Behavioral Biomarkers of Sleepiness

Thomas J. Balkin, Ph.D.

Department of Behavioral Biology, Walter Reed Army Institute of Research, Silver Spring, MD

For those endeavoring to develop better methods of measuring/quantifying sleepiness, the "Holy Grail" is a measure that is maximally objective, completely unobtrusive, exquisitely sensitive, and absolutely specific (i.e., varies *only* as a function of sleepiness). By these criteria, physiological measures (e.g., based on brain activity such as EEG, fMRI, near-infrared spectroscopy, etc.) would appear to hold the most promise. However, from an operational standpoint, the utility of a sleepiness measure is derived not from its ability to sensitively reflect the brain's extant level of sleepiness *per se*, but from the implications that this level of sleepiness has for the individual's current and near-term ability to safely and efficiently perform operationally-relevant tasks. Thus, an ideal operationally-relevant sleepiness measure is one that is unobtrusively embedded in the actual operational task, and allows sleepiness-related performance deficits to be distinguished from performance deficits due to other causes. Toward this end, we have developed a PVT-derived metric that incorporates the entire distribution of responses within a PVT session, and reflects changes in the *pattern* of performance that can be used to identify and quantify "state instability"—the putative physiological state that specifically underlies sleepiness-induced performance deficits.
Citation: Balkin TJ. Behavioral biomarkers of sleepiness. *J Clin Sleep Med* 2011;7(5):Supplement S12-S15.

There can be little doubt that sleepiness constitutes a significant problem for modern society. Although difficult to estimate with precision,[1] it is likely that in operational environments sleepiness contributes to accidents, errors, and inefficiencies that cost the US economy tens of billions of dollars per year.[2] Sleepiness likewise contributes directly to human suffering—a recent report by the Automobile Association of America suggests that sleepiness is a causal factor in 16.5% of fatal accidents on US highways (report available at http://www.aaafoundation.org/resources/index.cfm?button = research). Indeed, the mission of the National Sleep Foundation (NSF, http://www.sleepfoundation.org), founded in 1990, has been to increase public awareness of sleep disorders and sleepiness-related issues such as drowsy driving—and findings from the annual NSF "Sleep in America Poll" consistently suggest that sleepiness impacts the lives of a considerable portion of the American public in a variety of ways (e.g., see the 2009 NSF poll on "sleep and safety" at http://www.sleepfoundation.org/sites/default/files/2009%20POLL%20HIGHLIGHTS.pdf).

Although increased public awareness of the deleterious effects of sleepiness is an important first step, the direct benefits of such awareness—i.e., the impact that such awareness has on the behavior of individuals within society—is likely to be small. This is because 1) sleepy individuals are not particularly good at self-assessment of sleepiness and its effects on performance (e.g., Belenky et al.[3]; Schmidt et al.[4]; for review, see Balkin et al.[5]) and 2) the subjective experience of sleepiness is essentially a "moving target," with those suffering from various sleep disorders, or those who are chronically sleep restricted, subjectively habituating to the experience of sleepiness despite the fact that, from an objective standpoint, there has been no reduction in the actual physiological need for sleep. The dissociation of subjective and objective measures of sleepiness has recently been illustrated by Rupp et al.,[6] who compared the effects of 7 nights of sleep restriction (3 h time in bed [TIB]) in two groups. In one group, TIB was extended to 10 h per night for the week preceding the sleep restriction phase. For the other (Habitual TIB) group, TIB for the week prior to sleep restriction was maintained at each individual's typical duration (based on prior assessment with wrist actigraphy)—this was approximately 7 h per night. It was found that Psychomotor Vigilance Test (PVT) performance declined across the week of sleep restriction for both groups, but the rate of decline was slower for the Extended TIB group. As shown in **Figure 1**, the Extended TIB group both started and finished the sleep restriction phase with relatively greater levels of alertness/mental resources than did the Habitual TIB group. Of particular relevance, however, were the findings from the "recovery" phase of the study, consisting (for both groups) of 8 h TIB across 5 consecutive nights. As also illustrated in **Figure 1**, a single night of recovery sleep produced much greater recovery of next-day performance in the Extended TIB group than in the Habitual TIB group.

In stark contrast, and as illustrated in **Figure 2**, subjective alertness, as measured with the Stanford Sleepiness Scale (SSS), was comparably restored by the first night of recovery sleep (8 h TIB) in both groups. Based on these results, it is clear that subjective sleepiness is not experienced along a fixed, absolute scale. Instead, it varies as a function of prior sleep/wake history.

Since the relationship between subjective sleepiness and objective measures of sleepiness is malleable, it is clear that the potential usefulness of such subjective measures in operational environments (e.g., for assessing fitness for duty/risk of accidents) is limited. Ideally, what is needed is a sleepiness "biomarker"—an objective, physiological measure that reflects those changes in brain state/functioning that underlie sleepiness-mediated performance deficits. It is important to note that the ideal sleepiness biomarker should reflect the sleepy brain's

Figure 1—Mean (SE) response speed on the Psychomotor Vigilance Task across three phases of the study: Baseline (B), sleep restriction (SR - 3 hours TIB per night, and recovery sleep (R - 8 hours TIB per night).
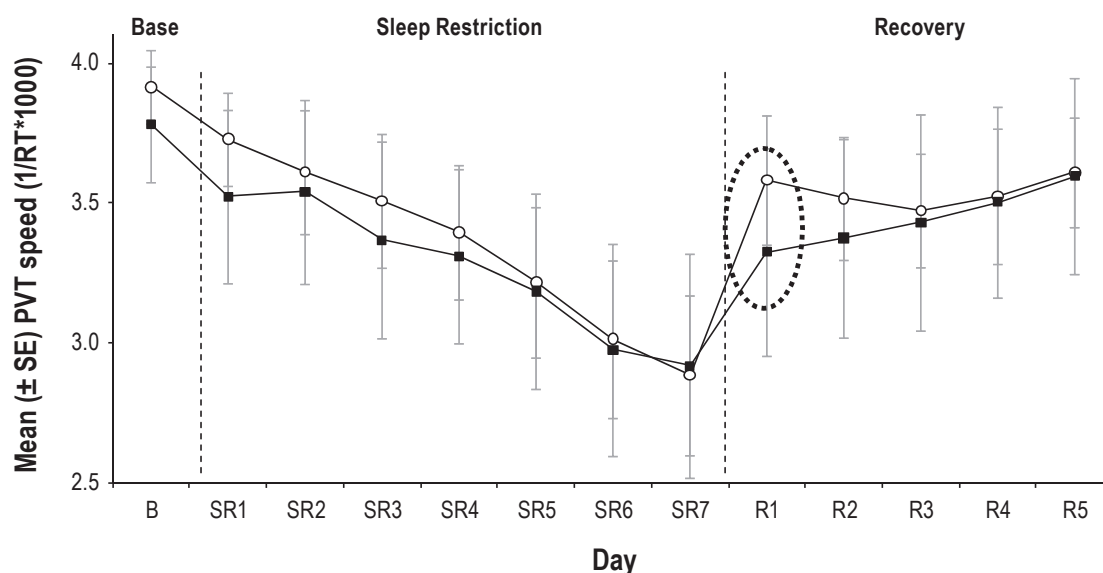
Shaded squares represent the group that maintained habitual nightly sleep times for 1 week prior to baseline. Open circles represent the group that extended nightly sleep times (to 10 hours TIB) for 1 week prior to baseline. Following the first night of recovery sleep (day R1), group differences in performance were still evident (dashed oval). TIB, Time in Bed; B, Baseline; SR, Sleep Restriction; PVT, Psychomotor Vigilance Test. Adapted from Rupp et al., 2009.[6]

Figure 2—Mean (SE) subjective sleepiness ratings on the Stanford Sleepiness Scale across three phases of the study: Baseline (B), sleep restriction (SR - 3 hours TIB per night, and recovery sleep (R - 8 hours TIB per night).
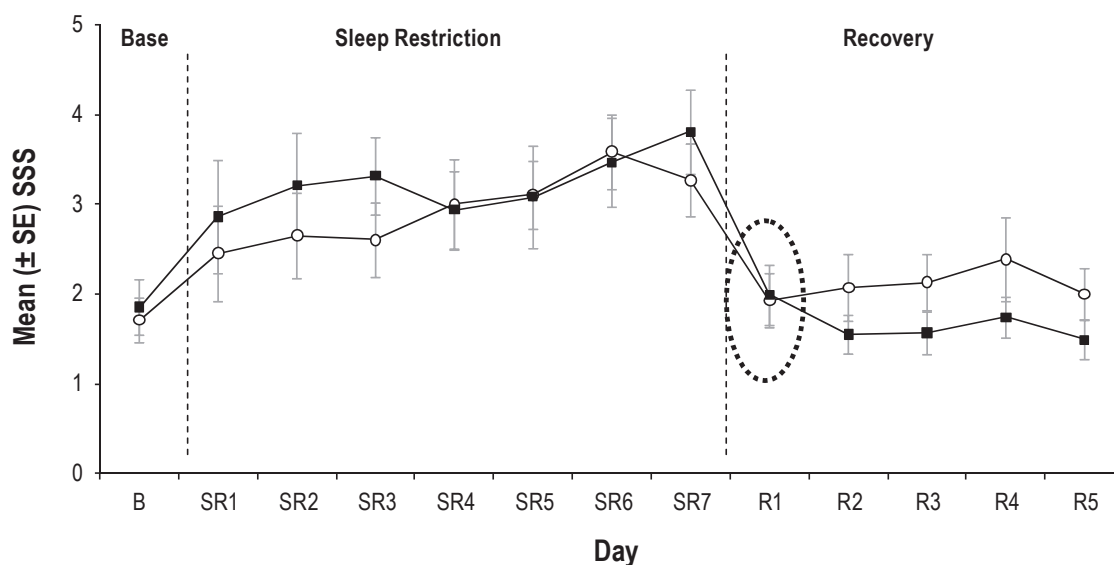
Shaded squares represent the group that maintained habitual nightly sleep times for 1 week prior to baseline. Open circles represent the group that extended nightly sleep times (to 10 hours TIB) for 1 week prior to baseline. Following the first night of recovery sleep (day R1), there were no group differences in subjective sleepiness ratings (dashed oval). TIB, Time in Bed; B, Baseline; SR, Sleep Restriction; SSS, Stanford Sleepiness Scale. Adapted from Rupp et al., 2009.[6]

capacity to actually *perform relevant tasks*, rather than (for example) merely reflecting the extent of the brain's physiological sleep deficit. This is because, from a practical standpoint, what ultimately matters—i.e., the reason that objective biomarkers of sleepiness are desirable—is not so that the extent to which the brain "hungers for" sleep can be quantified. Rather, what matters is the extent to which such biomarkers reflect an individual's ability to actually function effectively in real-world situations. And the relationship between "level of sleep debt" and "real-world functioning" is not straightforward.

The lack of concordance between sleep need and performance is perhaps best illustrated by the phenomenon of "sleep inertia." Sleep inertia refers to the profound deficits in alertness and performance that are universally experienced immediately upon awakening from sleep—performance deficits that rapidly reverse over the first 15-20 min of continued wakefulness.[7] Clearly, sleep inertia is a brain state in which both the subjective and objective manifestations of sleepiness are divorced from the brain's actual level of sleep need.[8] Logically, the brain's need for sleep should be at a nadir immediately upon awakening, and accrue during subsequent time spent awake. Therefore, if sleepiness was merely and invariably a direct reflection of the brain's underlying physiological need for sleep, the first few minutes of wakefulness would be characterized by optimal alertness and performance. Since this is not the case, it is clear that a biomarker that perfectly reflects "physiological sleep need" would not necessarily be useful for predicting sleepiness and its attendant performance deficits.

This dissociation between "apparent sleep deficit" and performance is also illustrated by the fact that there are considerable individual differences in the extent to which sleep loss impacts performance on a variety of tasks—differences that cannot be explained by, for example, variations in habitual nightly sleep duration.[9] The physiological basis of these individual differences in resilience in the face of sleep loss is not known, but findings from functional brain imaging studies suggest one strong possibility: that the performance deficits that characterize sleepiness result specifically from deactivation of the prefrontal cortex (PFC); and individual differences in the ability to sustain performance during sleep loss vary as a function of the sleepy brain's ability to recruit and utilize resources from non-PFC regions.

Functional brain imaging studies have revealed that sleep deprivation is characterized by reductions in brain metabolism, with the greatest reductions evident in thalamus, inferior parietal/superior temporal cortices, anterior cingulate, and the prefrontal cortices (PFC).[10] Similar studies performed during the sleep inertia period have revealed reduced activity in PFC—although activity in subcortical regions (including the thalamus) was actually elevated at this time. Because both sleep deprivation and sleep inertia result in objective and subjective sleepiness, and because deactivation of the PFC is the most salient finding common to both states, it can be concluded that sleepiness is exclusively a function of the level of (de)activation in the PFC.[11] Consistent with this conclusion, researchers have recently been able to successfully predict—based on what is currently known about the differential functions mediated by various PFC subregions (e.g., see Damasio[12]—specific sleepiness-induced deficits in olfaction,[13] moral judgment[14] and decision-making,[15] to name but a few).

Together, the fact that 1) both sleep loss and sleep inertia are characterized by PFC deactivation, and 2) performance deficits can successfully be predicted based what is known about the specific functions of affected PFC regions, it is reasonable to postulate that "PFC deactivation" constitutes a satisfactory "sleepiness biomarker"—but, again, it is not necessarily a good biomarker of "sleepiness-related *performance deficits*."

This is illustrated by the findings of Drummond et al.[16] First, consistent with prior research, they showed that performance on a verbal learning task was associated with activation of the temporal cortex during the normal (well-slept, alert) state. Also consistent with prior research, they found that this brain region was relatively deactivated following 35 h of sleep deprivation. However, they then compared the scans of subjects who were relatively resilient to the effects of sleep loss (i.e., who were better able to maintain performance on the verbal learning task following 35 h of sleep deprivation) vs. those subjects who were most vulnerable to sleep loss (i.e., whose performance on the verbal learning task was relatively impaired following 35 h of sleep deprivation). This comparison revealed something surprising—that during performance of the verbal learning task following sleep loss the resilient subjects showed relative activation in the parietal cortex, a region that had *not* been activated during performance of this task in the well-slept (non-sleep-deprived) state. This suggests that although sleep loss results in regional brain deactivation, and although such deactivation may underlie deficits in specific aspects of cognitive performance, such performance deficits can be mitigated by recruiting brain resources from *other* cortical regions—i.e., regions not typically activated during performance of the task in the well-slept/normally alert state. Thus, although sleep loss results in regional cortical deactivation, and a measure of regional cortical deactivation might therefore constitute an adequate biomarker of increased sleepiness, *per se*—such a measure would *not* constitute an adequate biomarker of "sleepiness-related *performance deficits*" since this measure only reflects the fact that the brain is sleepy—it does not reflect the ability of the brain to *cope* with that sleepiness.

In the absence of biomarkers that reflect the *capacity of the brain to perform despite sleepiness* (which, to reiterate, would be considerably more useful than biomarkers that merely reflect sleepiness, *per se*), the next best thing would be a "behavior biomarker of sleepiness"—that is, a behavioral measure that is not only sensitive to sleep loss/sleepiness, but that actually reflects some aspect of brain physiology that is *specific* to sleepiness. Such a measure would not only reflect performance deficits (which usually is, or at least should be, the ultimate goal) but it would also reflect the physiological basis of the performance deficit (i.e., sleepiness). Therefore, such a measure would provide information useful for predicting operational performance capacity and, by virtue of providing a "differential diagnosis" of sleepiness-induced performance deficit vs. all other possible causes of performance deficits, it would inform decisions regarding appropriate interventions to restore performance.

What might a behavioral biomarker of sleepiness look like? Reaction time (RT) can be used to illustrate.

RT has previously been shown to be sensitive to a variety of factors, including (to name just a few) sleep loss,[17] alcohol,[18] various medications,[19] and "time on task."[20] RT has also been shown to be exquisitely sensitive to mild traumatic brain injury (mTBI) events—with such events reliably producing extended mean RTs.[21] Thus, it has low specificity. The latter problem is illustrated in the equation below (adapted from Warm[22]:

$$Pv = \frac{f(M, S, U, B, C)\, E}{BS}$$

in which *Pv* is mean RT performance on a vigilance test. *Pv* is a function of several factors including "signal modality"

(*M*, the sensory modality of the signal being tested); salience (*S*, the meaningfulness of the signal being presented, e.g., the subject's name vs. a neutral tone); the uncertainty of the signal (*U*, which depends on the mathematical likelihood of a signal presentation; the "background events density" (*B*, for example, the density of "competing signals" that are presented, and must be distinguished from the target signal); and "signal complexity" (*C*, reflecting, for example, the amount of mental processing required to identify the signal).

To make matters more complicated, in addition to interacting with each other, each of these factors can themselves interact with a potentially infinite variety of environmental factors (*E*, e.g., environmental noise, ambient temperature, vibration)—at least as potential distracters, if not as factors that more directly influence vigilance performance.

Also, the influence of each of these factors is potentially mediated by "brain state" (*BS*)—i.e., the changeable capacity of the brain to perceive, process, and react to all of the factors that reside in the numerator of the depicted equation. *BS* could refer specifically to the brain's level of sleepiness, intoxication, hypoxia, or mTBI event-induced disruption of white matter function, etc.

It is therefore clear that, although sensitive, the potential utility of mean RT is limited by its lack of specificity. However, it might be hypothesized that adequate specificity could achieved by analyzing the RT signal in a manner that more directly reflects the primary, underlying physiological basis of the performance deficit. For example, it may prove possible to differentiate RT performance deficits caused by sleepiness from RT performance deficits caused by mild traumatic brain injury (mTBI). This is because the mean RT deficits resulting from sleep loss are thought to reflect "state instability"[23]– a moment-to-moment waxing and waning of alertness that produces a distribution of RTs that is characterized by an extended "right tail" (i.e., some "normal" RTs and many more "long" RTs than are evident in normally-alert subjects). In contrast, mTBI events are thought to produce functional disruption of the long, white-matter tracts involved in psychomotor (e.g., RT) performance. So, by virtue of disrupting the pathways that mediate RT performance, mTBI events *should* (unlike sleepiness) eliminate the possibility of *any* fast/normal RTs. That is, mTBI events should effectively "cut off" the left (fast) end of the tail of the distribution of RTs.

Thus, although mTBI and sleep loss both result in similarly extended mean RTs, the shape of the distributions of RTs that characterize mTBI vs. sleep loss *should* be quite different—differences that should, accordingly, be measurable using statistics designed to assess the shapes of such distributions. One such statistic, the Shannon-Jensen Divergence Metric (SJDM), is currently being developed by Army researchers (Rajaraman et al., manuscript in preparation). It is anticipated that future work will include validation of the SJDM, and development of similar "behavioral biomarkers."

REFERENCES

1. Webb, WB. The cost of sleep-related accidents: A reanalysis. *Sleep* 1995;18:276–280.
2. Leger D. The cost of sleep-related accidents: a report for the National Commission on Sleep Disorders Research. *Sleep* 1994;17:84–93.
3. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study. *J Sleep Res* 2003;12:1–12.
4. Schmidt EA, Schrauf M, Simon M, Fritzsche M, Buchner A, Kincses WE. Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accid Anal Prev* 2009;41:1087–93.
5. Balkin TJ, Rupp T, Picchioni D, Wesensten NJ. Sleep loss and sleepiness: current issues. *Chest* 2008;134:653–60.
6. Rupp TL, Wesensten NJ, Bliese PD, Balkin TJ. Banking sleep: Realization of benefits during subsequent sleep restriction and recovery. *Sleep* 2009;32:311–21.
7. Lubin A, Hord DJ, Tracy ML, Johnson LC. Effects of exercise, bedrest and napping on performance decrement during 40 hours. *Psychophysiology* 1976;13:334–9.
8. Balkin TJ, Badia P. Relationship between sleep inertia and sleepiness: cumulative effects of four nights of sleep disruption/restriction on performance following abrupt nocturnal awakenings. *Biol Psychol* 1988;27:245–58.
9. Van Dongen HP, Baynard MD, Maislin G, Dinges DF. Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. *Sleep* 2004;27:423–33.
10. Thomas M, Sing H, Belenky G, et al. Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity. *J Sleep Res* 2000;9:335–52.
11. Balkin TJ, Braun AR, Wesensten NJ, et al. The process of awakening: a PET study of regional brain activity patterns mediating the re-establishment of alertness and consciousness. *Brain* 2002;125:2308-19.
12. Damasio AR. On some functions of the human prefrontal cortex. *Ann N Y Acad Sci* 1995;769:241–51.
13. Killgore,WD, McBride SA. Odor identification accuracy declines following 24 h of sleep deprivation. *J Sleep Res* 2006;15:111–6.
14. Killgore WD, Killgore DB, Day LM, Li C, Kamimori GH, Balkin TJ. The effects of 53 hours of sleep deprivation on moral judgment. *Sleep* 2007;30:345–52.
15. Killgore WD, Balkin TJ, Wesensten NJ. Impaired decision making following 49 h of sleep deprivation. *J Sleep Res* 2006;15:7–13.
16. Drummond SP, Brown GG, Gillin JC, Stricker JL, Wong EC, Buxton RB. Altered brain response to verbal learning following sleep deprivation. *Nature* 2000;403:655–7.
17. Balkin TJ, Bliese PD, Belenky G, et al. Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *J Sleep Res* 2004;13:219–27.
18. Warren JW. The effect of pure alcohol on the reaction time, with a description of a new Chronoscope. *J Physiol* 1887;8:311–48.
19. McLay R, Spira J, Reeves D. Use of computerized neuropsychological testing to help determine fitness to return to combat operations when taking medication that can influence cognitive function. *Mil Med* 2010;175:945–6.
20. Wesensten NJ, Belenky G, Thorne DR, Kautz MA, Balkin TJ. Modafinil vs. caffeine: Effects on fatigue during sleep deprivation. *Aviat Space Environ Med* 2004;75:520–5.
21. Sosnoff JJ, Broglio SP, Ferrara MS. Cognitive and motor function are associated following mild traumatic brain injury. *Exp Brain Res* 2008;87:563–71.
22. Warm J. Vigilance and target detection. In: Workload transition: Implications for individual and team performance, National Academy of Sciences, 1993/2000.
23. Doran SM, Van Dongen HP, Dinges DF. Sustained attention performance during sleep deprivation: evidence of state instability. *Arch Ital Biol* 2001;39:253–67.

ACKNOWLEDGMENTS

CORRESPONDENCE

Address correspondence to: Thomas J. Balkin, Ph.D., Department of Behavioral Biology, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Suite 2A26, Silver Spring, MD 20910; Tel: (301) 319-9350; Fax: (301) 319-9979; E-mail: thomas.balkin@us.army.mil

DISCLOSURE STATEMENT

Dr. Balkin has indicated no financial conflicts of interest.