BEHAVIORAL SLEEP MEDICINE, *3*(4), 227–246

Initial Validation of the SLEEP-50 Questionnaire

Victor I. Spoormaker
Department of Clinical Psychology
Utrecht University, The Netherlands

Ingrid Verbeek
Center for Sleep and Wake Disorders
Kempenhaeghe, XXcountryXX

Jan van den Bout
Department of Clinical Psychology
Utrecht University, The Netherlands

Ed C. Klip
Department of Medical Psychology
University of Groningen, The Netherlands

Initial psychometric properties of the SLEEP-50 questionnaire, designed to detect sleep disorders as listed in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., Text Revision), were examined. The sample consisted of 377 college students, 246 sleep patients, 32 nightmare sufferers, and 44 healthy volunteers. The internal consistency was high (Cronbach's $\alpha = .85$); test–retest correlations fell between .65 and .89. Principal component analysis with a direct oblimin rotation revealed a factor structure that closely matched the designed structure. Sensitivity and specificity scores were promising for all sleep disorders; the agreement between all clinical diagnoses and SLEEP-50-classifications was substantial ($\kappa = .77$). These initial findings indicate that the SLEEP-50 seems able to detect a variety of sleep disorders. The SLEEP-50 can aid in screening for common sleep disorders in the general population.

Requests for reprints should be sent to Victor I. Spoormaker, Department of Clinical Psychology, Utrecht University, P.O. Box 80.140, 3508 TC Utrecht, The Netherlands. E-mail: v.i.spoormaker@fss.uu.nl

Although highly prevalent, sleep disorders are generally underdiagnosed in the adult population (Kupperman et al., 1995). This can possibly be overcome by the use of short sleep questionnaires, which could work as a screening device for sleep disorders (Kapuniai, Andrew, Crowell, & Pearce, 1988; Roth et al., 2002).

Sleep questionnaires are often designed to measure different aspects of sleep: quality of sleep, sleepiness during the day, impact of sleep problems on daily functioning, one specific sleep disorder, or two or more sleep disorders.

Questionnaires that assess the quality of sleep, such as the Pittsburgh Sleep Quality Index (Buysse, Reynolds, Monk, Berman, & Kupfer, 1992), are not designed to detect one or more specific sleep disorders. The same applies to questionnaires that assess daytime sleepiness, such as the Epworth Sleepiness Scale (Johns, 1991) and the Stanford Sleepiness Scale (Hoddes, Zarcone, Smythe, Philips, & Dement, 1973). Questionnaires like the Insomnia Impact Scale (Hoellscher, Ware, & Bond, 1993) and the Quality of Life of Insomnia (De Sousa, 1996) are aimed at identifying the impact of sleep problems on daily functioning.

Other sleep questionnaires are directed at one sleep disorder. Sleep apnea can be assessed by the Survey Screen for Sleep Apnea (Maislin et al., 1995), the Sleep and Health Questionnaire (Kump et al., 1994), and the Hawaii Sleep Questionnaire (Kapuniai et al., 1988). Insomnia can be assessed by questionnaires such as Spielman's Insomnia Symptom Questionnaire (Spielman, Saskin, & Thorpe, 1987) and the Insomnia Severity Index (Bastien, Vallières, & Morin, 2001), although the latter scale is designed for outcome rather than prevalence studies.

The Berlin Questionnaire (Netzer, Stoohs, Netzer, Clark, & Strohl, 1999) and the Basic Nordic Sleep Questionnaire (Partinen & Gislason, 1995) can detect two groups of sleep problems, whereas the Sleep Disorders Questionnaire (SDQ; Douglass et al., 1994) is able to assess four groups of sleep disorders. Due to its length, however, the latter scale is not a practical instrument to screen for sleep disorders in the general population. The Dutch version of the SDQ (Sweere et al., 1998) is somewhat shorter (and the only validated sleep questionnaire in the Netherlands) but this questionnaire has other limitations, such as unknown reliability coefficients and a correct prediction for less than a third of the healthy participants.

Recently the Global Sleep Assessment Questionnaire (GASQ; Roth et al., 2002) has been validated. It proved to measure and predict the most common sleep disorders, such as sleep apnea, insomnia, insomnia associated with a mental disorder, restless legs, periodic limb movement disorder (PLMD), and parasomnias. However, there were few healthy participants in this study, so conclusions about its ability to correctly predict persons without any sleep disorder are still preliminary. Moreover, narcolepsy was not included.

At this time there is no short sleep questionnaire that can adequately predict the sleep disorders as listed in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., Text Revision [DSM–IV–TR]; American Psychiatric Association,

2000) and that can effectively distinguish these sleep disorders from sleep complaints. Therefore the SLEEP-50 was developed.

METHOD

Instrument

The SLEEP-50 is a self-administered questionnaire (see the Appendix) about the intensity of a person's subjective sleep complaints. It was designed to detect both the sleep complaints and the sleep disorders as listed in the *DSM–IV–TR,* as well as factors influencing sleep. The first version of the SLEEP-50 was designed by the first author and followed the descriptions and criteria from the *DSM–IV,* leading to nine subscales.

Subscales. The subscales for the SLEEP-50 are Sleep Apnea (Items 1–8), Insomnia (Items 9–16), Narcolepsy (Items 17–21), Restless Legs/PLMD (Items 22–25), Circadian Rhythm Sleep Disorder (Items 26–28), Sleepwalking (Items 29–31), Nightmares (Items 32–36), Factors Influencing Sleep (Items 37–43), and the Impact of Sleep Complaints on Daily Functioning (Items 44–50).

Note that Item 32 checks whether frightening dreams are present. If not, persons should not fill out Items 33 to 36. Item 33 is required for checking the *DSM–IV* definition of nightmares, where waking up is a necessary criterion. Items 34 and 35 are necessary to distinguish nightmares (vivid memory and quick orientation) from night terrors (amnesia and slow orientation), whereas Item 36 could also aid in distinguishing these two parasomnias.

The Impact subscale was necessary for all the diagnoses of sleep disorders because the first seven subscales ask about sleep complaints only. According to the *DSM–IV* a sleep disorder, like any other mental disorder, can only be diagnosed if there are significant impairments in daily functioning. The SLEEP-50 checks for sleep complaints with the subscales of Items 1 to 36 and detects a sleep disorder with the Impact subscale.

Time frame, item format, and interpretation guidelines. The questionnaire starts with this statement: "Please respond to what extent a statement (item) has been applicable to you during the past 4 weeks." Each item is scored on a 4-point-scale: 1 (*not at all*), 2 (*somewhat*), 3 (*rather much*), and 4 (*very much*). This intensity scale was preferred over a frequency scale, because several items (e.g., Items 1, 3, and 4) could provide invalid answers with a frequency scale. For example, people that snore may not adequately respond to Item 1 ("I am told that I snore") on a frequency scale because it need not be said to them several times a

week or month. Yet, if a person has been told that he or she snored only once or twice during the last 4 weeks, he or she can judge that item to be very applicable.

Moreover, a sleep questionnaire is not an objective way of assessing sleep complaints, and therefore the (subjective) judgment of the participants is important as well. One can have only one nightmare a month that is highly distressing or four nightmares a month that are not distressing. Additionally, the frequency may very well differ from the intensity for several items; for example, falling asleep on a social occasion twice a month can be very distressing, whereas waking up with a dry mouth twice a month might not be equally distressing. With a frequency scale the scores on these two items would not differ because they are equally frequent, thus complicating the interpretation guidelines.

As a quick check, a score of 3 (*rather much*) or 4 (*very much*) on an item would indicate the presence of a sleep symptom of a specific sleep disorder, a procedure also used with other questionnaires with the same answering format (Hovens, Bramsen, & van der Ploeg, 2000). To check for the presence of a sleep disorder, at least one item on the Impact scale should also be endorsed with a score of 3 or 4.

Moreover, scores on the items could be summed as well, leading to a total score for each subscale. With no items endorsed a subscale would get the minimum score of amount of items × 1; with all items endorsed a subscale would get the maximum score of amount of items × 4. The sum of a specific subscale then determines the final prediction of whether a certain sleep disorder is present (depending on the optimal cutoff value). Note that for diagnosing a sleep disorder not only the specific subscale (e.g., Insomnia) needs to exceed a certain cutoff point, but so does the Impact subscale. If the score on the Impact subscale is below that cutoff value, no sleep disorder can be diagnosed (sleep complaints are present without significant impairments in daily functioning).

Initial testing. The initial testing started with contacting five Dutch sleep researchers and clinicians who were asked to comment on the SLEEP-50. This led to the inclusion of three items (22, 42, and 43) and a revision of Item 20 ("with intense emotions" was added to make it more recognizable for persons with narcolepsy).

Afterward, the questionnaire was filled out by two groups: 56 college students (bachelor's in psychology) who were contacted via two statistics classes and 30 participants with various confirmed sleep disorders who where selected via snowball sampling (a sampling method that consists of identifying participants who are then used to refer researchers to other participants). Both groups could comment on the questionnaire in an informal setting or by e-mail. This led to the rephrasing of several items (e.g., for Items 1, 3, and 4 "I am told" rather than "My bed partner has told me" as not everyone had a bed partner) and the exclusion of one item (everyone scored high on "drinking tea or coffee during the evening" as this is a very

common habit in the Netherlands).[1] Moreover, the last two items were renamed as additional questions to clarify the distinction in answer format. Also, Item 32 included a second answer format where participants could estimate the amount of frightening dreams a week. This estimation was excluded because it led to some confusion and because the correlation between this estimation and the answer on Item 32 was very high, $r(84) = .94$, $p < .001$.

After analyzing the results of the sleep patients it was found that the sum of the scores on a particular sleep subscale combined with the sum of the scores on the Impact subscale resulted in better predictions than the quick check (number of items endorsed on a specific sleep subscale and on the Impact subscale). With the summed scores, not only the endorsement of an item counted, but also the relative intensity. The quick check yielded no better predictions when a different score was used for the endorsement of an item (e.g., 2, 3, or 4 vs. 1 or score 4 vs. 1, 2, or 3).

Participants and Procedures

College students. As part of determining the internal consistency and construct validity, 500 questionnaires were handed out to psychology students (no inclusion criteria) who were approached via lectures in three different bachelor-level courses at Utrecht University. In a 10-min speech the purpose and relevance of the study were explained ("Sleep disorders, although highly prevalent, are not often detected. A questionnaire that is able to screen for sleep disorders might be helpful in recognizing sleep disorders. This study is about how well this sleep questionnaire detects sleep disorders"). Afterward, the SLEEP-50 was handed out. Respondents were asked to send it back together with the written consent. Of the 500 students who received a questionnaire, 336 returned the SLEEP-50, for a response rate of 67% (*M* age = 22.1 years, *SD* = 4.2; 56% female). This response rate is moderate, but students did not receive any incentive.

For determining the test–retest reliability, the same procedure was followed during another lecture with master's students, except that after 3 weeks the students were asked to fill out the same questionnaire. Fifty questionnaires were handed out; these students received a small monetary incentive (10 euro). The response rate was higher (41 filled out the SLEEP-50 twice for a response rate of 83%; *M* age = 22.3 years, *SD* = 3.4; 85% female).

Sleep patients. Three-hundred consecutive patients were approached during intake at the Center for Sleep and Wake Disorders Kempenhaeghe at Heeze, the Netherlands (no inclusion criteria; all patients were approached). They re-

[1]Drinking tea or coffee in the evening is so common in the Netherlands that it was not informative for this sample, but it might be important in other cultures.

ceived the SLEEP-50 by mail together with a letter stating the purpose and the relevance of the study. They were asked to return the sleep questionnaire and the written consent at the sleep clinic. Two-hundred and fifty-two participants returned the SLEEP-50. As 6 persons dropped out during intake at the sleep clinic, 246 useful questionnaires were returned for a response rate of 82.3% (*M* age = 47.6 years, *SD* = 12.2; 44% female).

Polysomnographic measures were conducted at the sleep center of Kempenhaeghe, where polysomnography is not indicated for the routine evaluation of insomnia. Insomnia is diagnosed primarily with a detailed medical, psychiatric, and sleep history (the latter one measured with a sleep log and unstructured sleep questionnaires). Polysomnography is indicated for insomnia when a comorbid sleep disorder or sleep state misperception (SSM) is suspected or the initial diagnosis is uncertain. Of the 65 sleep patients with insomnia as the primary diagnosis, 42 did not receive polysomnography. Of the 30 sleep patients with an affective disorder as the primary diagnosis, 19 did not receive polysomnography. All other sleep patients did (n = 185).

Nightmare sufferers. Another group consisted of nightmare sufferers who had experienced nightmares (*DSM–IV–TR* definition) for more than a year. The rationale for including a group with this specific sleep disorder is twofold. Although nightmares are prevalent in the general population (estimated at around 3–7%; Hublin, Kaprio, Partinen, & Koskenvuo, 1999; Ohayon, Guilleminault, & Caulet, 1996), nightmare sufferers do not often seek help. Of all sleep patients in this study, none had a nightmare disorder as the primary diagnosis. Moreover, research indicates that polysomnographic recordings tend to decrease nightmare frequency (Fisher, Byrne, Edwards, & Kahn, 1970) and are not the method of choice for assessing nightmares (Spoormaker et al., in press).

Thirty-eight nightmare sufferers who volunteered for a treatment study on nightmares were asked to participate in this study. They received the SLEEP-50 by mail together with a letter stating the purpose and the relevance of this study. They were asked to return the questionnaire with the written consent and to schedule an appointment with the first author for a 1-hr unstructured interview about their sleep (e.g., complaints, hygiene, and daily functioning). They would receive a monetary incentive (25 euro). Thirty-two nightmare sufferers returned the SLEEP-50 and came to their appointment for a response rate of 89% (*M* age = 25.8 years, *SD* = 6.2; 84% female). The interview took place before any treatment. At the interview, all 32 participants reported suffering from nightmares and daily functioning limitations. The high response rate can be explained by the incentive and participants' motivation to do anything possible to overcome their nightmares.

Healthy volunteers. The healthy volunteers were recruited via an advertisement in a local newspaper that said that participants who were satisfied with their

sleep could participate in a study on sleep. They could register by telephone or e-mail. A total of 178 registered (58 by telephone, 120 via e-mail). One hundred participants were randomly selected and received the SLEEP-50 by mail together with a letter stating the purpose and the relevance of the study. They were asked to return the questionnaire with the written consent and to schedule an appointment with the first author for a 1-hr unstructured interview about their sleep (e.g., complaints, hygiene, and daily functioning). They would receive a monetary incentive as well (25 euro). Forty-four healthy volunteers returned the questionnaire and came to their appointment for a response rate of 44% (*M* age = 41.4 years, *SD* = 14.5; 55.4% female). The effort that was asked of participants (and the possible lack of motivation) may explain the low response rate.

The 1-hr appointment was necessary for minimizing the chance that any of these healthy volunteers had a sleep disorder, although diagnoses from the sleep clinic would have been more informative. Yet, none of the participants reported any sleep complaint, inadequate sleep hygiene, or problems with daily functioning during the interview. These data were important to check whether the SLEEP-50 can adequately distinguish people with a sleep disorder from people with no sleep disorder, next to adequately predicting which disorder is present.

Statistical Analyses

The internal consistency was measured by Cronbach's alpha (all participants including the test–retest group's first assessment); for the test–retest reliability Pearson correlation coefficients (*r* values) were determined for all subscales and the total score of the SLEEP-50. These *r* values were tested for significance with Pearson's *r* test (one-tailed).

Principal component analysis with a direct oblimin rotation was used for the construct validity analysis (all participants including the test–retest group's first assessment). We preferred this oblique rotation to an orthogonal rotation (e.g., varimax) because several factors could be interrelated. For example, the impact of sleep complaints was expected be related to various other factors (e.g., insomnia). However, an oblique rotation could also result in independent factors if that provides a better fit. Note that the factors influencing sleep (Items 37–43) were excluded because they were not theoretically related to one factor or another subscale. Moreover, the nightmare items (Items 33–36) were only filled out by a minority (participants with nightmares) and therefore were also excluded.

To investigate the predictive validity, sensitivity (proportion correctly predicted with a sleep disorder) and specificity scores (proportion correctly predicted without that sleep disorder) were established with optimized cutoff scores. The starting point for a specific sleep disorder was the mean of the relevant subscale (e.g., the mean score of sleep apnea patients on the SLEEP-50 Apnea subscale). All values lower than this mean were analyzed on the sensitivity and specificity of that partic-

ular value. The value where both the sensitivity and specificity were highest would be chosen as the optimal cutoff point. For example, for classifying sleep apnea, cutoff values for the Apnea subscale were (a) ≥ 14 (sensitivity .89, specificity .82), (b) ≥ 15 (sensitivity .85, specificity .88), and (c) ≥ 16 (sensitivity .79, specificity .90). Here the second option was chosen as the optimal cutoff value.

Kappa was computed to evaluate the agreement between the primary clinical diagnoses and the primary classifications of the SLEEP-50.

RESULTS

Reliability

The internal consistency for the entire scale—minus additional nightmare Items 33 to 36 and factors influencing sleep in Items 37 to 43—was high (Cronbach's α = .85; see Table 1). Deletion of any item did not increase or decrease the alpha with more than .02. Alphas were low for several subscales due to the small amount of items, except for sleepwalking, as these items were rarely endorsed.

The general test–retest reliability, tested in the student sample with an interval of 3 weeks, was good: $r(39) = .78$, $p < .001$. Adequate scores were found for all designed subscales except for sleepwalking. Closer examination revealed that there was no variance in the scores on Item 30 (evidence of action performed during the night); no participant endorsed this item. Exclusion of this item increased the test–retest correlation of the Sleepwalking scale, $r(39) = .65$, $p < .001$.

Not all items had a good item-total correlation (see Table 2). Only the insomnia and impact items showed relatively high item-total correlations, whereas about half of apnea, narcolepsy, and restless legs items showed an item-total correlation of greater than .20. This is, however, not problematic because the questionnaire was designed to detect and distinguish different sleep disorders. The scores on the subscales are important, not the total score (e.g., endorsing a sleepwalking item does not necessarily mean that one endorses items measuring other sleep complaints).

Construct Validity

A principal component analysis with a direct oblimin rotation (Table 2) revealed 10 factors that were able to explain 67.5% of the variance. The factors did not correlate more than .30 with another, and most correlations fell between .00 and .20.

The factor structure fits the originally designed structure accurately. Only the Narcolepsy subscale is problematic, with Item 19 (sleep attacks) loading on the factor impact of sleep complaints, and Item 17 (hypnagogic hallucinations) loading on both narcolepsy and nightmares, although the latter loading was somewhat

TABLE 1
Test–Retest Reliabilities (n = 41) and Cronbach's Alphas (N = 699) for the SLEEP-50 Subscales and the Total Score

	Apnea	*Insomnia*	*Narcolepsy*	*RLS/PLMD*	*Circadian Rhythm*	*Sleepwalking*	*Nightmares*	*Factors*	*Impact*	*Total*
Items	*1–8*	*9–16*	*17–21*	*22–25*	*26–28*	*29–31*	*32*	*37–43*	*44–50*	*1–32, 44–50*
Cronbach's α	.51	.85	.52	.70	.47	.84	—	—	.86	.85
Pearson's r	.81*	.77*	.71*	.74*	.81*	–.07	.89*	.73*	.76*	.78*

Note. Cronbach's alpha was not computed for the Nightmares subscale as only Item 32 of this subscale was filled out by all participants, and Items 33 to 36 were only filled out by participants who endorsed Item 32. Cronbach's alpha was not computed for the Factors subscale as this subscale did not consist of theoretically related items. RLS/PLMD = restless legs/periodic limb movement disorder.

*$p < .001$.

TABLE 2
Pattern Matrix of the Item Loadings on Various Factors—Principal Component Analysis With Direct Oblimin Rotation ($N = 699$); Factor Loadings ≥ .30

	Insomnia	*Impact*	*Apnea I Breathing*	*Sleepwalking*	*RLS/ PLMD*	*Circadian Rhythm*	*Narcolepsy*	*Apnea II Mouth*	*Nightmares*	*Apnea III Sweating*	*Corrected Item-Total Correlation*
Eigenvalue	7.06	3.80	2.72	2.18	1.88	1.65	1.48	1.29	1.15	1.06	
Apnea											
1. Snoring			.38								.13
2. Sweating										.76	.14
3. Holding breath			.70								.07
4. Waking up gasping for air			.84								.27
5. Dry mouth								.75			.15
6. Waking up short of breath			.63								.32
7. Sour taste								.73			.39
8. Headache			.48								.26
Insomnia											
9. Difficulty falling asleep	.54										.58
10. Disturbing thoughts	.43								.37		.62
11. Worrying/unable to relax	.59								.34		.53
12. Waking up	.84										.60
13. Not being able to sleep again	.92										.54
14. Difficulty continuing sleep	.90										.32
15. Sleeping lightly	.75										.55
16. Sleeping too little	.79										.53
Narcolepsy											
17. Hypnagogic hallucinations							.56		.44		.39
18. Sleeping on social occasions							.65				.13
19. Sleep attacks		*.70*									.19
20. Cataplexy			.32				.47				.50

21. Sleep paralysis						.66			.21
RLS/PLMD									
22. Kicking legs while sleeping				.43					.09
23. Cramp or pain in legs				.86					.13
24. Shocks in legs				.65					.25
25. Difficulty keeping legs still				.80				.42	.27
Circadian rhythm									
26. Wanting a different rhythm					.58				.12
27. Going to bed at different times					.56				.20
28. Doing shift work					.66				.21
Sleepwalking									
29. Sleepwalking			.86						–.13
30. Waking up elsewhere			.87						–.07
31. Proof of nightly action			.86						.00
Nightmares									
32. Frightening dreams							.87		.19
Impact of sleep complaints									
44. Tired at getting up		.78							.50
45. Difficulty staying alert		.50							.52
46. Wanting more energy		.78							.59
47. Easily irritated		.54							.39
48. Difficulty concentrating		.50							.54
49. Worrying about sleep	.31	.57			.35				.56
50. Sleeping badly in general	*.78*								.66

Note. Italicized factor loadings consider items originally designed for another subscale. RLS/PLMD = restless legs/periodic limb movement disorder.

lower. It was also surprising that Insomnia Items 10 and 11 (disturbing thoughts and worrying or unable to relax) also loaded on this nightmare factor. Negative thoughts and anxiety or tension may induce nightmares, although Item 49 (worrying about sleep) did not load on the nightmare factor. Item 49 loaded on circadian rhythm, insomnia, and impact, but highest on the last factor. Item 50 (sleeping badly in general) loaded on the insomnia factor but not on impact of sleep complaints. Subsequently, Item 19 was included on the Impact scale, and Item 50 was included on the Insomnia scale. For Item 17 the factor with the highest loading was chosen: narcolepsy.

Furthermore it is worth noting that the Apnea subscale was split into three factors, which did not correlate more than .20 with one another. One factor consisted of breathing problems (headache at waking up included), one consisted of a dry or sour mouth at waking, and the third consisted only of Apnea Item 2 (sweating) together with a loading of Item 25 (difficulty keeping legs still). Although Item 2 can still be seen as an Apnea item, it may focus on more than sleep apnea alone. Of 52 women aged 45 to 55, 39 endorsed this item with a score of 2 or higher, indicating that it may have measured a menopausal sleep complaint as well.

Predictive Validity

Table 3 shows that participants with a specific sleep disorder scored highest on the SLEEP-50 subscale designed to measure that sleep disorder. However, three distinct diagnoses were not specifically measured by the SLEEP-50: hypersomnia, affective disorder, and SSM. Table 3 shows that the 3 participants with hypersomnia scored low to medium on all subscales except on the Impact subscale. Moreover, the 30 participants with an affective disorder did not score highest on any of the SLEEP-50 subscales, so the amount and intensity of sleep complaints could not predict this diagnosis. However, Items 10 (disturbing *thoughts*), 11 (worrying and unable to relax), 42 (feeling sad), and 43 (no interest in daily occupations) could, as participants with affective disorders scored significantly higher on these four items than participants with insomnia, $t(93) = 2.1$, $p < .05$, and participants with other sleep disorders, $t(276) = 3.8$, $p < .001$.

Participants with SSM scored highest on the Insomnia subscale, although not significantly higher than participants with insomnia, $t(71) = 1.5$, $p > .10$. These two groups could be distinguished by the two additional Items A (rating of own sleep) and B (amount of hours slept), as participants with SSM reported a lower amount of hours slept, $t(71) = 3.9$, $p < .001$, and rated their sleep lower, $t(71) = 2.7$, $p < .01$, than participants with insomnia. All 8 participants with SSM reported sleeping less than 5 hr a night, and 7 reported sleeping less than 4 hr.

A sleep disorder was present if the score on the Impact subscale was 15 or higher (see Table 4). For the optimal cutoff points the value was taken where both the sensitivity and specificity were highest. Optimal sensitivity and specificity

TABLE 3
SLEEP-50 Subscale Scores for Sleep Patients, Nightmare Sufferers, and Healthy Participants

		Apnea[a]		*Insomnia*[b]		*Narcolepsy*[c]		*RLS/PLMD*[c]		*Circadian Rhythm*[d]		*Sleepwalking*[d]		*Nightmares*[e]		*Impact*[f]	
Sleep Disorder	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Apnea	81	16.9	3.1	14.0	4.5	4.8	1.3	6.2	2.3	5.6	2.4	3.1	0.2	1.3	0.7	18.9	5.1
Insomnia	65	11.4	2.6	23.0	4.9	4.8	1.7	5.2	1.7	7.5	1.8	3.0	0.0	1.3	0.7	19.2	3.4
Affective disorder	30	12.6	3.5	19.3	4.9	5.6	2.4	5.3	1.6	7.4	2.5	3.2	0.5	1.7	1.1	19.4	4.4
SSM	8	10.0	1.6	25.8	6.2	5.3	1.9	6.0	1.8	7.4	0.5	3.3	0.4	1.5	0.5	19.5	2.9
Narcolepsy	3	12.3	2.5	11.7	1.5	6.7	0.6	6.0	0.5	6.7	1.7	3.0	0.0	1.0	0.0	17.7	1.2
RLS/PLMD	29	12.1	3.2	18.7	6.4	4.6	1.0	9.1	2.3	5.9	2.9	3.1	0.3	1.1	0.2	16.4	4.8
Circadian rhythm	23	12.3	2.2	15.5	4.5	5.1	1.9	6.5	2.9	9.7	2.8	4.1	2.3	1.0	0.0	17.2	3.5
Sleepwalking	4	10.4	1.4	11.5	2.8	5.0	1.4	5.0	1.4	6.8	1.7	10.5	2.1	2.0	1.4	15.3	2.4
Nightmares	32	11.7	3.4	17.9	5.1	6.0	2.1	6.5	1.5	5.7	1.5	3.0	0.2	3.2	0.4	16.6	4.7
Hypersomnia	3	10.7	1.8	16.7	3.2	5.3	1.3	5.7	1.9	5.3	2.2	3.0	0.0	1.3	0.2	20.3	3.9
Healthy	44	9.9	1.5	10.7	1.5	4.8	0.6	4.6	0.8	4.0	1.3	3.0	0.0	1.2	0.6	8.4	1.3

Note. RLS/PLMD = restless legs/periodic limb movement disorder; SSM = sleep state misperception.
[a]Range = 8–32. [b]Range = 9–36. [c]Range = 4–16. [d]Range = 3–12. [e]Range = 1–4. [f]Range = 7–28.

TABLE 4
Optimal Cutoff Values and Scoring Procedures

Sleep Disorder	*Optimal Cutoff Value*	*Sensitivity*	*Specificity*
Apnea	≥ 15 on apnea	0.85	0.88
Insomnia	≥ 19 on insomnia	0.71	0.75
Affective disorder	≥ 12 on Items 10, 11, 43, and 44	0.77	0.73
SSM	≥ 19 on insomnia and estimated amount of hours slept < 4	0.88	0.92
Narcolepsy	≥ 7 on narcolepsy	0.67	0.86
RLS/PLMD	≥ 7 on RLS/PLMD[a]	0.83	0.72
Circadian rhythm	≥ 8 on circadian rhythm	0.83	0.69
Sleepwalking	≥ 7 on sleepwalking	1.00	1.00
Nightmares	≥ 3 on Item 32 and ≥ 9 on Items 33–35	0.84	0.77
Hypersomnia	None of the above and ≥ 15 on impact	1.00	0.79
All sleep disorders	≥ 15 on impact	0.84	0.77

Note. SSM = sleep state misperception; RLS/PLMD = restless legs/periodic limb movement disorder.

[a]≥ 8 on RLS/PLMD showed a sensitivity of 0.72 and a specificity of 0.8.

scores were lowest for insomnia and highest for sleepwalking. For restless legs and PLMD, two optimal cutoff scores were found, one where the sensitivity was higher (≥ 7) and one where the specificity was higher (≥ 8).

Finally, the primary clinical diagnoses were compared to the primary classifications of the SLEEP-50 (Table 5). The SLEEP-50 correctly predicted 80% of all participants. The kappa for the entire scale (the measure of agreement controlled for chance) was .77, indicating a substantial agreement between the clinical diagnoses and the SLEEP-50 classifications. Distinguishing insomnia from affective disorders was difficult, as 11 participants with insomnia were predicted to have an affective disorder, and 5 participants with an affective disorder were predicted to have insomnia. It was surprising that almost all of the incorrectly classified participants with apnea were predicted to be healthy. Moreover, 10 of the healthy participants were predicted to have a sleep disorder, although 3 of them were predicted to have an affective disorder.

DISCUSSION

Before discussing the results, several limitations should be mentioned. Test–retest reliabilities were obtained in a small and healthy student sample. The sample sizes for SSM, narcolepsy, hypersomnia, and sleepwalking were too low for valid conclusions considering the predictive validity. Moreover, the majority of insomnia patients did not receive polysomnography. The same applies to healthy partici-

TABLE 5
Number of Correctly (Italicized) and Incorrectly Classified Participants

	Clinical Diagnosis										
SLEEP-50 Classification	*Apnea*	*Insomnia*	*Affective*	*SSM*	*RLS/ PLMD*	*Circadian*	*Nightmares*	*Narcolepsy*	*Sleepwalking*	*Hypersomnia*	*Healthy*
Apnea	*69*	1	1	0	1	0	0	0	0	0	2
Insomnia	1	*46*	5	1	3	1	1	0	0	0	4
Affective disorder	0	11	*23*	0	0	1	1	0	0	0	3
SSM	0	4	0	*7*	0	0	0	0	0	0	0
RLS/PLMD	0	0	0	0	*24*	2	0	0	0	0	0
Circadian rhythm	0	2	1	0	1	*19*	0	0	0	0	0
Nightmares	0	1	0	0	0	0	*27*	0	0	0	0
Narcolepsy	0	0	0	0	0	0	3	*2*	0	0	1
Sleepwalking	0	0	0	0	0	0	0	0	*4*	0	0
Hypersomnia	0	0	0	0	0	0	0	0	0	*3*	0
Healthy	11	0	0	0	0	0	0	1	0	0	*34*
Total	81	65	30	8	29	23	32	3	4	3	44

Note. SSM = sleep state misperception; RLS/PLMD = restless legs/periodic limb movement disorder.

pants and nightmare sufferers: Their diagnoses were not obtained in the sleep center. It is possible that some of the healthy participants actually had a sleep disorder. However, if that were so, the kappa would probably become higher (10 of the healthy participants were incorrectly classified with a disorder, yet this could have been the correct prediction for some of them). In addition, polysomnography is not the method of choice for diagnosing nightmares because nightmares tend to occur less often in the sleep laboratory (Fisher et al., 1970).

Yet, this initial validation study showed promising results. The SLEEP-50 had a high internal consistency and good test–retest reliabilities. A factor structure that closely resembled the originally designed structure indicated acceptable construct validity. Sensitivities and specificities were promising for all sleep disorders. Moreover, the agreement between SLEEP-50 classifications and clinical diagnoses was substantial.

Whereas the predictive validity of the SLEEP-50 for SSM, narcolepsy, sleepwalking, and hypersomnia is promising but very preliminary, the SLEEP-50 seems able to adequately predict sleep apnea, insomnia, restless legs or PLMD, circadian rhythm sleep disorder, and nightmares. In addition, the majority of healthy participants were classified correctly. In other studies on global sleep disorder questionnaires with comparable sensitivity scores there were either very few healthy participants (Roth et al., 2002) or many healthy participants who were incorrectly classified with a sleep disorder (Sweere et al., 1998).

Insomnia was hardest to predict, especially because it was difficult to distinguish insomnia from affective disorders. Because the SLEEP-50 focuses on sleep problems and not on affective complaints, additional questionnaires should be used to check for comorbid mental complaints or disorders.

For restless legs and PLMD two cutoff values with different specificities were provided. It may be important that a sleep disorder like restless legs and PLMD or sleep apnea uses a cutoff point with a high specificity (i.e., low number of false positives) because this classification would typically result in polysomnography or another costly laboratory test (Roth et al., 2002). This does not apply to insomnia, where the specificity need not necessarily be higher than the sensitivity, because persons with insomnia are more likely to be treated without further costly testing (Roth et al., 2002).

In addition, the results suggest that the SLEEP-50 may be able to detect SSM. Both the reported amount of hours slept and the rating of the sleep was significantly lower in participants with SSM than in participants with insomnia. As these findings can aid in recognizing SSM, future research needs to examine these findings in a larger sample.

In conclusion, these initial findings suggest that the SLEEP-50 is able to detect the most prevalent *DSM–IV–TR* sleep disorders: sleep apnea, insomnia, restless legs and PLMD, circadian rhythm sleep disorder, and nightmares. The SLEEP-50 is a practical global sleep questionnaire addressing the intensity of sleep com-

plaints, whereas another practical global sleep questionnaire (the GASQ; Roth et al., 2002) addresses the frequency of sleep complaints. The findings indicate that the SLEEP-50 can distinguish sleep complaints from sleep disorders and that the SLEEP-50 may be able to detect less common sleep disorders as well. The SLEEP-50 can aid in recognizing sleep disorders in the general population.

ACKNOWLEDGMENTS

This research was supported by Grant XXXXXXXX from the Research Institute of Psychology and Health.

We thank Maggie Stroebe for her help with preparing the article and Nijs Lagerweij for his statistical advice.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.

Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine, 2,* 297–307.

Baumel, M. J., Maislin, G., & Pack, A. I. (1997). Population and occupational screening for obstructive sleep apnea: Are we there yet? *American Journal of Respiratory and Critical Care Medicine, 155,* 9–14.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28,* 193–213.

De Sousa, J. C. R. P. (1996). Quality of life in insomnia in university psychology students. *Human Psychopharmacology, 11,* 169–184.

Douglass, A. B., Bornstein, R., Nino-Murcia, G., Keenan, S., Miles, L., Zarcone, V. P., et al. (1994). The Sleep Disorders Questionnaire I: Creation and multivariate structure of SDQ. *Sleep, 17,* 160–167.

Fisher, C., Byrne, J., Edwards, A., & Kahn, E. (1970). A psychophysiological study of nightmares. *Journal of the American Psychoanalytic Association, 18,* 747–782.

Hoddes, E., Zarcone, V., Smythe, H., Philips, R., & Dement, W. C. (1973). Quantification of sleepiness: A new approach. *Psychophysiology, 10,* 431–436.

Hoelscher, T. J., Ware, J. C., & Bond, T. (1993). Initial validation of the Insomnia Impact Scale. *Sleep Research, 22,* 149.

Hovens, J. E., Bramsen, I., & van der Ploeg, H. M. (2000). *Manual for the self-rating inventory for posttraumatic stress disorder (SRIP).* Lisse, Netherlands: Swets Test Publishers.

Hublin, C., Kaprio, J., Partinen, M., & Koskenvuo, M. (1999). Nightmares: Familial aggregation and association with psychiatric disorders in a nationwide twin cohort. *American Journal of Medical Genetics, 88,* 329–336.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. *Sleep, 14,* 540–545.

Kapuniai, L. E., Andrew, D. J., Crowell, D. H., & Pearce, J. W. (1988). Identifying sleep apnea from self reports. *Sleep, 11,* 430–436.

Kump, K., Whalen, C., Tishler, P. V., Browndler, I., Ferrette, V., Strohl, K. P., et al. (1994). Assessment of the validity and utility of a sleep-symptom questionnaire. *American Journal of Respiratory and Critical Care Medicine, 150,* 735–741.

Kuppermann, A., Lubeck, D. P., Mazonson, P. D., Patrick, D. L., Stewart, A. L., Buesching, D. P., et al. (1995). Sleep problems and their correlates in a working population. *Journal of General Internal Medicine, 10,* 25–35.

Maislin, G., Pack, A. L., Kribbs, N. B, Schwartz, P. L., Schwartz, A. R., Kline, L. R., et al. (1995). A survey screen for prediction of apnea. *Sleep, 18,* 158–166.

Netzer, N. C., Stoohs, R. A., Netzer, C. M., Clark, K., & Strohl, K. P. (1999). Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. *Annals of Internal Medicine, 131,* 185–536.

Ohayon, M. M., Priest, R. G., Guilleminault, C., & Caulet, M. (1996). Nightmares: Their relationships with mental disorders and sleep disorders. *European Neuropsychopharmacology, 6*(Suppl. 3), 136–137.

Partinen, M., & Gislason, T. (1995). Basic Nordic Sleep Questionnaire (BNSQ): A quantified measure of subjective sleep complaints. *Journal of Sleep Research, 4,* 150–155.

Roth, T., Zammit, G., Kushida, C., Doghramji, K., Mathias, S. D., Wong, J. M., et al. (2002). A new questionnaire to detect sleep disorders. *Sleep Medicine, 3,* 99–108.

Spielman, A. J., Saskin, P., & Thorpy, M. J. (1987). Treatment of chronic insomnia by restriction of time in bed. *Sleep, 10,* 45–56.

Sweere, Y., Kerkhof, G. A., Weerd, A. W. de, Kamphuisen, H. A. C., Kemp, B., & Schimsheimer, R. J. (1998). The validity of the Dutch Sleep Disorders Questionnaire (SDQ). *Journal of Psychosomatic Research, 45,* 549–555.