

Validation of an Automated Wireless System for Sleep Monitoring During Daytime Naps

Nicola Cellini

*Department of General Psychology
University of Padua*

Elizabeth A. McDevitt, Ashley A. Ricker, Kelly M. Rowe, and
Sara C. Mednick

*Department of Psychology
University of California, Riverside*

An automated wireless system (WS) for sleep monitoring was recently developed and validated for assessing nighttime sleep. Here, we aimed to evaluate the validity of the WS to correctly monitor daytime sleep during naps compared to polysomnography (PSG). We found that the WS underestimated wake, sleep onset latency, and wake after sleep onset. Meanwhile, it overestimated total sleep time, sleep efficiency, and duration of REM sleep. Sensitivity was moderate for wake (58.51%) and light sleep (66.92%) and strong for deep sleep (83.46%) and REM sleep (82.12%). These results demonstrated that the WS had a low ability to detect wake and systematically overscored REM sleep, implicating the WS as an inadequate substitute for PSG in diagnosing sleep disorders or for research in which sleep staging is essential.

A continuous challenge in sleep research is the development of cost-effective systems for sleep monitoring in real-world environments. Several portable monitoring devices have been developed as an alternative to polysomnography (PSG), the gold standard for sleep recording. Generally, compared to PSG, these devices are less expensive, less invasive, more user-friendly, and more able to acquire several days of recordings in nonlaboratory settings (Kelly, Strecker, & Bianchi, 2012). However, these devices (e.g., bed sensors, noncontact biomotion sensors) do not differentiate sleep stages and remain at the prototype level (Van de Water,

Correspondence should be addressed to Sara C. Mednick, Department of Psychology, University of California, Riverside, 900 University Ave., Riverside, CA 92521, USA. E-mail: smednick@ucr.edu

Holmes, & Hurley, 2011) until they can be accurately validated for clinical and research purposes.

An automated wireless sleep-stage monitoring system was recently developed (WS; Zeo, Inc., Newton, MA), consisting of three dry frontal electrodes integrated in an elastic headband placed on the forehead, which records electroencephalogram activity, muscle tone, and eye movements (Shambroom, Fabregas, & Johnstone, 2012). These data are wirelessly transmitted to a bedside base station, which automatically scores the data as wake or a specific sleep stage (i.e., light, deep, and REM sleep). The automated sleep staging, portability, and user-friendly configuration of this device makes the WS a promising substitute for PSG. The WS could potentially be used for several clinical and research purposes, such as assessing the effect of short naps on performance in a workplace environment, or providing ecologically valid measures of afternoon naps in children in preschool and kindergarten. Clinically, the device could be employed to continuously monitor the duration and quality of sleep in shift workers, or the effect of treatments on sleep disorders such as insomnia.

To date, three laboratory studies have validated the WS against concurrent nighttime PSG-recording. A company-sponsored study compared the WS to PSG and actigraphy (Actiwatch 64) in healthy adults and reported overall good agreement between the WS and PSG (Shambroom et al., 2012). Moreover, the WS performed better than actigraphy in discriminating sleep from wakefulness. Tonetti and colleagues (2013) found moderate to high agreement, measured by Cohen's kappa, for the different sleep stages between PSG and the WS in a group of healthy young adults. Griessenberger, Heib, Kunz, Hoedlmoser, and Schabus (2012) tested the WS, an automated sleep staging system and manual scoring (using the AASM criteria; Iber, Ancoli-Israel, Chesson, & Quan, 2007) against a semiautomatic sleep staging system (Somnolyzer 24x7) in a mixed sample of insomnia patients and healthy sleepers (Griessenberger et al., 2012). They reported moderate overall agreement between the WS and the Somnolyzer, but worse performance compared to the other staging systems. These studies concluded that the WS was useful for sleep monitoring at home with some weaknesses related to overscoring of REM sleep and an underestimation of wakefulness (i.e., sleep onset and wake after sleep onset [WASO]). Notably, despite not being sold or marketed as a scientific device, the WS has already been used as the primary sleep monitor in published studies (Gumenyuk et al., 2011; Kudesia & Bianchi, 2012; Scullin, 2012).

Prior validation studies assessed performance of the WS during nocturnal sleep. To the best of our knowledge, the WS has yet to be validated for daytime sleep (naps). Although few studies have investigated differences between daytime and nocturnal sleep architecture (Milner & Cote, 2009), daytime sleep is known to have increased wakefulness and stage 1 sleep (15% vs. 2%; Nishida & Walker, 2007; Payne et al., 2009; Wamsley, Tucker, Payne, Benavides, & Stickgold, 2010; Wamsley, Tucker, Payne, & Stickgold, 2010), and lower sleep efficiency than nocturnal sleep. Decreased sleep efficiency (SE) during daytime naps (e.g., between 68% and 77% SE; Kanady, Drummond, & Mednick, 2011; Nishida & Walker, 2007; Tucker et al., 2006) is mainly due to the "weight" of sleep onset latency in a short sleep period. These differences pose an interesting challenge for monitoring daytime sleep with a device that uses an algorithm rendered on nighttime PSG-data recording (Shambroom et al., 2012), and has already shown limitations in its ability to detect wakefulness (Griessenberger et al., 2012). The aim of the current study was to determine the validity of the WS for monitoring sleep during daytime naps.

METHOD

Participants

Thirty healthy, nonsmoking young adults (17 female, $M_{\text{age}} = 20.3$, $SD = 2.76$) gave informed consent to participate in the study, which was approved by the University of California, Riverside Human Research Protections Program. Exclusionary criteria included: (a) having an irregular sleep-wake schedule (reporting a habitual time in bed [TIB] longer/shorter than 7–9 hrs per night); (b) having a sleep disorder. Sleep disorders were screened by interviewing the subject at the first meeting and asking about potential symptoms of insomnia, apnea, narcolepsy, restless leg syndrome/periodic leg movements; (c) any personal or immediate family (i.e., first-degree relative) history of diagnosed significant psychopathology; (d) personal history of head injury with loss of consciousness greater than 2 minutes, or seizures; (e) history of substance dependence; (f) current use of any psychotropic medications; and (g) any cardiac, respiratory, or other medical condition that may affect metabolism. Participants received financial compensation or course credit for participating in the study.

Procedure

The study took place at the Sleep and Cognition Lab in the Department of Psychology at the University of California, Riverside. Each participant completed one PSG-recorded nap while wearing the WS between 1:30 p.m. and 4:00 p.m. Sleep was monitored online by a trained sleep technician. Participants were awakened after they had accrued 90 min of total sleep time, or they had spent two hr in bed, whichever occurred first. Participants were also pulled out of bed if they spent more than 30 minutes continuously awake without falling asleep. The WS and PSG computer clocks were synchronized before each recording in order to compare the WS and PSG records.

Polysomnography

PSG recordings were collected using Astro-Med Grass Heritage Model 15 amplifiers with Grass Gamma software. Scalp electroencephalogram and electrooculogram electrodes were referenced to unlinked contralateral mastoids (C3/A2, C4/A1, O1/A2, LOC/A2, and ROC/A1), and muscle tone electromyogram electrodes were attached under the chin according to the International 10-20 system (Jasper, 1958). Raw data were digitized at a sampling rate of 256 Hz and visually scored in 30 s epochs following the American Academy of Sleep Medicine (AASM) rules for sleep staging (Iber et al., 2007). In order to reduce the possible level of variability, scoring was performed by a single well-trained sleep technician. The technician demonstrated a reliability of 88% with his own scores.

Wireless System Recording

The WS (Zeo, Inc., Newton, MA) used a headband with a single bipolar dry fabric sensor (and a single ground lead), which acquired EEG, EOG, and EMG signals. Data were transmitted to a bedside base station and automatically scored in 30 s epochs following Rechtschaffen

and Kales' (1968) sleep scoring criteria using a proprietary algorithm. For the technical specifications of the WS, see Shambroom et al. (2012).

Data Analysis

All analyses were confined to the period between lights off and lights on, which was marked on the PSG recording and synchronized with WS.

Sleep Summary

The following sleep parameters were calculated for the two systems: total sleep time (TST), defined as the number of minutes scored as sleep between lights off and lights on; sleep-onset latency (SL), minutes between lights off and the first epoch scored as any stage other than wake; wake after sleep onset (WASO), minutes scored as wake after sleep onset; sleep efficiency (SE), the ratio between TST and total time in bed (i.e., minutes from lights out to lights on); and the number of awakenings lasting at least 2 minutes ($NA \geq 2$). In order to compare wake and sleep staging of PSG and WS, we also calculated total minutes spent in wake, light sleep (stage N1 and N2), deep sleep (stages N3), and REM sleep.

Wilcoxon Matched-Pairs Rank Sum tests were used to compare nonnormally distributed sleep parameters, and paired *t*-tests to compare normally distributed variables (determined by Shapiro-Wilk tests), between WS and PSG. Intraclass correlation coefficients (ICCs) were calculated to assess the relationship between the systems. Statistical significance was set at $p < .05$ for all analyses.

Bland-Altman Statistic

Bland-Altman statistics (Altman & Bland, 1983; Bland & Altman, 1999), a technique that plots the difference score between two measures against their average, assessed the concordance between sleep summaries derived from the WS and PSG. In order to determine the significance of the "bias," we computed the upper and the lower limits of the mean difference between the two methods based on 95% confidence intervals.

Epoch-by-Epoch Agreement

In order to assess the ability of the WS to correctly identify different sleep stages, we computed two dichotomous variables for each stage: specific sleep stage identification from visual scoring of the PSG data (e.g., N1 = 1, N2 = 1, and the remaining epochs = 0) and automated sleep stage scoring by the WS (e.g., light sleep = 1 and the remaining epochs = 0). We created a 2×2 contingency table for each sleep stage by combining these two dichotomous variables, which was used to compute the following parameters: accuracy (the agreement rate between PSG and WS), sensitivity (ability of the WS to detect a specific stage that was also scored from PSG), predictive positive value (PPV; the percentage of epochs scored as a specific stage by the WS that were also scored in the same way via PSG) and Cohen's kappa (a measure of interrater reliability that reflects the percentage of scoring agreement between two methods not

due to chance). According to the Landis and Koch scale (Landis & Koch, 1977), we considered a kappa coefficient of 0–0.2 as slight agreement, 0.2–0.4 as fair agreement, 0.4–0.6 as moderate agreement, 0.6–0.8 as substantial agreement, and 0.8–1.0 as almost perfect agreement.

By comparing wake to nonwake epochs, we provided an index of the ability of the device to discriminate sleep from wakefulness. For the sleep/wake comparison, we also computed specificity, the ability of the WS to detect wakefulness when the PSG also scored wakefulness.

RESULTS

Sleep Summary

Results from Shapiro-Wilk tests showed that all sleep variables were nonnormally distributed ($p < .05$), with the exception of light sleep ($p = .499$). The Wilcoxon Matched-Pairs Rank-Sum tests showed that WS significantly overestimated TST, SE, and minutes of REM, but underestimated SL, WASO, and wake. No differences were found in $NA \geq 2$, light and deep sleep (Table 1).

ICCs revealed significant associations between the two systems for all of the parameters. The weakest relationship between devices was found for REM sleep and the strongest for TST and deep sleep (Table 2).

Bland-Altman

Also the Bland-Altman statistic (Table 3) highlighted an overestimation of TST, SE, and minutes of REM and a systematic underestimation of wake as a consequence of underestimating both SL and WASO (Figure 1).

Epoch-by-Epoch Agreement

Accuracy, sensitivity, PPV, and Cohen's kappa values for each sleep stage are reported in Table 4. Wake and light sleep showed moderate sensitivity, whereas REM and deep sleep

TABLE 1
Sleep Measures (Mean \pm SD) for WS and PSG

	PSG	WS	Z	p
TST (min)	58.85 \pm 16.43	66.07 \pm 17.33	3.68	.001
SL (min)	9.85 \pm 6.11	7.82 \pm 5.68	1.97	.049
WASO (min)	13.60 \pm 13.01	8.42 \pm 9.01	3.09	.002
SE (%)	72.40 \pm 14.15	80.96 \pm 12.98	3.62	.001
$NA \geq 2$ (nr.)	2.27 \pm 1.76	1.70 \pm 1.47	1.89	.059
Wake (min)	23.45 \pm 14.66	16.23 \pm 12.08	3.68	.001
Light (min)	36.28 \pm 13.46	32.67 \pm 13.19	1.38*	.177
Deep (min)	17.53 \pm 15.73	14.90 \pm 13.22	1.67	.095
REM (min)	5.03 \pm 8.29	18.50 \pm 17.39	4.51	.001

* = *t*-value.

TABLE 2
ICCs Between Systems for Each Sleep Variable

	ICCs
TST (min)	0.75***
SL (min)	0.53***
WASO (min)	0.62***
SE (%)	0.61***
NA \geq 2 (nr.)	0.58***
Wake (min)	0.62***
Light (min)	0.49**
Deep (min)	0.83***
REM (min)	0.24*

*** = $p < .001$; ** = $p < .01$; * = $p < .05$.

TABLE 3
Bland-Altman Statistics Between WS and PSG

	TST	SL	WASO	SE	NA \geq 2	Wake	Light	Deep	REM
Mean difference	7.22	-2.03	-5.19	8.56	-0.57	-7.22	-3.62	-2.69	13.47
Standard deviation	10.27	4.95	9.20	11.61	1.43	10.27	13.40	8.51	15.44
Upper limit	10.96	-0.26	-1.90	12.72	-0.55	-3.55	1.18	0.36	18.90
Lower limit	3.55	-3.80	-8.48	4.41	-1.08	-10.90	-8.42	-5.74	7.95
Over- or underestimation	Over	Under	Under	Over	Under	Under			Over

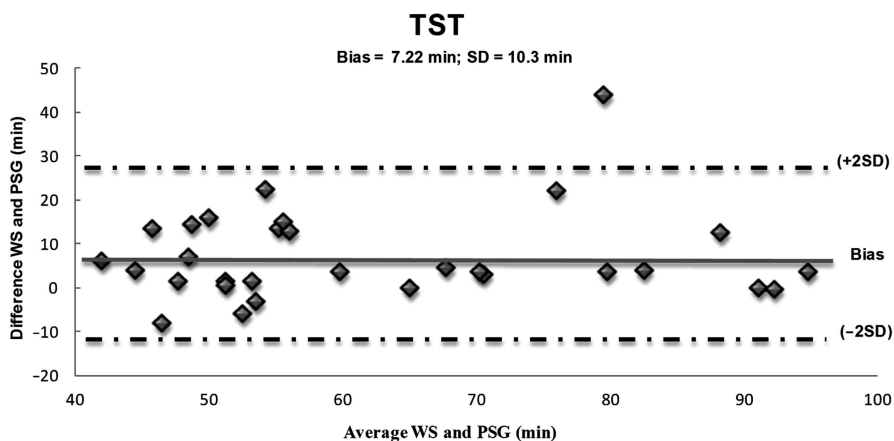


FIGURE 1 Bland-Altman plots of PSG and WS for sleep parameters. Only the significant parameters are shown. The y-axis indicates the differences between the WS score minus the PSG score, whereas the x-axis showed the average of their scoring. The bias represents the mean difference between the devices for a specific parameter, with values above zero meaning an overestimation and the values below zero meaning an underestimation of the WS relative to PSG. Dashed-dotted lines represent Bias \pm 2 SD. (Continued).

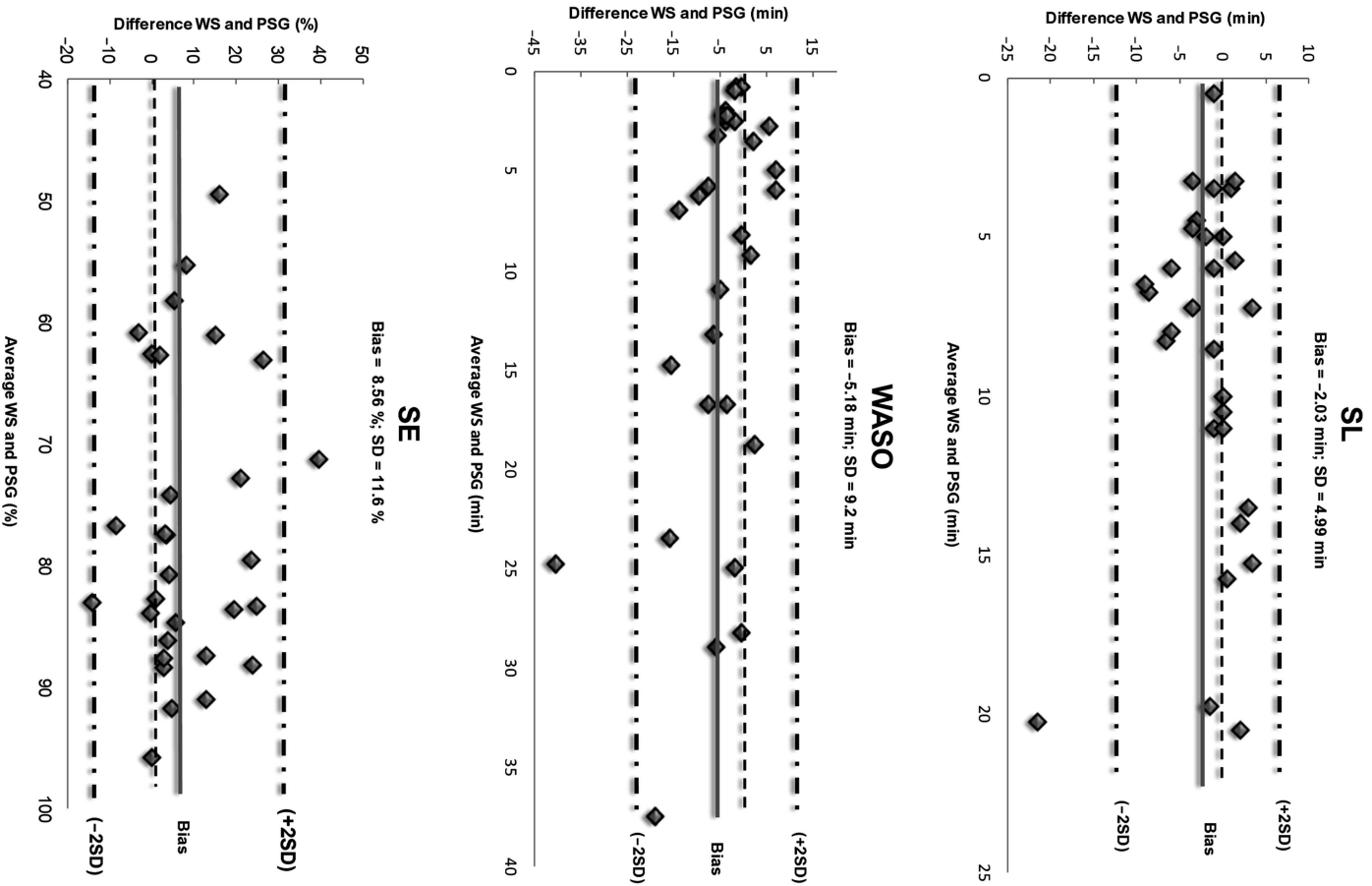


FIGURE 1 (Continued).

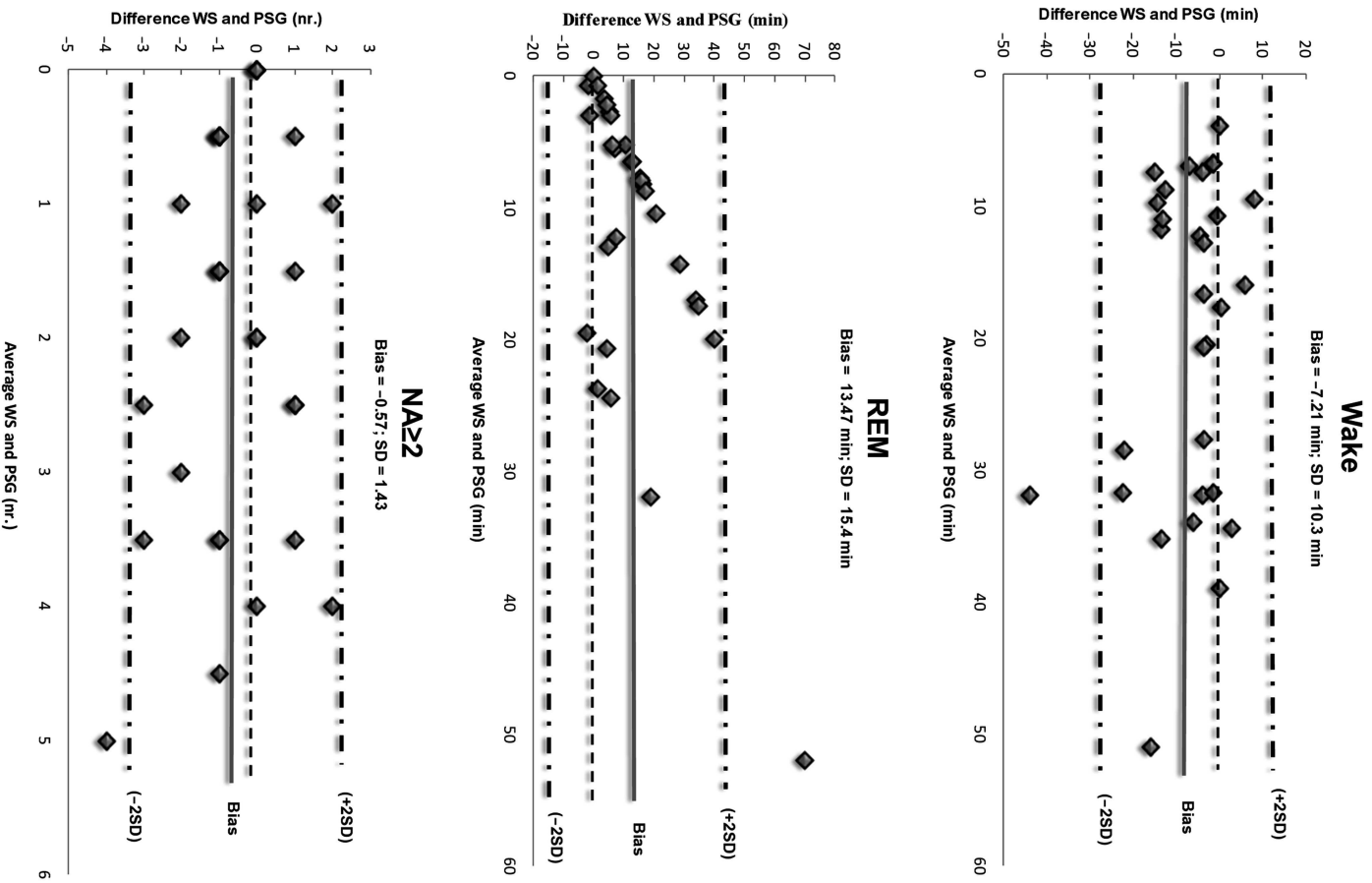


FIGURE 1 (Continued).

TABLE 4
Accuracy, Sensitivity, PPV, and Cohen's Kappa Values for Each
Sleep Stage

	<i>Wake</i>	<i>Light</i>	<i>Deep</i>	<i>REM</i>
Accuracy (%)	86.79	74.94	90.81	81.03
Sensitivity (%)	58.51	66.92	83.46	82.12
PPV (%)	86.64	75.60	72.03	22.34
Cohen's kappa	0.62	0.49	0.72	0.28

showed strong sensitivity. REM sleep showed extremely poor PPV. Cohen's kappa revealed an agreement between PSG and WS that was not due to chance, ranging from moderate (light sleep) to substantial (wake, deep sleep). However, Cohen's kappa highlighted very low agreement for REM sleep. Specifically, the WS correctly identified 82.1% of PSG-defined REM sleep, but it also defined 27.6% and 22.9% of wake and light sleep epochs as REM, respectively. These results also explain the reduced sensitivity of WS for wake and light sleep.

Regarding the sleep/wake comparison, accuracy and Cohen's kappa values statistically remained the same (86.8% and 0.62 respectively), whereas sensitivity, specificity, and PPV of the sleep/wake comparison were 96.8%, 58.51%, and 86.8%, respectively. Thus, the WS showed validity for detecting sleep, but only a moderate capacity to correctly identify wake.

DISCUSSION

The current study examined the ability of WS to accurately discriminate daytime sleep stages. We utilized multiple analyses to completely describe the performance of this device. Our main finding was the reduced ability of WS to detect wakefulness. Although Cohen's kappa showed a moderate/substantial agreement (0.62), the moderate specificity (58.51%) and the underestimation of SL, WASO, and $NA \geq 2$ (and consequent overestimation of TST and SE) suggests poor wake detection by the WS. These data confirm previous findings for nighttime sleep (Griessenberger et al., 2012; Shambroom et al., 2012; Tonetti et al., 2013). Thus, the limitation of the WS is likely an intrinsic feature of the device that is not specific to daytime sleep monitoring.

The second weakness of the WS is REM classification. Only considering accuracy (81.03%) and sensitivity (82.12%), the WS showed high epoch-by-epoch agreement with PSG. However, the WS also showed an overestimation of REM sleep (e.g., 1,110 epochs scored as REM by WS compared to 302 epochs scored as REM by PSG). This bias was confirmed by low PPV (22.34%), Cohen's kappa (0.28), and ICC (0.24) values. As reported in previous validation studies (Griessenberger et al., 2012; Shambroom et al., 2012), this overscoring of REM mostly occurred at the expense of wake (27.6%) and light sleep (22.9%). This may result from the absence of an independent EOG channel in the WS (Shambroom et al., 2011; Tonetti

et al., 2013). Additionally, sensor placement in the prefrontal area, where alpha activity is low, has been implicated in the WS's difficulty detecting quiet wakefulness (Tonetti et al., 2013).

It should be noted that contrary to previous studies (Shambroom et al., 2012; Tonetti et al., 2013), we failed to find significant differences between the WS and PSG sleep scoring methods for deep sleep. This may be due to the sleep stage scoring system applied. In the current study, we used AASM criteria (Iber et al., 2007), whereas Shambroom and colleagues (2012) used Rechtschaffen and Kales (1968), the same criteria used to develop the WS scoring algorithm. Although AASM criteria were reported to increase scoring of WASO, N1 and N3, and decrease scoring of N2 compared to Rechtschaffen and Kales criteria (Moser et al., 2009), the WS showed no differences in deep sleep stages compared to PSG. Nonetheless, the different scoring system could potentially explain the absence of significant differences for deep sleep.

For the sleep/wake comparison, WS sensitivity was high (96.8%), indicating good ability of the device to detect sleep when PSG also scores sleep. This sensitivity value was even higher than actigraphy reported for daytime sleep (92–96% in Kanady et al., 2011; 86–94% in Cellini, Buman, McDevitt, Ricker, & Mednick, 2013). WS also showed poor specificity (58.51%), similar to that of actigraphy for naps (40–66% in Kanady et al., 2011; 36–64% in Cellini et al., 2013), confirming the limited ability of this device to correctly detect wakefulness. These results corroborate previous reports on this WS. Tonetti et al. (2013) reported 97.6% sensitivity and 56.1% specificity, and Griessenberger et al. (2012) reported 40.8% specificity in a mixed sample of insomniacs and healthy sleepers (sensitivity for total sleep stages was not reported in this study), whereas Shambroom et al. (2012) reported a higher specificity of 64%. These differences could be the consequence of different samples and different sleep contexts (i.e., daytime and nighttime sleep).

A potential limitation of the current study relates to the age range of our sample. We studied healthy, young adults ($M_{\text{age}} = 20.3$), which may limit the generalization of our findings to other populations such as infants or older adults who exhibit greater variations in sleep quantity (e.g., total sleep time, minutes spent in different stages) and quality (e.g., amplitude of slow-wave sleep, fragmentation index; Cajochen, Münch, Knoblauch, Blatter, & Wirz-Justice, 2006). In addition, it is worth noting that we collected data in a laboratory setting, but home validation with concurrent ambulatory PSG could yield different results. Finally, further direct comparison among the WS, PSG, and different brands of actigraphy is warranted.

In conclusion, WS monitored daytime naps comparably to previous studies of nighttime sleep. Moreover, sleep/wake comparisons showed performance similar to actigraphy for daytime sleep (Kanady et al., 2011). Given the automatic sleep staging capability and relatively lower price compared to actigraphy (Cellini et al., 2013), the WS may be an excellent alternative to actigraphy for objective sleep monitoring outside the lab (Tonetti et al., 2012). Thus, the WS is a relatively inexpensive device for ecological sleep/wake monitoring that could be used as a screening tool to improve sleep hygiene and for general assessment of sleep quality and quantity. However, the systematic overscoring of REM and low ability to detect wake combine to make this system an inadequate substitute for PSG in diagnosing sleep disorders, or for research in which sleep staging is essential. As such, we recommend further improvements in the scoring algorithm and electrode placement to provide more effective wake and REM detection.

ACKNOWLEDGMENTS

Wireless systems were donated by Zeo, Inc., to Sara C. Mednick. Nicola Cellini, Elizabeth A. McDevitt, Ashley A. Ricker, and Kelly M. Rowe declare the absence of any conflicts of interest.

REFERENCES

- Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *The Statistician*, 32, 307–317.
- Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8(2), 135–160.
- Cajochen, C., Münch, M., Knoblauch, V., Blatter, K., & Wirz-Justice, A. (2006). Age-related changes in the circadian and homeostatic regulation of human sleep. *Chronobiology International*, 23(1–2), 461–474.
- Cellini, N., Buman, M. P., McDevitt, E. A., Ricker, A. A., & Mednick, S. C. (2013). Direct comparison of two actigraphy devices with polysomnographically-recorded naps in healthy young adults. *Chronobiology International*, 30(5), 691–698.
- Griessenberger, H., Heib, D. P. J., Kunz, A., Hoedlmoser, K., & Schabus, M. (2012). Assessment of a wireless headband for automatic sleep scoring. *Sleep and Breathing*, 17(2), 747–752.
- Gumenyuk, V., Roth, T., Korzyukov, O., Jefferson, C., Bowyer, S., & Drake, C. L. (2011). Habitual short sleep impacts frontal switch mechanism in attention to novelty. *Sleep*, 34(12), 1659–1670.
- Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. (2007). *The AASM manual for the scoring of sleep and associated events: Rules, terminology, and technical specifications*. Westchester, IL: American Academy of Sleep Medicine.
- Jasper, H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10(2), 371–375.
- Kanady, J., Drummond, S., & Mednick, S. (2011). Actigraphic assessment of a polysomnographic-recorded nap: A validation study. *Journal of Sleep Research*, 20(1, Pt. 2), 214–222.
- Kelly, J. M., Strecker, R. E., & Bianchi, M. T. (2012). Recent developments in home sleep-monitoring devices. *ISRN Neurology*, 2012, 768–794.
- Kudesia, R. S., & Bianchi, M. T. (2012, April 12). Decreased nocturnal awakenings in young adults performing Bikram yoga: A low-constraint home sleep monitoring study. *ISRN Neurology*. doi:10.5402/2012/153745
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Milner, C. E., & Cote, K. A. (2009). Benefits of napping in healthy adults: Impact of nap length, time of day, age, and experience with napping. *Journal of Sleep Research*, 18(2), 272–281.
- Moser, D., Anderer, P., Gruber, G., Parapatics, S., Loretz, E., Boeck, M., . . . Dorffner, G. (2009). Sleep classification according to AASM and Rechtschaffen & Kales: Effects on sleep scoring parameters. *Sleep*, 32(2), 139–149.
- Nishida, M., & Walker, M. P. (2007). Daytime naps, motor memory consolidation and regionally specific sleep spindles. *PLoS One*, 2(4), e341.
- Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L.-W., Wamsley, E. J., Tucker, M. A., . . . Stickgold, R. (2009). The role of sleep in false memory formation. *Neurobiology of Learning and Memory*, 92(3), 327–334.
- Rechtschaffen, A., & Kales, A. (1968). *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects*. Los Angeles: University of California, Brain Information Service/Brain Research Institute.
- Scullin, M. (2012, June 18). Sleep, memory, and aging: The link between slow-wave sleep and episodic memory changes from younger to older adults. *Psychology and Aging*. doi:10.1037/a0028830
- Shambroom, J. R., Fabregas, S. E., & Johnstone, J. (2012). Validation of an automated wireless system to monitor sleep in healthy adults. *Journal of Sleep Research*, 21(2), 221–230.
- Tonetti, L., Cellini, N., De Zambotti, M., Fabbri, M., Martoni, M., Fabregas, S., . . . Natale, V. (2013). Polysomnographic validation of a wireless dry headband technology for sleep monitoring. *Physiology & Behavior*, 118, 185–188.

- Tonetti, L., Fábregas, S. E., Fabbri, M., Occhionero, M., Erbacci, A., Martoni, M., & Natale, V. (2012). Comparison of a wireless dry headband technology for sleep monitoring with actigraphy in healthy adults. *Biological Rhythm Research*, *44*(2), 333–338.
- Tucker, M. A., Hirota, Y., Wamsley, E. J., Lau, H., Chaklader, A., & Fishbein, W. (2006). A daytime nap containing solely non-REM sleep enhances declarative but not procedural memory. *Neurobiology of Learning and Memory*, *86*(2), 241–247.
- Van de Water, A., Holmes, A., & Hurley, D. (2011). Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography: A systematic review. *Journal of Sleep Research*, *20*(1 Pt. 2), 183–200.
- Wamsley, E. J., Tucker, M., Payne, J. D., Benavides, J. A., & Stickgold, R. (2010). Dreaming of a learning task is associated with enhanced sleep-dependent memory consolidation. *Current Biology*, *20*(9), 860–855.
- Wamsley, E. J., Tucker, M. A., Payne, J. D., & Stickgold, R. (2010). A brief nap is beneficial for human route-learning: The role of navigation experience and EEG spectral power. *Learning & Memory*, *17*(7), 332–336.