate and BD samples reported depressive symptoms in the mild range (<11) on the QIDS-SR, while the MDD+HS sample endorsed depressive symptoms in the moderate range (BDI-II≥20).

Means and standard deviations for all self-report and prospective measures can be found in Table 2. One-way ANOVAs with Scheffe's test for post-hoc comparisons were used to compare the three groups, and Student's t-tests with Bonferroni corrections were used to compare two groups. The three subgroups differed on Epworth Sleepiness Scale Scores (MDD>Undergraduate>Bipolar Disorder), the two psychiatric subgroups reported greater PSQI Daytime Dysfunction and functional impairment than the undergraduate group, and the MDD group reported longer sleep durations on the PSQI sleep duration item. The two psychiatric subgroups did not differ on any prospective measure except diary-reported total sleep time ($p<0.001$). To examine validity in the sections that follow, the two psychiatric subgroups were combined to form the Scale Validation sample.

3.2 Factorial Validity

An exploratory factor analysis using principal axis factoring was performed in the Scale Development Sample. We considered both promax (allowing factors to correlate) and varimax (constraining factors to orthogonality) factor rotations, with the goal of extracting the highest number of factors that were interpretable. Results from varimax rotation with Kaiser normalization rotation are reported below. Inspection of the break in slope on the scree plot indicated that two factors should be retained in the final solution. These two factors together accounted for 56% of the total variance. The first factor ("Distress/ Impairment/Difficulty") contained all of the distress/impairment items and the difficulty waking up in the morning item, with all item factor loadings ≥ 0.53. The second factor ("Hypersomnia Symptoms") contained items related to sleeping during the day and feeling sleepy during the daytime, with item factor loadings ≥ 0.60. The items regarding sleeping too much at night and sleep attacks in the daytime did not reliably load onto either factor.

However, given threat to construct validity with these items removed, they were retained and added to the second factor (containing other sleep items) for interpretability.

A confirmatory factor analysis was performed in the Scale Validation Sample, with questionnaire items loading onto the two factors as described above. However, fit statistics from this initial model were problematic (χ^2/ df = 2.52, CFI = 0.89, TLI = 0.85, RMSEA = 0.12, SRMR = 0.083). Examination of modification indices strongly suggested model improvement with the "difficulty waking up in the morning" item added to the Hypersomnia Symptoms factor. Fit statistics were improved with this revised model (χ^2/ df = 2.01, CFI = 0.93, TLI = 0.90, RMSEA = 0.097, SRMR = 0.073). The present model also fit the data significantly better than a single-factor solution ($\chi^2(1) = 19.6$, $p<0.001$) or a two-factor solution that did not allow for covariances among latent factors ($\chi^2(1) = 47.6$, $p<0.001$). In sum, the final scale consisted of two factors, a Hypersomnia Symptoms factor and a Distress/Impairment factor. The Hypersomnia Symptoms factor included items related to "sleep too much at night," "difficulty waking up in the morning or from naps," "sleep during the day," "feel sleepy during the daytime," and "sleep attacks." The Distress/Impairment subscale included remaining items pertaining to "satisfied," "interfere," "noticeable" and "worried/distressed."

3.3 Internal Consistency

Internal consistency of HSI overall was good (Scale Development $\alpha = 0.82$; combined Scale Validation $\alpha = 0.84$). Examining internal consistency within factors, reliability was high for the Distress/Impairment factor (Scale Development $\alpha = 0.82$, combined Scale Validation $\alpha = 0.88$) but attenuated for the Hypersomnia Symptoms factor (Scale Development $\alpha = 0.61$, combined Scale Validation $\alpha = 0.65$), which may reflect heterogeneity in the Hypersomnia Symptoms construct and is not necessarily a threat to the measure's utility (Schmitt, 1996). Pearson correlation coefficients between the two factors, and between each factor and the HSI total, were all moderate to high (Scale Development 0.61–0.90, Scale Validation 0.59–0.89), suggesting good construct validity.

3.4 Convergent Validity

Convergent validity was assessed by evaluating Pearson correlations between the HSI (total score and subscales) and Epworth Sleepiness Scale, PSQI Daytime Dysfunction subscale, PSQI Sleep Duration and Sheehan Disability Scales. Results from both samples are presented in Table 3. Correlations were all moderate to strong, and in expected directions, in both the Scale Development and Scale Validation BD Samples. As expected, self-reported measures of sleep (e.g. Epworth Sleepiness Scale) correlated more strongly with the HSI Hypersomnia Symptoms Factor, while self-reported measures of impairment (Sheehan Disability Scale, PSQI Daytime Dysfunction) were more strongly associated with the HSI Distress/Impairment factor, supporting convergent validity. The PSQI Sleep Duration item showed only a weak relationship to the HSI total and subscales, though was associated with the HSI Hypersomnia Symptoms item in the expected duration in the Scale Validation Sample. Overall, the correlations indicates good convergent validity, and suggest the HSI subscales represent content consistent with disturbed sleep and distress or impairment, respectively.

3.5 Construct Validity

Construct validity was demonstrated by comparing HSI scores against sleep diary and actigraphy total sleep time and time in bed in the Scale Validation sample. Results are presented in Table 4. Small to moderate associations were observed between most diary and actigraphy variables and the HSI total and subscale scores. Interestingly, the HSI Hypersomnia Symptoms subscale was more consistently associated with sleep diary measures, suggesting concordance between self-reported sleep disturbance and self-reported diary measures, while the HSI Distress/Impairment subscale was more consistently associated with actigraphy measures of sleep.

3.6 Criterion Validity

Because individuals in the MDD+HS all met research diagnostic criteria for hypersomnolence (Ohayon et al., 2012) while the BD subgroup exhibited a range of sleep complaints, we expected HSI scores to be higher in the depression subgroup than in the bipolar subgroup. Moreover, as the HSI was designed to assess psychiatric hypersomnia, we expected HSI scores among psychiatric patients to be higher than those of the undergraduate sample. To establish this criterion validity, we conducted a one-way ANOVA with Scheffe's test for post-hoc comparisons to compare the three subgroups. Results are listed in Table 2. In support of criterion validity, the MDD+HS sample demonstrated significantly greater HSI total scores than the BD and undergraduate samples ($p<0.001$), as well as significantly greater HSI Hypersomnia Symptoms subscale scores compared with the other two subgroups ($p<0.001$). Both psychiatric groups reported greater scores on the HSI Distress/Impairment subscale compared to the Undergraduate sample ($p<0.001$).

3.7 Criterion Cutoff Score

To identify a HSI cutoff score that would maximally differentiate individuals with hypersomnia from individuals with confirmed absence of hypersomnia, we compared our MDD+HS group to a separate sample of individuals with confirmed absence of Axis I psychopathology and not meeting operational criteria for hypersomnolence (N=23, 78% female, mean age 28.8±5.4 years). A cutoff score of 10 was found to maximally differentiate between MDD+HS and this control group.

4. DISCUSSION

The goal of the present research was to develop and evaluate a brief self-report instrument to assess the severity of psychiatric hypersomnia, along with associated distress and impairment. Because such a measure does not currently exist, and given the prevalence and associated consequences of hypersomnia within psychiatric disorders, having a simple, short assessment available may help in addressing the burden of hypersomnia. The first aim of the present investigation was to explore the factor structure, reliability and validity of the HSI in a large Scale Development sample of undergraduates. Analyses suggested a two-factor structure of the instrument that together accounted for 56% of the total variance. Internal consistency of the instrument overall and for the Distress/Impairment factor was high, though Chronbach's alpha coefficients were reduced for the Hypersomnia Symptoms factor. Rather than viewed as a threat to reliability (Schmitt, 1996), this reduced alpha possibly

reflects the heterogeneity within the Hypersomnia Symptoms factor, which evaluates long sleep, excessive sleepiness and excessive sleep inertia. Previous research has called into question the overlap between these features (Kaplan et al., 2015; Ohayon et al., 2012). Supporting good convergent validity, correlations between the HSI total and subscales and the Epworth Sleepiness Scale, PSQI Daytime Dysfunction, and Sheehan Disability Scales were all high. Higher scores on the PSQI Sleep Duration item (i.e., longer reported sleep durations) were weakly associated with *lower* scores on the HSI Distress/Impairment factor, which likely reflects low or variable sleep durations commonly seen among college students (Lund et al., 2010).

The second aim was to confirm the factor structure and establish reliability and validity in a Scale Validation sample of individuals with psychiatric diagnoses. Here we focused on individuals diagnosed with BD who had a range of sleep complaints, as well as individuals meeting criteria for MDD as well as operational criteria for HS derived from US epidemiologic survey data (Ohayon et al., 2012). A confirmatory factor analysis lent support for the two-factor structure of the HSI, with improvement in model fit noted when “difficulty waking up in the morning or from naps” was moved to the Hypersomnia Symptoms factor. Of the hypersomnia symptoms, the “difficulty waking up” item was endorsed most highly in our undergraduate Scale Development sample, which may reflect insufficient sleep opportunity, evening chronotype, hypersomnolence or a combination of these factors. It is possible, then, that this item thus captured general sleep distress in the Scale Development Sample, which was not present in the Scale Validation sample, explaining why model fit improved with its move to the Hypersomnia Symptoms factor in the latter sample. Consistent with the Scale Development sample, internal consistency was generally high and convergent validity was supported by significant associations between the HSI total and subscales and self-report measures of sleepiness and impairment.

Within the Scale Validation subgroups, higher scores on the HSI generally corresponded to greater durations of prospectively-measured sleep and bedrest, lending support for construct validity. However, these correlations were modest. Instead, as a self-report instrument, the HSI total and subscales showed a much stronger relationship to other self-report measures of sleep disturbance and distress (i.e. Epworth Sleepiness Scale, Sheehan Disability Scale) in the BD sample. Previous research has shown that self-report sleep measures such as the PSQI and the ESS similarly show weak relationships to objectively-measured sleep data in community samples (Buysse et al., 2008) but the PSQI is still regarded as an essential measure to capture subjective sleep experiences in insomnia treatment trials (Buysse et al., 2006). Thus the HSI may still be important in establishing subjective complaints about sleep-related behaviors, distress and impairment, even if relationships to actual sleep data are modest.

Criterion validity was established by comparing HSI scores between our three groups: undergraduates, BD and MDD+HS. We expected that the HSI total and subscale scores should be greatest for the MDD+HS group relative to the other two groups, and that scores among the psychiatric samples should be higher than that of our undergraduate sample. We found support for the former but not for the latter. That is, we found the MDD+HS group demonstrated higher total and Hypersomnia Symptoms subscale scores relative to the other

two samples. However, while both psychiatric groups endorsed greater Distress/Impairment relative to the undergraduate sample, we found high rates of sleepiness and reported sleep disturbance in our undergraduate sample that were generally on par with, if not superior to, rates within BD. This is consistent with research suggesting the college years are a period of heightened vulnerability to daytime sleepiness and poor quality sleep (see Hershner and Chervin, 2014 for review). Using college samples as analogues to patient populations provides advantages in cost, feasibility and ease of recruiting large samples necessary to evaluate EFAs with multiple response items. Indeed, initially developing a measure with an undergraduate sample and then establishing validity in a specific population has been used to develop other sleep questionnaires over the past decade (e.g. Kanady and Harvey, 2015; Koffel and Watson, 2010). However, we also recognize that our undergraduate sample was both younger and more diverse (particularly regarding Asian participants) than our psychiatric samples, which makes direct comparison difficult. The racial composition of this undergraduate sample also did not reflect general college or population samples more broadly.

While convergent validity was established with comparisons to other accepted scales and subscales, our primary aim in developing this instrument was to assess hypersomnia symptoms and associated distress/impairment, not to create an instrument that would be maximally discriminable from insomnia symptoms. Hypersomnia and insomnia symptoms can and do overlap in patients with neuropsychiatric disorders, particularly mood disorders (Geoffroy et al., 2018; Liu et al., 2007; Soehner et al., 2014). Future research is needed to develop specific scales that can both quantify the full nature and type of sleep disturbance (e.g. hypersomnia and insomnia) as well as the degree of impairment and distress attributable to each these specific sleep complaints.

Finally, we preliminarily explored an optimal cutoff score for the HSI by comparing our MDD+HS group to a separate matched sample of control individuals who did not meet operational criteria for hypersomnolence. The optimal cutoff score, determined by maximizing the sum of sensitivity and specificity, was determined to be 10. However, we emphasize that our criterion and comparison groups were small and replication in larger samples is warranted.

The present research evaluated the HSI as an instrument to assess hypersomnia severity and impairment at a single point in time. However, an important feature of any measure is establishing its stability across time and/or sensitivity to clinical intervention. For example, the ISI (Bastien et al., 2001) is a recommended/essential measure of insomnia (Buysse et al., 2006) that has shown good responsivity to treatment (Morin et al., 2011) and is often a primary outcome measure in insomnia treatment trials. Evaluating the HSI longitudinally was beyond the scope of this initial validation paper. However, we can offer one hint as to its temporal stability. In a subset of the present participants with bipolar disorder (N=26) who were followed over time, we did re-administer HSI at a one-month follow-up without any sleep intervention between visits. High test-retest reliability in this short follow-up window was observed ($r = 0.69$, Pearson), suggesting both stability of hypersomnia symptoms and good test-retest reliability of the HSI. Whether the HSI is sensitive to treatment response is

unknown; hence, we encourage future research to examine sensitivity to treatment as well as optimal cutoff scores to determine treatment response.

The HSI was developed for, and evaluated in, hypersomnia associated with mood disorders. The present validation samples included those with unipolar MDD and bipolar disorder, and included both euthymic and syndromal mood states. One potential limitation of the present research is that we did not evaluate the performance of the HSI in hypersomnia associated with other mood disorders (i.e., per the DSM-5, premenstrual dysphoric disorder and persistent depressive disorder), nor did we evaluate its performance in specific subtypes (i.e., depression with atypical specifier). Perhaps more importantly, hypersomnia is a cardinal symptom of disorders collectively referred to as Central Disorders of Hypersomnolence in the ICSD-3, including Narcolepsy Type I and II and Idiopathic Hypersomnia. Evaluating the utility of HSI in these groups will thus be an important area for future research. Finally, we wish to underscore that the definitions of hypersomnia across diagnostic nosologies have evolved over the last decade (Plante, 2015), and newer diagnostic systems were published after the period in which the HSI was first developed and initial evaluation had begun. It is likely that the definition of hypersomnia will change further still, and future revisions to this preliminary instrument are needed to address the changing definitions of hypersomnia.

In sum, the HSI (Appendix A) is an accessible, brief measure of hypersomnia symptoms and related distress/impairment, which may of value in research and clinical settings. Given the high rates of hypersomnia across mood disorders, as well as its impact on illness severity and quality of life, assessing hypersomnia is an encouraging first step towards addressing it clinically.

Acknowledgments

Support: This work is supported by grants from the American Sleep Medicine Foundation (76-JF-12), the Brain and Behavior Research Foundation (19193), and NIMH K23MH099234 (DTP); National Science Foundation Graduate Research Fellowship Program and Stanford Child Health Research Institute (KAK); and R34MH080958 and R01MH105513 (AGH).

Appendix A.: The Hypersomnia Severity Index.

1. For these next few questions, please consider your SLEEP IN THE PAST MONTH. To what extent do you think that you:

	Not at All	A Little	Somewhat	A Lot	Very Much
Sleep too much at night?	0	1	2	3	4
Have difficulty waking up in the morning or from naps?	0	1	2	3	4
Sleep during the day?	0	1	2	3	4
Feel sleepy during the daytime?	0	1	2	3	4

2. How SATISFIED/dissatisfied are you with your current sleep pattern?

Very satisfied		Moderately satisfied		Very dissatisfied
0	1	2	3	4

3. To what extent do you consider your sleep problem to INTERFERE with your daily functioning (e.g., daytime fatigue, ability to function at work/daily chores, concentration, memory, mood, etc.)?

Not at all	A little	Somewhat	Much	Very much
0	1	2	3	4

4. How NOTICEABLE to others do you think your sleeping problem is in terms of impairing the quality of your life?

Not at all Noticeable	Barely	Somewhat	Much	Very much Noticeable
0	1	2	3	4

5. How WORRIED/DISTRESSED are you about your current sleep problem?

Not at all	A little	Somewhat	Much	Very much
0	1	2	3	4

6. Do you ever have “sleep attacks,” defined as unintended sleep in inappropriate situations?

Not at all		Sometimes		All the time
0	1	2	3	4

References

American Academy of Sleep Medicine, 2005 International Classification of Sleep Disorders: Diagnostic and Coding Manual. 2nd Ed, 2nd ed. American Academy of Sleep Medicine, Westchester.

American Academy of Sleep Medicine, 2014 International classification of sleep disorders–third edition (ICSD-3). American Academy of Sleep Medicine, Darien, IL.

American Psychiatric Association, 2013 The Diagnostic and Statistical Manual of Mental Disorders: DSM 5, 5th ed. American Psychiatric Association, Washington, DC.

Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F, 2002 Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. Journal of Psychosomatic Research 53 (3), 737–740. [PubMed: 12217446]

Bastien CH, Vallieres A, Morin CM, 2001 Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Medicine 2 (4), 297–307. [PubMed: 11438246]

Billiard M, Dolenc L, Aldaz C, Ondze B, Besset A, 1994 Hypersomnia associated with mood disorders: a new perspective. Journal of Psychosomatic Research 38 Suppl 1, 41–47.

Buysse D, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM, 2006 Recommendations for a standard research assessment of insomnia. Sleep 29, 1155–1173. [PubMed: 17040003]

Buysse DJ, Hall ML, Strollo PJ, Kamarck TW, Owens J, Lee L, Reis SE, Matthews KA, 2008 Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical/polysomnographic measures in a community sample. Journal of Clinical Sleep Medicine 4 (6), 563–571. [PubMed: 19110886]

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ, 1989 The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Research 28 (2), 193–213. [PubMed: 2748771]

Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, Morin CM, 2012 The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 35 (2), 287–302. [PubMed: 22294820]

Carpenter JS, Andrykowski MA, 1998 Psychometric evaluation of the Pittsburgh Sleep Quality Index. Journal of Psychosomatic Research 45 (1), 5–13. [PubMed: 9720850]

Cho HJ, Lavretsky H, Olmstead R, Levin MJ, Oxman MN, Irwin MR, 2008 Sleep disturbance and depression recurrence in community-dwelling older adults: a prospective study. American Journal of Psychiatry 165 (12), 1543–1550. [PubMed: 18765482]

Edinger JD, Bonnet MH, Bootzin RR, Doghramji K, Dorsey CM, Espie CA, Jamieson AO, McCall WV, Morin CM, Stepanski EJ, 2004 Derivation of research diagnostic criteria for insomnia: report of an American Academy of Sleep Medicine Work Group. Sleep 27, 1567–1596. [PubMed: 15683149]

First M, Spitzer R, Gibbons M, Williams J, 1997 Structured Clinical Interview for DSM IV Axis 1 Disorders: Clinician Version. American Psychiatric Press, Washington, DC.

Geoffroy PA, Hoertel N, Etain B, Bellivier F, Delorme R, Limosin F, Peyre H, 2018 Insomnia and hypersomnia in major depressive episode: Prevalence, sociodemographic characteristics and psychiatric comorbidity in a population-based study. Journal of Affective Disorders 226, 132–141. [PubMed: 28972930]

Gershon A, Thompson WK, Eidelman P, McGlinchey EL, Kaplan KA, Harvey AG, 2012 Restless pillow, ruffled mind: sleep and affect coupling in interepisode bipolar disorder. Journal of Abnormal Psychology 121 (4), 863–873. [PubMed: 22845651]

Goldstein TR, Bridge JA, Brent DA, 2008 Sleep disturbance preceding completed suicide in adolescents. Journal of Consulting and Clinical Psychology 76 (1), 84–91. [PubMed: 18229986]

Grandner MA, Kripke DF, Yoon IY, Youngstedt SD, 2006 Criterion validity of the Pittsburgh Sleep Quality Index: Investigation in a non-clinical sample. Sleep and Biological Rhythms 4 (2), 129–139. [PubMed: 22822303]

Grigolon RB, Trevizol AP, Cerqueira RO, Lee Y, Mansur RB, McIntyre RS, Brietzke E, 2018 Hypersomnia and Bipolar Disorder: A systematic review and meta-analysis of proportion. Journal of Affective Disorders 246, 659–666. [PubMed: 30611064]

Hair JF, Anderson RE, Tatham RL, Black WC, 1998 Multivariate Data Analysis with Readings, 5th ed. Prentice-Hall, Englewood Cliffs.

Harvey AG, Soehner AM, Kaplan KA, Hein K, Lee J, Kanady J, Li D, Rabe-Hesketh S, Ketter TA, Neylan TC, Buysse DJ, 2015 Treating insomnia improves mood state, sleep, and functioning in bipolar disorder: A pilot randomized controlled trial. Journal of Consulting and Clinical Psychology 83 (3), 564–577. [PubMed: 25622197]

Hershner SD, Chervin RD, 2014 Causes and consequences of sleepiness among college students. Nat Sci Sleep 6, 73–84. [PubMed: 25018659]

Hu LT, Bentler PM, 1995 Evaluating model fit, in: Hoyle RH (Ed.), Structural Equation Modeling: Concepts, Issues and Applications. Sage, Thousand Oaks, pp. 76–99.

Johns MW, 1991 A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 14, 540–545. [PubMed: 1798888]

Johns MW, 1992 Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 15 (4), 376–381. [PubMed: 1519015]

Kanady JC, Harvey AG, 2015 Development and Validation of the Sleep Inertia Questionnaire (SIQ) and Assessment of Sleep Inertia in Analogue and Clinical Depression. Cognitive Therapy and Research 39 (5), 601–612. [PubMed: 26451063]

Kaplan KA, Harvey AG, 2009 Hypersomnia across mood disorders: A review and synthesis. Sleep Medicine Reviews 13, 275–285. [PubMed: 19269201]

Kaplan KA, McGlinchey EL, Soehner A, Gershon A, Talbot LS, Eidelman P, Gruber J, Harvey AG, 2015 Hypersomnia subtypes, sleep and relapse in bipolar disorder. Psychological Medicine 45 (8), 1751–1763. [PubMed: 25515854]

Kaplan KA, Talbot LS, Gruber J, Harvey AG, 2012 Evaluating sleep in bipolar disorder: comparison between actigraphy, polysomnography, and sleep diary. Bipolar Disorders 14 (8), 870–879. [PubMed: 23167935]

King AC, Oman RF, Brassington GS, Bliwise DL, Haskell WL, 1997 Moderate-intensity exercise and self-rated quality of sleep in older adults. A randomized controlled trial. Journal of the American Medical Association 277 (1), 32–37. [PubMed: 8980207]

Knutson KL, Ryden AM, Mander BA, Van Cauter E, 2006 Role of sleep duration and quality in the risk and severity of type 2 diabetes mellitus. Archives of Internal Medicine 166 (16), 1768–1774. [PubMed: 16983057]

Koffel E, Watson D, 2010 Development and initial validation of the Iowa Sleep Disturbances Inventory. Assessment 17 (4), 423–439. [PubMed: 20484713]

Liu X, Buysse DJ, Gentzler AL, Kiss E, Mayer L, Kapornai K, Vetro A, Kovacs M, 2007 Insomnia and hypersomnia associated with depressive phenomenology and comorbidity in childhood depression. Sleep 30 (1), 83–90. [PubMed: 17310868]

Lund HG, Reider BD, Whiting AB, Prichard JR, 2010 Sleep patterns and predictors of disturbed sleep in a large population of college students. Journal of Adolescent Health 46 (2), 124–132. [PubMed: 20113918]

McCall C, McCall WV, 2012 Comparison of actigraphy with polysomnography and sleep logs in depressed insomniacs. Journal of Sleep Research 21 (1), 122–127. [PubMed: 21447050]

Morin CM, Belleville G, Belanger L, Ivers H, 2011 The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 34 (5), 601–608. [PubMed: 21532953]

Ohayon MM, Dauvilliers Y, Reynolds CF 3rd, 2012 Operational definitions and algorithms for excessive sleepiness in the general population: implications for DSM-5 nosology. Archives of General Psychiatry 69 (1), 71–79. [PubMed: 22213791]

Plante DT, 2015 Hypersomnia in Mood Disorders: a Rapidly Changing Landscape. Curr Sleep Med Rep 1 (2), 122–130. [PubMed: 26258003]

Plante DT, Cook JD, Goldstein MR, 2017 Objective measures of sleep duration and continuity in major depressive disorder with comorbid hypersomnolence: a primary investigation with contiguous systematic review and meta-analysis. Journal of Sleep Research 26 (3), 255–265. [PubMed: 28145043]

Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH, 1996 The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychological Medicine 26, 477–486. [PubMed: 8733206]

Schmitt N, 1996 Uses and abuses of coefficient alpha. Psychological Assessment 8 (4), 350.

Sheehan DV, Harnett-Sheehan K, Raj BA, 1996 The measurement of disability. International Clinical Psychopharmacology 11(suppl 3), 89–95.

Soehner AM, Kaplan KA, Harvey AG, 2014 Prevalence and clinical correlates of co-occurring insomnia and hypersomnia symptoms in depression. Journal of Affective Disorders 167, 93–97. [PubMed: 24953480]

Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh DK, Lyon K, Smith PL, Schwartz AR, Redline S, Pack AI, Dinges DF, 1997 An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep 20 (10), 835–843. [PubMed: 9415942]

Worthington J, Fava M, Davidson K, Alpert JE, Nierenger A, Rosenbaum J, 1995 Patterns of improvement in depressive symptoms with fluoxetine treatment. Psychopharmacology Bulletin 31, 223–226. [PubMed: 7491372]

Young R, Biggs J, Ziegler V, Meyer D, 1978 A rating scale for mania: reliability, validity and sensitivity British Journal of Psychiatry 133, 429–425. [PubMed: 728692]

Zimmerman M, McGlinchey JB, Posternak MA, Friedman M, Boerescu D, Attiullah N, 2005 Differences between minimally depressed patients who do and do not consider themselves to be in remission. Journal of Clinical Psychiatry 66, 1134–1138. [PubMed: 16187771]

Highlights

- Hypersomnia is common across mood disorders and associated with significant illness burden
- We validated a self-report instrument to measure its severity, distress and impairment
- The Hypersomnia Severity Index showed good convergent validity against existing sleep measures
- Favorable construct validity was established against two weeks of sleep diaries and actigraphy

Table 1.

Demographics and Depression Symptoms in Scale Development (Undergraduate) and Scale Validation (Psychiatric) Samples

	Undergraduate (N=381)	Bipolar Disorder (N=89)	Major Depressive Disorder (N=21)
Demographics, Mean (SD) or No. (%)			
Age	20.9 (3.0)	35.2 (11.1)	28.1 (5.9)
Gender, No. Female	194 (55.0)	55 (61.8)	17 (81.0)
Race			
African American	7 (1.8)	7 (8.0)	1 (5.0)
Asian	177 (46.7)	13 (14.9)	1 (5.0)
Caucasian	187 (49.3)	61 (70.1)	18 (90.0)
Other/Biracial	8 (2.1)	6 (6.9)	
Ethnicity, No. Hispanic	37 (9.9)	6 (6.9)	1 (4.7)
QIDS-SR Total Score	8.3 (7.0)	9.9 (5.2)	--
BDI-II	--	--	23.2 (6.2)

Note. BDI-II= Beck Depression Inventory, 2nd Ed; QIDS-SR= Quick Inventory of Depressive Symptomatology, Self-Report Version.

Table 2.

Self-Report Instruments and Prospective Sleep Measures in Scale Development (Undergraduate) and Scale Validation (Psychiatric) Samples

	Undergraduate (N=381)	BD (N=89)	MDD+HS (N=21)
Self-Report Measures, Mean (SD)			
Hypersomnia Severity Index	14.0 (6.1)$_b$	14.8 (7.1)$_b$	20.9 (4.3)$_a$
Hypersomnia Symptoms Subscale	7.6 (3.3)$_b$	6.5 (3.9)$_c$	10.8 (2.0)$_a$
Distress/Impairment Subscale	6.5 (3.5)$_b$	8.8 (4.3)$_a$	10.1 (2.7)$_a$
Epworth Sleepiness Scale	8.9 (4.5)$_b$	7.4 (4.6)$_c$	12.2 (2.6)$_a$
PSQI Daytime Dysfunction	0.4 (0.6)$_b$	1.4 (0.76)$_a$	1.7 (0.6)$_a$
PSQI Sleep Duration	403.2 (103.3)$_b$	419.4 (116.6)$_b$	508.6 (59.6)$_a$
Sheehan Disability Scale	10.6 (6.1)$_b$	14.1 (7.3)$_a$	--
Sleep Measures, Mean (SD)			
Diary Total Sleep Time, min.	--	440.9 (85.9)$_b$	495.3 (43.9)$_a$
Diary Time in Bed, min.	--	524.1 (88.9)$_a$	524.4 (45.0)$_a$
Actigraphy Total Sleep Time, min.	--	427.7 (82.3)$_a$	451.1 (47.6)$_a$
Actigraphy Time in Bed , min.	--	529.2 (82.3)$_a$	520.4 (45.4)$_a$

Note. BD= Bipolar Disorder; MDD+HS= Major Depressive Disorder and Hypersomnolence Disorder; PSQI= Pittsburg Sleep Quality Index. Means within a row not sharing the same subscript differ from one another at $p < 0.05$. The subscales presented reflect the final organization of items, see text.

Table 3.

Correlation between HSI and Self-Report Instruments in Scale Development (Undergraduate) and Scale Validation (Psychiatric) Samples

	HSI Total	HIS Hypersomnia Symptoms	HSI Distress/ Impairment
Scale Development Sample			
Epworth Sleepiness Scale	0.44**	0.48**	0.32**
PSQI Daytime Dysfunction	0.43**	0.40**	0.39**
PSQI Sleep Duration	−0.07	0.02	−0.12*
Sheehan Disability Scale	0.61**	0.48**	0.61**
Scale Validation Sample: BD			
Epworth Sleepiness Scale	0.38**	0.42**	0.33**
PSQI Daytime Dysfunction	0.29**	0.30**	0.27**
PSQI Sleep Duration	−0.06	0.09	−0.22*
Sheehan Disability Scale	0.31*	0.12	0.40**
Scale Validation Sample: MDD+HS			
Epworth Sleepiness Scale	−0.06	0.03	−0.11
PSQI Daytime Dysfunction	0.15	−0.03	0.25
PSQI Sleep Duration	0.19	0.17	0.18
Sheehan Disability Scale	--	--	--

Table 4.

Correlation between HSI and Prospective Sleep Measures in Scale Validation (Psychiatric) Samples

	HSI Total	HIS Hypersomnia Symptoms	HSI Distress/ Impairment
Scale Validation: BD			
Diary Total Sleep Time, min.	0.07	0.17	−0.12
Diary Time in Bed, min.	0.29**	0.29**	0.15
Actigraphy Total Sleep Time, min.	0.42**	0.27*	0.42**
Actigraphy Time in Bed, min.	0.35**	0.24	0.31*
Scale Validation: MDD+HS			
Diary Total Sleep Time, min.	0.04	0.09	−0.01
Diary Time in Bed, min.	0.14	0.17	0.10
Actigraphy Total Sleep Time, min.	0.35	0.26	0.36
Actigraphy Time in Bed, min.	0.32	0.27	0.31