

Reliability and Efficacy of the Epworth Sleepiness Scale: Is There Still a Place for It?

Matthew T Scharf

Sleep Center, Division of Pulmonary and Critical Care, Department of Medicine and Department of Neurology, Rutgers- Robert Wood Johnson Medical School, New Brunswick, NJ, USA

Correspondence: Matthew T Scharf, Division of Pulmonary and Critical Care Medicine, Rutgers-Robert Wood Johnson Medical School, MEB 535, 1 Robert Wood Johnson Place, New Brunswick, NJ, 08901, USA, Tel +1 732 235-7840, Fax +1 732 235-7944, Email Matthew.scharf@rwjms.rutgers.edu

Abstract: The Epworth sleepiness scale (ESS) is a commonly used questionnaire to evaluate patients for excessive daytime sleepiness (EDS). The ESS has been validated as a measure of EDS, but a number of studies have shown more test–retest variability in clinical settings compared to the original validation study. This observation of higher-than-expected test–retest variability has called into question the utility of the ESS as a clinical tool to assess EDS. The purpose of this review article is to summarize how studies of test–retest variability in clinical populations compare to the original validation study of Johns and to highlight where they differ. Furthermore, use of the ESS as a continuous variable (with no specified cutoff value) versus a categorical variable (normal versus high) is described. These observations are put into a clinical context by comparing the test–retest variability observed on the ESS with that of the multiple sleep latency test (MSLT). Finally, how contributors to ESS scores differ within certain subpopulations is described. The ESS remains an important tool to measure EDS in patient populations, but an awareness of its limitations needs to be considered.

Keywords: excessive daytime sleepiness, EDS, test-retest reliability, Cohen’s kappa, multiple sleep latency test

Evaluation of Excessive Daytime Sleepiness

Excessive daytime sleepiness (EDS) is a common complaint for which patients are referred to sleep clinic. EDS is important not only because it is a dysphoric feeling for patients and is associated with impairments in functioning such as an increased risk for motor vehicle crashes¹ and occupational injury,² but also because it is associated with increased mortality.^{3,4} In one poignant study that assessed the contribution of EDS and obstructive sleep apnea (OSA) to mortality in elderly patients, those with obstructive sleep apnea (OSA) and EDS had increased mortality; those with either OSA or EDS alone did not have increased mortality.⁵ Similarly, in patients with moderate-severe OSA, the risk of major adverse cardiac events was higher in those with EDS.⁶ These studies suggest that EDS may be a critical component in linking certain diseases, such as OSA, to deleterious outcomes including mortality. In fact, exclusion of patients with EDS in certain randomized controlled trials of OSA treatment has cast doubts on the generalizability of the findings.⁷ However, despite the fact that EDS is clearly important, there is no consensus on the optimal way to assess EDS. There are objective measures such as the multiple sleep latency test (MSLT) and maintenance of wakefulness test (MWT), and self-reported measures including single-item questions,^{3,4} two-item questionnaires,^{8,9} and more comprehensive validated questionnaires including the Epworth Sleepiness Scale (ESS).¹⁰ The ESS has been used extensively in clinical and research settings and will be the focus of this review.

Development of the Epworth Sleepiness Scale

The ESS was developed by Murray Johns at Epworth Hospital in Australia and was first reported in 1991.¹⁰ The ESS was designed to be a simple test to administer and interpret. It asked patients to rate their likelihood of “dozing” in eight different scenarios with a minimum score of 0 indicating “would never doze” to a maximum score of 3 indicating a “high chance of dozing.” The total score can range from a minimum of 0 indicating a low level of sleepiness to a maximum of

24 indicating a very high level of sleepiness. Patients were asked to refer to “your usual way of life in recent times” rather than how they are feeling at the moment.

In the original validation study, ESS scores were higher in patients with OSA, narcolepsy, and idiopathic hypersomnia compared to control subjects and were lower in patients with insomnia. Furthermore, the ESS score was associated with OSA severity such that patients with worse OSA had higher ESS scores. In a small subset of patients for whom MSLTs were available, the ESS score was inversely correlated with mean sleep latency.¹⁰ The associations between the ESS and OSA and the ESS and the MSLT were further detailed in subsequent studies by Johns.^{11,12}

One important consideration with any measure is test–retest reliability. In other words, in the absence of an intervention, one would expect the score to be stable over time. Johns showed that in healthy medical students, ESS scores were very similar when done at two separate time points spaced 5 months apart. He also demonstrated that in patients with OSA who were subsequently treated with continuous positive airway pressure (CPAP), their ESS scores decreased.¹³ These observations suggest that in healthy individuals, the ESS scores are stable and that in sleepy patients, the scores decrease with appropriate intervention. Of note, test–retest reliability in non-treated OSA patients or other clinical populations was not assessed.

Test-Retest Reliability in Clinical Settings

To determine whether Johns' observations of test–retest reliability in healthy medical students are similar in clinical populations, a number of studies have assessed test–retest reliability of the ESS in patients presenting for clinical care. The ESS score observed on one visit was compared to the ESS score observed on a subsequent visit. These studies have demonstrated some important differences when comparing ESS test–retest reliability in clinical populations to healthy subjects (Table 1; Supplemental Document).^{14–17}

Features that are common to Johns' original study and studies on clinical populations are that the average scores between the two administrations are similar and the correlation coefficients are generally similar. The divergence between these two groups emerges on measures of intra-individual variability. For example, in the study of Johns, only 3% of subjects had a discrepancy in the ESS scores of ≥ 5 between the two administrations of the ESS. In the clinical populations included in Table 1, a discrepancy in the ESS scores of ≥ 5 was seen in 14% of subjects in the study with

Table 1 Studies Comparing Test–Retest Reliability of the ESS

	Johns	Nguyen	Campbell	Taylor (GP/OX)	Taylor (OX/SPC)	Walker
Difference in ESS scores						
≥ 2	48%	61%	61%	78%	46%	64%
≥ 3	18%	41%	46%	66%	37%	48%
≥ 5	3%	23%	21%	42%	14%	15%
≥ 7	N/A	10%	8%	27%	5%	7%
ESS #1 mean (S.D.)	7.4 (3.9)	11.1 (5.2)	13.7 (4.9)	12.2 (5.3)	11.9 (5.5)	9.8 (5.4)
ESS #2 mean (S.D.)	7.6 (3.8)	11.2 (5.3)	13.5 (4.6)	14.6 (5.2)	12.8 (5.5)	10.2 (6.2)
n	87	142	154	88	112	85
Correlation Coefficient	0.82	0.73	0.67	0.35	0.82	0.82
Cohen's kappa (ESS ≥ 11)	N/A	0.52	0.53–0.70	0.09	0.66	0.67
Time b/t ESS #1 and #2	5 months	Average=71 days	N/A	Median=91 days	Median=111 days	Average=25 days
Population	Healthy	Clinical	Clinical	Clinical	Clinical	Clinical

Notes: See Supplemental Section for methodological details on how this data was obtained. These data are presented from Johns,¹³ Nguyen et al,¹⁴ Campbell et al,¹⁵ Taylor et al,¹⁶ and Walker et al.¹⁷

Abbreviations: ESS, Epworth sleepiness scale; GP, general practitioner; OX, oximetry; S.D., standard deviation, SPC, sleep specialist.

the *lowest* intra-individual variability.¹⁶ Similarly, in the study of Johns, 18% of subjects had a discrepancy of ≥ 3 between the two administrations of the ESS, but in the clinical populations, there was a discrepancy of ≥ 3 in 37% of subjects in the study with the *lowest* intra-individual variability.¹⁶

Since there was higher test–retest variability observed in clinical populations compared to healthy, young subjects, it would be logical that some covariates or factors prevalent in clinical populations could explain this finding. However, age,^{14,16,17} apnea–hypopnea index,^{14–17} body mass index (BMI),^{15,17} sex,^{14,15,17} race,¹⁷ presence of diabetes,¹⁵ presence of hypertension,¹⁵ occupation,¹⁶ type of OSA testing,¹⁷ and time interval between ESS administrations^{14,16,17} did not account for the test–retest variability on ESS scores. In fact, only one study identified any covariates or factors associated with the variability on ESS scores, age, and that contribution was very small in magnitude such that the entire model, including age, only explained 3% of the variability in ESS scores.¹⁵ Therefore, the contributors to the test–retest variability in ESS scores observed in clinical populations remain unknown. Furthermore, the reasons why there was relatively high variability observed in one part of one of the included studies (Taylor oximetry/general practitioner), compared to the other part of that study (Taylor oximetry/specialist) as well as the other included studies (Nguyen, Campbell, and Walker) remain unknown (Table 1).

Test–Retest Variability with the Epworth Sleepiness Scale as a Categorical Measure (Normal versus High)

In research settings, the ESS score is often evaluated as a continuous measure; there is no specified cutoff and essentially, ESS scores are viewed as higher or lower. In contrast, in some research settings and many clinical settings, ESS scores are assessed as normal versus high; a cutoff is specified and scores above that point are considered to indicate EDS. In the original validation study of Johns, the control subjects had an ESS score of 2–10, and scores ≥ 11 were only found in the evaluated subjects with a sleep disorder.¹⁰ Johns and Hocking made a similar observation in a subsequent study.¹⁸ Since then, there is a common convention of using an ESS ≥ 11 to indicate EDS.^{9,19–26} In clinical settings, diagnostic testing and treatment are often predicated upon an “abnormal” score so the use of cutoff is often necessary.

With use of a cutoff, one could ask if it is the use of a measure as a categorical rather than a continuous variable that explains the results or is it the use of that particular cutoff? In other words, if a different cutoff is specified, will the results be substantially different? In a study assessing ESS normalization, defined as improvement of the ESS from ≥ 11 to < 11 , in patients with OSA treated with CPAP, the authors found a threshold of hours of CPAP use beyond which there was relatively little improvement in the rate of ESS normalization. A similar threshold for hours of CPAP use was obtained with use of other ESS cutoffs (8 or 12) to define normalization of the ESS.¹⁹ These results suggest that it is the use of the ESS as a categorical variable, rather than the use of the specific cutoff of 11, that is more important with this type of use of the ESS.

An important question to consider when assessing test–retest variability of the ESS in clinical populations is as follows: are the differences meaningful? Specifically, what is the chance that an ESS score remains normal or abnormal with two administrations of the ESS with no intervention between tests? One way to address this question is with use of Cohen’s kappa with use of a cutoff score of the ESS. Cohen’s kappa (for categorical variables) measures how often scores will remain on either side of a cutoff value with separate testing. In this context, if there was a group of patients for whom the ESS was repeated and the values that were < 11 remained < 11 every time, and the values that were ≥ 11 remained ≥ 11 every time, Cohen’s kappa would be 1, indicating excellent agreement. If, however, the values < 11 increased to ≥ 11 half the time and the values that started ≥ 11 decreased to < 11 half the time (i.e. no better than a coin flip), Cohen’s kappa would be 0, indicating poor agreement. Using the convention of defining EDS as an ESS ≥ 11 , values for Cohen’s kappa with repeat administrations of the ESS in clinical populations generally range between 0.5 and 0.7, suggesting moderate to substantial agreement²⁷ in classifying the scores as normal versus high (Table 1). Interestingly, in data derived from one of these studies,¹⁷ Cohen’s kappa was in the range of 0.5–0.7 with ESS cut-offs of either 8, 9, 10, 11, 12, 13 or 14, suggesting that there is moderate to substantial agreement with use of a threshold within this range, not necessarily due to the specific threshold of 11. This author has previously argued that these values are similar to those observed in a wide array of clinically-used measures including polysomnographic determination of OSA severity.¹⁷ This

suggests that the test–retest reliability of the ESS in determining EDS with the use of a cutoff of an $ESS \geq 11$ is consistent with other tests in widespread clinical use.

Perhaps most illuminating would be to compare the ESS to another test used to assess EDS in clinical populations. The MSLT is in widespread clinical use and is part of the diagnostic evaluation for narcolepsy and idiopathic hypersomnia.²⁸ An MSLT is performed by giving a patient five nap opportunities during the day spaced 2 hours apart. The sleep latency is measured for each nap, and a mean sleep latency is calculated from these naps. Whether rapid eye movement (REM) sleep is observed in each nap is also noted and sleep-onset REM periods are reported. A mean sleep latency of less than 8 minutes is considered to indicate an abnormal degree of sleepiness, and the number of sleep-onset REM periods is further used to identify narcolepsy. Obviously, for the results to be meaningful, subsequent testing in the absence of an intervention should reveal similar results. However, ample evidence has cast doubt on this. In one study of patients with either narcolepsy without cataplexy, idiopathic hypersomnia or physiologic hypersomnia, the results of two separate MSLTs were compared. There was no correlation between the two tests, a change in mean sleep latency crossing the 8-minute threshold occurred in 42% of patients and a change in sleep-onset REM periods crossing the threshold of 2 occurred in 31% of patients.²⁹ From the data presented in the study, Cohen's kappa for mean sleep latency for normal versus abnormal (8 minute cutoff) was 0.11 indicating poor agreement. Similar results were obtained in a subsequent study examining test–retest reliability of the MSLT in patients with narcolepsy without cataplexy, idiopathic hypersomnia and unspecified EDS. Furthermore, a longer test–retest interval was associated with an improved mean sleep latency and MSLT normalization. Interestingly, there was better test–retest reliability in patients with narcolepsy with cataplexy.³⁰ It should be noted that the inter-test interval in these studies was 4.2 (mean) and 1.8 (median) years, respectively, and whether there would be similar variability if the inter-test interval was substantially shorter is unknown. Nonetheless, these observations and other similar results suggest a limited utility for the MSLT in testing patients with disorders other than narcolepsy without cataplexy.³¹ More broadly, this highlights the variability in measuring EDS over time and demonstrates that even this “objective” measure has important caveats.

A fundamental question to consider when assessing test–retest variability is as follows: are the observed changes a limitation of the test, or alternatively, reflective of actual variability in the object of measure? In other words, with serial measurements of EDS, is there some natural variability that occurs over time that is being identified by EDS-assessment tools? Changes that may occur in everyday life including changes in sleep duration, sleep timing, use of stimulants such as caffeine, use of medications and psychosocial stressors likely vary over time and some of these changes may be reflected in measures of EDS. It is likely that this will explain at least part of the variability observed in serial measurements of the ESS over time. It would also be interesting to know whether there is less test-rest variability associated with very high or very low scores on the ESS.

Limitations of the ESS

There is no clear definition of EDS and, as noted above, there is no real “gold standard” test. It would therefore seem logical to assume that individual tools used to assess EDS may perform differently in certain groups and populations. One area that has received attention in this regard is sex. In adjusted analyses of data from the large community-based Sleep Heart Health Study, women and men showed a similar rate of EDS in a simple sleepiness question, but women were less likely to have an abnormal ESS score of ≥ 11 suggesting that the ESS may not have identified sleepiness in women as well as in men.⁸ When looking at ESS normalization (improvement from ≥ 11 to < 11) in patients with OSA started on CPAP therapy, one study showed female sex was associated with an increased likelihood of ESS normalization,²¹ while another study showed that female sex was associated with a decreased likelihood of ESS normalization.²⁶ Consistent with sex-specific effects on the ESS, in a study of patients with epilepsy, the ESS was associated with OSA risk and antiepileptic drugs in men, but in women it was not associated with these covariates and was associated with depression.³² These studies suggest that the ESS may not be an identical measure in women and men. It therefore seems likely that the ESS score may be due to different variables in different groups and the nature of these interactions requires further elucidation.

Conclusion

The question of what constitutes EDS remains enigmatic. There are various tools to measure EDS, but there is no ideal method. The ESS is an inexpensive, widely available and commonly used questionnaire to assess EDS. Although the ESS is a well-validated measure of EDS, a number of studies have demonstrated higher test–retest variability in clinical populations compared to the original validation study, calling into question the appropriateness of the routine use of ESS in clinical settings. With categorical use of the ESS where a cutoff of ≥ 11 is used to classify patients as normal versus high, test–retest variability of the ESS is consistent with other tools in clinical use and may even compare favorably to the MSLT. One limitation of the ESS is that it is likely to be a measure of different variables in different populations, and a given score in one group may be comprised differently than a similar score in a different group. Clinically, determining whether a patient has EDS remains a challenge. The ESS is a useful tool but should be viewed in a larger clinical context along with the patient's presentation, history and other diagnostic testing.

Acknowledgments

The author thanks the faculty and staff of the Comprehensive Sleep Disorders Center at Rutgers Robert Wood Johnson Medical School for assistance in data collection in some reported studies. The author thanks Dr. Elizabeth Taylor for sharing the raw data from her study. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure

The author reports no conflicts of interest in this work.

References

1. Ward KL, Hillman DR, James A, et al. Excessive daytime sleepiness increases the risk of motor vehicle crash in obstructive sleep apnea. *J Clin Sleep Med*. 2013;9(10):1013–1021. doi:10.5664/jcsm.3072
2. Melamed S, Oksenberg A. Excessive daytime sleepiness and risk of occupational injuries in non-shift daytime workers. *Sleep*. 2002;25(3):315–322. doi:10.1093/sleep/25.3.315
3. Empana JP, Dauvilliers Y, Dartigues JF, et al. Excessive daytime sleepiness is an independent risk indicator for cardiovascular mortality in community-dwelling elderly: the three city study. *Stroke*. 2009;40(4):1219–1224. doi:10.1161/STROKEAHA.108.530824
4. Newman AB, Spiekerman CF, Enright P, et al. Daytime sleepiness predicts mortality and cardiovascular disease in older adults. The Cardiovascular Health Study Research Group. *J Am Geriatr Soc*. 2000;48(2):115–123. doi:10.1111/j.1532-5415.2000.tb03901.x
5. Gooneratne NS, Richards KC, Joffe M, et al. Sleep disordered breathing with excessive daytime sleepiness is a risk factor for mortality in older adults. *Sleep*. 2011;34(4):435–442. doi:10.1093/sleep/34.4.435
6. Xie J, Sert Kuniyoshi FH, Covassin N, et al. Excessive daytime sleepiness independently predicts increased cardiovascular risk after myocardial infarction. *J Am Heart Assoc*. 2018;7(2). doi:10.1161/JAHA.117.007221
7. Pack AI, Magalang UJ, Singh B, Kuna ST, Keenan BT, Maislin G. Randomized clinical trials of cardiovascular disease in obstructive sleep apnea: understanding and overcoming bias. *Sleep*. 2021;44(2). doi:10.1093/sleep/zsaa229
8. Baldwin CM, Kapur VK, Holberg CJ, Rosen C, Nieto FJ; Sleep Heart Health Study G. Associations between gender and measures of daytime somnolence in the Sleep Heart Health Study. *Sleep*. 2004;27(2):305–311. doi:10.1093/sleep/27.2.305
9. Kapur VK, Baldwin CM, Resnick HE, Gottlieb DJ, Nieto FJ. Sleepiness in patients with moderate to severe sleep-disordered breathing. *Sleep*. 2005;28(4):472–477. doi:10.1093/sleep/28.4.472
10. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540–545. doi:10.1093/sleep/14.6.540
11. Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea. The Epworth Sleepiness Scale. *Chest*. 1993;103(1):30–36. doi:10.1378/chest.103.1.30
12. Johns MW. Sleepiness in different situations measured by the Epworth Sleepiness Scale. *Sleep*. 1994;17(8):703–710. doi:10.1093/sleep/17.8.703
13. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep*. 1992;15(4):376–381. doi:10.1093/sleep/15.4.376
14. Nguyen AT, Baltzan MA, Small D, Wolkove N, Guillon S, Palayew M. Clinical reproducibility of the Epworth Sleepiness Scale. *J Clin Sleep Med*. 2006;2(2):170–174. doi:10.5664/jcsm.26512
15. Campbell AJ, Neill AM, Scott DAR. Clinical reproducibility of the Epworth Sleepiness Scale for patients with suspected sleep apnea. *J Clin Sleep Med*. 2018;14(5):791–795. doi:10.5664/jcsm.7108
16. Taylor E, Zeng I, O'Dochartaigh C. The reliability of the Epworth Sleepiness Score in a sleep clinic population. *J Sleep Res*. 2019;28(2):e12687. doi:10.1111/jsr.12687
17. Walker NA, Sunderram J, Zhang P, Lu SE, Scharf MT. Clinical utility of the Epworth Sleepiness Scale. *Sleep Breath*. 2020;24(4):1759–1765. doi:10.1007/s11325-020-02015-2
18. Johns M, Hocking B. Daytime sleepiness and sleep habits of Australian workers. *Sleep*. 1997;20(10):844–849. doi:10.1093/sleep/20.10.844
19. Weaver TE, Maislin G, Dinges DF, et al. Relationship between hours of CPAP use and achieving normal levels of sleepiness and daily functioning. *Sleep*. 2007;30(6):711–719. doi:10.1093/sleep/30.6.711

20. Antic NA, Catcheside P, Buchan C, et al. The effect of CPAP in normalizing daytime sleepiness, quality of life, and neurocognitive function in patients with moderate to severe OSA. *Sleep*. 2011;34(1):111–119. doi:10.1093/sleep/34.1.111
21. Budhiraja R, Kushida CA, Nichols DA, et al. Predictors of sleepiness in obstructive sleep apnoea at baseline and after 6 months of continuous positive airway pressure therapy. *Eur Respir J*. 2017;50(5). doi:10.1183/13993003.00348-2017
22. Pepin JL, Viot-Blanc V, Escourrou P, et al. Prevalence of residual excessive sleepiness in CPAP-treated sleep apnoea patients: the French multicentre study. *Eur Respir J*. 2009;33(5):1062–1067. doi:10.1183/09031936.00016808
23. Koutsourelakis I, Perraki E, Economou NT, et al. Predictors of residual sleepiness in adequately treated obstructive sleep apnoea patients. *Eur Respir J*. 2009;34(3):687–693. doi:10.1183/09031936.00124708
24. Berger M, Hirotsu C, Haba-Rubio J, et al. Risk factors of excessive daytime sleepiness in a prospective population-based cohort. *J Sleep Res*. 2020;30(2):e13069. doi:10.1111/jsr.13069
25. Gottlieb DJ, Whitney CW, Bonekat WH, et al. Relation of sleepiness to respiratory disturbance index: the Sleep Heart Health Study. *Am J Respir Crit Care Med*. 1999;159(2):502–507. doi:10.1164/ajrccm.159.2.9804051
26. Scharf MT, Zhang P, Walker NA, et al. Sex differences in Epworth Sleepiness Scale normalization with continuous positive airway pressure. *J Clin Sleep Med*. 2022;18(9):2273–2279. doi:10.5664/jcsm.10048
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174. doi:10.2307/2529310
28. American Academy of Sleep Medicine. Central disorders of hypersomnolence. In: *The International Classification of Sleep Disorders – Third Edition (ICSD-3) Online Version*. Westchester, IL: American Academy of Sleep Medicine; 2014.
29. Trotti LM, Staab BA, Rye DB. Test-retest reliability of the multiple sleep latency test in narcolepsy without cataplexy and idiopathic hypersomnia. *J Clin Sleep Med*. 2013;9(8):789–795. doi:10.5664/jcsm.2922
30. Lopez R, Doukkali A, Barateau L, et al. Test-retest reliability of the multiple sleep latency test in central disorders of hypersomnolence. *Sleep*. 2017;40(12). doi:10.1093/sleep/zsx164
31. Trotti LM. Twice is nice? Test-retest reliability of the Multiple Sleep Latency Test in the central disorders of hypersomnolence. *J Clin Sleep Med*. 2020;16(S1):17–18. doi:10.5664/jcsm.8884
32. Jo S, Kim HJ, Kim HW, Koo YS, Lee SA. Sex differences in factors associated with daytime sleepiness and insomnia symptoms in persons with epilepsy. *Epilepsy Behav*. 2020;104(Pt A):106919. doi:10.1016/j.yebeh.2020.106919

Nature and Science of Sleep

Dovepress

Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/nature-and-science-of-sleep-journal>