SHIP VIA: Odyssey

Odyssey Address: 206.107.42.197
ARIEL Address: 130.108.121.58

Email: fill@www.libraries.wright.edu

Borrower: OHUDAC
Lending String:

Journal Title: Journal of clinical neurophysiology ; official publication of the American Electroencephalographic S
ISSN: 0736-0258
Volume: 23 Issue: 1
Month/Year: 2006
Pages: 50-8

Article Author: Bonnet M
Article Title: ACNS clinical controversy; MSLT and MWT have limit

ILL Number: 29241822

Patron: TN; 187776
Request Date:

Call #: fill www
Location:

MAXCOST: $20.00
Borrowing Notes: GMR-RL, PLEASE ARIEL IF POSSIBLE.

Ship To:
Wright State University
Dunbar Library ILL (OhioLink #217)
3640 Colonel Glenn Highway
Dayton, OH 45435-0001

Phone: 1.937.775-2289
Fax: 1.937.775-2356

Please send resend requests to hslill@osu.edu or fax (614) 292-5717 or call (614) 292-4894.

NOTICE: This material may be protected by COPYRIGHT LAW (Title 17 U.S. Code).
ACNS Clinical Controversy: MSLT and MWT Have Limited Clinical Utility

Michael H. Bonnet

The Multiple Sleep Latency Test (MSLT) and the Maintenance of Wakefulness Test (MWT) are two commonly used laboratory-based objective tests to measure sleepiness and alertness, respectively. Data suggest both are extremely sensitive tests when measuring the effects of sleep deprivation within subjects, but are less sensitive for confirming sleepiness and response to treatment in groups of patients with different sleep disorders. Inconsistent and even sometimes paradoxical test results may be partly explained by data that show the MSLT and MWT are not selectively sensitive to either sleepiness or alertness, but sensitive to both the sleep and the arousal systems. Sleep latencies seen on both the MSLT and MWT are affected to varying degrees by a myriad of internal and external influences that can alter what we would prefer each test to show. If we continue to use these tests to measure sleepiness or alertness in patients with different sleep disorders, we need to understand more about the nature and impact of different sources of internal and external arousal so that we can better control the test environment. Improved understanding of the determinants of sleep onset is essential because excessive sleepiness has important consequences for both individuals and society.

Key Words: MSLT MWT Sleepiness Narcolepsy Arousal Sleep disorders.

(J Clin Neurophysiol 2006;23: 50 58)

The Multiple Sleep Latency Test (MSLT) was originally developed in 1977 to measure sleepiness in young normal subjects involved in sleep deprivation experiments (Carskadon and Dement, 1977, 1979). Following two nights of total sleep deprivation, six young healthy adults had opportunities to nap at 2-hour intervals across a day. If the subject fell asleep during the nap, the subject was awakened after one minute of sleep. Each nap opportunity was terminated within twenty minutes if the subject did not fall asleep. Sleep latency was measured from lights out on a particular nap to the first minute of stage 1 sleep. The subjects fell asleep more quickly as their time awake increased, and this provided face validity that sleep latency could be used to measure sleepiness. When used within this design framework, the MSLT has been shown to be sensitive to graded amounts of sleep loss within-subsjects, especially when comparing baseline and sleep deprivation periods (Bonnet, 2000; Carskadon and Dement, 1981; Roehrs et al., 1994).

Within a year of its original design, sleep researchers began using the MSLT to measure sleepiness in patients with narcolepsy (Mitler et al., 1979; Richardson et al., 1978) and then extended its use to objectively assess sleepiness in patients with other sleep pathologies (sleep apnea, periodic limb movements, insomnia) and to assess the effects of treatment on sleepiness and alertness. Using the MSLT to assess sleepiness in these patients and paradigms occasionally yielded what initially seemed to be unpredicted, marginal, or even paradoxical test results. These test results would seem to challenge the validity of the MSLT, but actually have prompted research to further understand what the MSLT is actually measuring.

An exhaustive review of the clinical use of the MSLT and the Maintenance of Wakefulness Test (MWT) has been recently published by a Task Force organized by the American Academy of Sleep Medicine (AASM) Standards of Practice Committee (Arand et al., 2005). The AASM Standards of Practice Committee used this evidence-based review, which included meta-analyses of extensive MSLT data, to revise practice parameters for clinical use of the MSLT and MWT (Littner et al., 2005). I was a member of a Standards of Practice Committee Task Force that prepared the review paper (but this presentation does not necessarily reflect the views of that committee or the AASM). The purpose of this article is not to repeat that review, but instead to focus on problems that occur when the MSLT and MWT are used to assess sleepiness and alertness in patients with different sleep disorders and in less controlled laboratory environments than the historical sleep deprivation paradigm for which they were designed. I hope discussion of these issues leads to a better understanding of how sleep tendency is measured and serves as a means to direct future research.
TESTS OF SLEEPINESS AND ALERTNESS

A number of subjective and objective tests have been used to measure sleepiness. An early commonly used subjective measure of sleepiness is the Stanford Sleepiness Scale (SSS). First developed in 1973, the SSS asks subjects to rate their degree of sleepiness at a single moment in time from seven descriptors ranging from “feeling active and vital; alert; wide awake” to “almost in reverse; sleep onset soon; losing struggle to remain awake” (Hoddes et al., 1973). Patients are often asked to complete a SSS just before each nap opportunity during a MSLT or MWT. The SSS is fast and easy to use. However, there are no reference values for the SSS, and it has not been validated with other physiologic measures. Early on, the reliability of the SSS in chronically sleepy patients was questioned (Carskadon and Dement, 1982; Roth et al., 1980). Because the SSS has seven discrete descriptions, the response spectrum is not necessarily linear, and some subjects or patients complain that their state fits into more than one category.

Another commonly used subjective measure of sleepiness is the Epworth Sleepiness Scale (ESS). First developed in 1991 as a trait measure of tendency to fall asleep in specific situations (Johns, 1991), the ESS is commonly used in clinical populations. The ESS provides a subjective measure of how sleepiness interferes with an individual’s common daily activities. It consists of a simple self-administered questionnaire asking patients to rate the likelihood (from 0 to 3) of dozing or falling asleep in eight sedentary situations. For each situation, patients estimate the likelihood of their falling asleep on a four-point scale (0 = never doze to 3 = high probability). Situations range from lying quietly in the afternoons when circumstances permit napping, to sitting talking with someone. Scores from each of the eight situations yield a total score ranging from 0 to 24. Johns used healthy medical students to establish a normal ESS value of 7.6 (Johns, 1991). He thought a score of 0 to 10 was within the range of normal, a score of 10 was the upper limit of normal alertness, and scores greater than 12 were consistent with pathologic sleepiness. The higher the ESS score, the greater the sleepiness. An ESS score of ≥15 suggests marked sleepiness.

Unfortunately, the SSS and ESS are less simple than they seem. Many people are unaware how sleepy they are, in part because their sleep disorders may be chronic or have developed insidiously. Others have difficulty distinguishing sleepiness from fatigue; some say they need to sleep, but given the opportunity cannot. The MSLT was developed by Carskadon and Dement (1977) to deal with some of the problems inherent in the SSS.

When the MSLT is done for clinical purposes, subjects or patients are allowed four or five opportunities to nap at 2-hour intervals across their usual “day” beginning 1.5 to 3 hours after awakening. Subjects are tested in the sleep laboratory, lying in a quiet, dark, temperature-controlled bedroom with their street clothes loosened. During each nap opportunity, they have 20 minutes to “try and fall asleep.” If sleep occurs in a given nap opportunity, the patient is allowed to sleep 15 minutes (of clock time, not sleep time) to observe whether REM sleep appears inappropriately early after the onset of sleep. REM sleep in adults normally first appears approximately 90 minutes after sleep onset (range 70 to 110 minutes). The inappropriately early appearance of REM sleep within 15 minutes after sleep onset is called a “sleep-onset REM period” (SOREMP). The MSLT must be performed immediately following PSG recorded during the individual’s major sleep period to document no other significant sleep disorder is present to explain the sleepiness complaint (Littner et al., 2005). At least 6 hours of sleep needed to be recorded in the prior PSG if the results of the MSLT are to be used to confirm a diagnosis of narcolepsy.

Sleep in polysomnography is typically scored (staged) by convention based on 30-second periods (epochs) from “lights out” to “lights on”. Sleep onset in the MSLT is typically determined as the time from lights out to the first epoch of any stage of sleep, including stage 1 non-REM sleep. Sleep onset is further defined as the first epoch of greater than 15 seconds of cumulative sleep in a 30-second epoch. A mean sleep latency (MSL) is calculated as the arithmetic mean of all four or five nap opportunities, counting those nonsleeping naps as a 20-minute sleep latency for the MSLT. REM latency is taken as the time of the first 30-second epoch of REM sleep regardless of the intervening stages of wakefulness or sleep (Littner et al., 2005). If the MSL is for suspected narcolepsy, a fifth nap is required if only one SOREMP occurs in the first four naps. A fifth nap is not necessary if a short mean sleep latency of ≤5 minutes and two or more SOREMPs occur among the first four naps.

The MSLT, when used as it was designed to test sleep deprivation under controlled experimental conditions, has face-value validity as a quantification of sleepiness since it correlates with subjective reports, varies with prior sleep time, and reflects the sedating and alerting effects of medication (Arand et al., 2005). The MSLT also has been shown to have high test-retest reliability, as high as 0.97 in normal subjects over a period of 4 to 14 months (Zwyghuizen-Dorenbos et al., 1988). However, the reliability of the MSLT falls when fewer than four nap opportunities are performed: 0.85 for three tests and 0.65 for two. The MSLT can be scored with high interrater reliability in clinical populations (Drake et al., 2000, Benbadis et al., 1995). Studies have shown that the MSL mean sleep latency and SOREMPs can be scored with high interrater reliability: r values range from 0.85 to 0.90 for MSL and are 0.91 for more than one SOREMP (Drake et al., 2000).

What are “normal values” for the MSLT? The current AASM practice parameter guidelines (using pooled data from normal subjects across all ages with sleep onset defined as the first epoch of any sleep stage) reported normal mean sleep latency values of 10.4 ± 4.3 minutes for a four-nap MSLT and 11.6 ± 5.2 minutes for a five-nap MSLT (Littner et al., 2005). Notice the first problem with using the MSLT outside the within-subject sleep deprivation paradigm for which it was developed: based on two standard deviations from the mean, 95% of the values from control populations on the four nap MSLT would fall between 1.8 to 19 minutes and 1.2 to 20 minutes for the five-nap MSLT. Such a wide range of
“normal values” for the MSLT suggests the test does not discriminate well between clinical and control populations.

Sleep specialists soon discovered further limitations of the MSLT when they tried to confirm that sleepiness had lessened in patients following treatment. Roth et al. (1980) found some severely sleepy patients often failed to show normalization of their MSLT results following seemingly effective treatment responses. Other criticisms of the MSLT are the floor effect where it is unclear if the scale remains linear (i.e., is the difference between a 2-minute latency and a 1-minute latency the same as the difference between a 20-minute latency and a 19-minute latency—linear—or the same as the difference between a 20-minute latency and a 10-minute latency?); the ceiling effect where limiting nap opportunities to 20 minutes may limit the ability to discriminate between and within more alert individuals; and the fact that the sleep laboratory environment is not representative of the workplace (Arand et al., 2005).

The Maintenance of Wakefulness Test (MWT) was developed in response to some of these criticisms of the MSLT. First developed in 1982 by Mitter et al. (1982a), the MWT was designed to measure the ability of an individual to remain awake in soporific conditions. It consists of four naps performed at 2-hour intervals with the first beginning about 1.5 to 3 hours after the patient’s usual wake-up time. During each trial, the patient sits in a chair in a dimly lit room and is given the same instruction, “Please sit still and remain awake for as long as possible. Look directly ahead of you, and do not look directly at the light.” Patients are not allowed to use extraordinary measures to stay awake such as “slapping the face or singing” (Littner et al., 2005). Trials end after 40 minutes if no sleep occurs, or after unequivocal sleep, defined as three consecutive epochs of stage 1, or one epoch of any other stage of sleep (Littner et al., 2005).

The MWT is most often done to assess the ability of a previously sleepy patient to stay awake when treated for the person’s condition. MWT is sometimes requested to assess the ability of an individual to stay awake, most often when the individual’s occupation involves public transportation or issues regarding public (and patient) safety. The MWT is occasionally performed on patients with narcolepsy or idiopathic hypersomnia who have been treated with stimulant medications to assess response to treatment, most often as part of drug study protocol. Mitter et al. (1982a) showed the MWT was able to demonstrate significant pre- and posttreatment differences in patients with excessive sleepiness (Mitter et al., 1982a).

There are limited normative data for the MWT. The AASM SPC Task Force culled normative data from five articles reporting MWT. The mean sleep latency for normal controls on a four-trial 40-minute MWT was 30.4 ± 11.2 minutes using latency to the first epoch of sleep. The upper limit of the 95% confidence interval was 40.0 minutes.

So what are “abnormal values” for a MWT? One study (Mitter et al., 1998) that used an MWT protocol with 20-minute trials found only 1% of 530 patients with narcolepsy were able to remain awake during all four 20-minute trials, compared with approximately 55% of controls (tested by them in another study). The MSL among the narcoleptics on their MWT20 was 6.0 ± 4.8 minutes. Of note, 15% of study subjects with narcolepsy fell within their 12-minute lower limit of normal for the mean sleep latency (which corresponds to the fifth percentile for normals). Patients who had narcolepsy with cataplexy were least able to remain awake. Patients with narcolepsy are able to stay awake for an average of only approximately 6 minutes compared with 19 minutes in normal controls (Mitter et al., 1982a).

Criticism of the MWT begins with the difficulty of interpreting MWT results with limited normative data and comparing MWT results from laboratories that have used widely varying testing protocols. The AASM SPC revised practice parameters contain recommendations for performing MWT and recommend a 40-minute nap protocol (Littner et al., 2005). Unfortunately, there is little data that show the results of the MWT done in a sleep laboratory directly correlate with sleep/wake behavior in the “real world.” Nor are there any data correlating MWT findings and risk for adverse consequences of sleepiness in daily life situations.

How good is the agreement between these different subjective (ESS and SSS) and objective (MSLT, MWT) measures of sleepiness and alertness? Because the MSLT and MWT measure the same thing (latency to sleep onset), one would expect the results to be correlated even if not similar in magnitude. This is not true. Only low (but usually statistically significant) correlations have been found between MWT and MSLT results in the same subjects (ranging from r = 0.29 to r = 0.52; Bonnet and Arand, 2001; Sangal et al., 1992, 1997). Such correlations indicate that the relationship accounts for only 10% to 25% of the total variance. Roth et al. (1980) showed there was frequent disagreement between the MSLT and SSS results. Chervin et al. (1997) found only a weak correlation between the ESS and mean sleep latencies on the MSLT (Chervin et al., 1997). Johns (1992) found only weak correlations between the MSLT and treatment responses among patients with obstructive sleep apnea and that a number of the questions on the ESS did not correlate with the sleep latency. After analyzing a large number of studies comparing the ESS and MSLT, Johns (2000) found the average correlation between them was low (r = 0.30).

As all of these tests (MSLT, MWT, ESS, SSS) have demonstrated some measure of test-retest reliability, the lack of agreement among tests implies that a simple single sleepiness dimension does not exist in clinical populations. Moreover, the lack of substantial correlations between tests implies that judgments of sleepiness may be test specific, indicating that there is more than a single component to sleepiness.

An even clearer understanding of the limitations of the MSLT and MWT in clinical practice occurs when we try to use them to assess sleepiness and alertness in patients with narcolepsy, obstructive sleep apnea, and insomnia.

NARCOLEPSY

The MSLT was first used as a clinical test in patients with narcolepsy, primarily to document their abnormally short REM sleep onset (SOREM) in daytime naps. The AASM SPC Task Force identified thirteen papers judged...
relatively free of inclusion bias to calculate mean sleep latency (MSL) values in patients with narcolepsy (Arand et al., 2005). Four papers that directly compared a total of 39 patients with narcolepsy with 40 controls had a MSL value for thenarcolepsy patients which was significantly shorter than for controls (3.0 ± 3.1 minutes versus 10.5 ± 4.6 min, P < 0.001). These MSL values are in close agreement with those reported for a larger group of narcolepsy patients without controls (n = 245) and a comparable large group of normals (3.1 ± 2.9 minutes versus 10.3 ± 4.3 minutes). Note once again that if these control values are used as population means and standard deviations, then the lower border of normal on the MSLT (mean ± 2 SDs) would be 1.3 minutes. The implication of these data is that a diagnosis of excessive sleepiness two standard deviations outside the population norm really requires a MSL of 1.3 minutes or less. An abnormally short onset of REM sleep within less than 15 minutes after sleep onset is a hallmark in the definition of narcolepsy. One might think that the presence of two or more SOREMPs coupled with an abnormally short MSL might be sufficiently reliable for distinguishing patients with narcolepsy from patients with other causes of excessive sleepiness or normals. Arand et al. (2005), using data culled from 10 articles, calculated that two or more SOREMPs had a sensitivity of 0.78 and a specificity of 0.93 for the diagnosis of narcolepsy. Although these figures seem impressive, SOREMPs also occur in normal subjects. Bishop et al. (1996) found 17% of 139 normal subjects had two or more REM onsets in their MSLT and a mean sleep latency of 6.2 minutes. Two SOREMPs on a MSLT are most often associated with narcolepsy-cataplexy, but similar SOREMPs can also be seen on occasion in sleep apnea, drug withdrawal, severe depression, and myotonic dystrophy (Miller et al., 1982b).

Such high SOREMP incidence values in concert with the broad range of MSL values in the “pathologic range” in normal subjects cast significant doubt on a diagnosis of narcolepsy without concurrent cataplexy. However, the diagnosis of cataplexy is usually based only subjective report (i.e., does not require objective laboratory confirmation), and patients can therefore false-report what they think is cataplexy (“my knees go weak when I get excited”) as well.

An example of how the MSLT can show a short MSL and two or more SOREMPs and result in the misdiagnosis of narcolepsy was recently published by Janjua et al. (2003). They reported a 19-year-old woman who was diagnosed as narcoleptic based on a clinical history of sleepiness since junior high school without any other clinical symptoms of narcolepsy. Her first PSG followed by a MSLT showed a MSL of 3.4 minutes and two SOREMPs. However, she did not respond well to stimulant medication and was referred to their sleep center for reevaluation.

At that time, her PSG and MSLT were repeated for four consecutive nights/day, and she was allowed to sleep as long as she desired on each night. Her sleep time on these four nights ranged from 10.2 to 13.8 hours. She had MSL of 6.8, 9.8, 14.2, and 18.1 minutes and no SOREMPs on her MSLT for the 4 days after extended nighttime sleep. It was concluded that she was a normal individual with a long sleep need. The misdiagnosis of narcolepsy occurred from testing her in a state of chronic sleep deprivation (which for her was sleeping about 8 hours per night). Because partial sleep deprivation is common and may predispose to decreased sleep latency and increased REM onset, MSLT data must be viewed “within the context of the individual patient history and as a part of other medical information and testing” (Arand et al., 2005). This certainly implies limitations for the use of MSLT as the objective test for narcolepsy.

**OBSTRUCTIVE SLEEP APNEA**

Sleep clinicians routinely search for historical reports of sleepiness in sedentary situations when evaluating patients with suspected obstructive sleep apnea (OSA), and one of the major treatment goals is improvement in daytime alertness. A long-time assumption has been that the severity of the OSA should correlate with the degree of sleepiness. However, this has not been clearly demonstrated in the literature. For example, Sauter et al. (2000) found no difference in MSL on MWT between patients with moderate or severe OSA. They also found that severe OSA patients self-rated themselves more alert on the Epworth Sleepiness Scale than patients with moderate OSA.

Patients with untreated OSA are more likely to have motor vehicle accidents (George and Smiley, 1999). They pose as great a risk to road safety as drinkers (Dawson and Reid, 1997). Again, one would presume that patients with more severe OSA would be sleepier and have more accidents. Unfortunately, OSA patients who actually had car accidents did not have shorter MSL on their MSLT, nor did they self-report themselves sleepier on ESS than OSA patients who had not had accidents (Aldrich, 1989; Young et al., 1997). Aldrich found mean MSL on MSLT in OSA patients who had accidents compared to those who did not was 8.0 minutes versus 7.8 minutes (Aldrich, 1989).

Another false belief is that MSL values on the MSLT and MWT should normalize after OSA patients are treated. Significantly longer MSL on the MSLT after effective OSA treatment have not always been found (Roth et al., 1980). Only two of four placebo-controlled studies found improved (longer MSL) MSLT values in OSA patients effectively treated with CPAP compared with those treated with placebo (Barbe et al., 2001; Engleman et al., 1994, 1997, 1998). Arand et al. (2005) recently calculated weighted MSL from these four studies and found weighted MSL on the MSLT in OSA patients effectively treated with CPAP was 9.8 ± 4.5 minutes compared with 8.3 ± 4.7 minutes on placebo. Engleman et al. (1999) found no significant improvement in MSL on MWT when comparing CPAP treatment with placebo.

The inability to find strong correlations between OSA severity indices and MSLT values and only marginal improvement in MSLT after seemingly effective CPAP treatment suggest that something is missing from our understanding of either the sleep disturbance associated with OSA or with the MSLT measure. Recent work by Vgontzas et al. (1998) showed reduced MSL in obese individuals without
OSA, and this suggests that other factors, such as metabolic disorders, may contribute to sleepiness, irrespective of OSA.

**INSOMNIA**

The MSLT test results in patients with primary insomnia are particularly paradoxical and pose the greatest challenge to our traditional understanding of what the test should show. These patients have chronically decreased sleep time that should cause partial sleep deprivation and short sleep latencies on the MSLT. However, MSLT studies in patients with psychophysiological insomnia have consistently shown just the opposite: despite decreased amounts of nighttime sleep, significantly longer MSL are found on their MSLT (Bonnet and Arand, 1995, Haynes et al., 1985, Schneider-Helmert, 1987, Stepanski et al., 1988). Average MSL latencies from those MSLT studies were 11.7 minutes for 55 control subjects and 14.5 minutes for 88 patients with primary insomnia.

These counterintuitive MSLT results are not limited to patients with insomnia. Arand et al. (2005) compared MSL from 451 normal subjects (from 11 different studies combined) who slept longer than 435 minutes on the night before their MSLT to 397 subjects (from 9 studies) who slept less than 435 minutes on the preceding night. The MSL was significantly shorter when subjects slept longer the night before (9.8 ± 4.8 minutes) compared with those who slept less (11.2 ± 4.1 minutes). Nothing in the traditional understanding of the MSLT as a measure of the sleep system can explain these consistent findings of longer MSLT values following both acute and chronically reduced sleep at night.

**WHY ARE MSLT FINDINGS PARADOXICAL?**

Whenever data do not agree with one’s presumptions, the most common response is to blame the data (“college students are not normal”) or shoot the messenger (“the MSLT is not a valid test”). These initial responses are almost always incorrect. More often, unrecognized factor(s) or uncontrolled variable(s) account for a significant amount of the experimental variance observed and have influenced the test results. These issues are apparent in the problems discussed and point to areas that require new understanding.

**OTHER DETERMINANTS OF SLEEP LATENCY**

**Trait Level of CNS Arousal**

Perhaps the most problematic issue with concept of the MSLT as a simple measure of sleep tendency is the consistent inability of the MSLT to deal with the problem of insomnia. Based on their disorder and by definition, these patients should be sleep deprived, and they should fall asleep rapidly when given a nap opportunity. The fact that these patients do not have short MSLT latencies implies that something is overriding the sleep system. The common finding in patients with primary insomnia is elevated arousal (Bonnet and Arand, 1997a). These patients have elevated heart rate (Bonnet and Arand, 1998a), elevated metabolic rate (Bonnet and Arand, 1995, 1997b), and elevated sympathetic nervous system activity (Bonnet and Arand, 1998a).

Evidence suggests that sleep latency varies as a function of level of arousal when amount of prior sleep is held constant. For example, several studies have shown that sleep latency is correlated with heart rate (Bonnet and Arand, 1999, 2000; Johns et al., 1976). However, heart rate and heart rate variability do not change after partial sleep deprivation (Muentener et al., 2000). This implies that these cardiac changes are both related to sleep latency and independent of sleep loss. Therefore, level of physiologic arousal is an independent determiner of sleep latency. As such, one can explain the consistent MSLT elevation in primary insomnia patients as due to their elevated level of physiologic arousal.

One study has specifically looked at individual differences in MSLT values in normal sleeping young adults to determine whether there are meaningful underlying differences in physiologic arousal even in relatively homogeneous normals (Bonnet and Arand, 2005b). In this study, young adults were chosen as normal sleepers and were split into several groups based on their MSLT and trait subjective sleepiness ratings. Two groups were chosen to report normal and equal trait subjective daytime alertness but to have either normal MSLT (> 10 minute) or short MSLT (< 7 minute) latencies. The subjects with the short MSLT were found to have a slower heart rate at rest and decreased sympathetic nervous system activity (based on spectral analysis of their heart period). These findings suggest that some of the variability in MSLT in normal young adults after normal nights of sleep in the laboratory can be related to their trait level of physiologic arousal as indexed by heart rate.

Level of physiologic arousal can also be an issue in patients with sleep apnea. Untreated OSA is associated with increased sympathetic activity (Cutter et al., 2004). When we try to predict sleep latency on the MSLT or MWT in patients with OSA, we must realize that these patients may have a number of physiologic problems independent of fragmented sleep that may influence their level of sleepiness. For example, it is known that simple aging is associated with progressive activation of the sympathetic nervous system (Seals and Bell, 2004), and this change may also be driven in part by obesity (Seals and Bell, 2004). In addition, sleep apnea patients may suffer from some degree of depression, anxiety, or insomnia (all of which have been associated with increased sympathetic activation). All of these physiologic factors probably combine with actual sleep disturbance in these patients to determine MSLT values, but have not been accounted for in the published research. When patients with sleep apnea are effectively treated, their sleep disturbance decreases and the sympathetic activation produced by the apnea decreases. This means that there is decreased pressure from the sleep system (which would increase the MSLT) and decreased pressure from the arousal system (which would decrease the MSLT). The net effect could be little change in sleep latency despite effective apnea treatment. Understanding how the many sources of arousal interact with the sleep system to determine sleep latency can improve our ability to predict how and when the MSLT should change.
STATE LEVEL OF CNS AROUSAL

Elevated arousal and long sleep latency is a trait in patients with primary insomnia. However, changes in physiologic level of arousal and related sleep latency can also be caused by a number of state influences. For example, studies have shown that physiologic activation, typically a 5-minute walk as compared with watching television, results in significantly longer sleep latencies in normal young adults (Bonnet and Arand, 1998b) and that the impact of one 5-minute walk on cardiac and sleep latency measures remains significant for at least 100 minutes under basal sleep conditions (Bonnet and Arand, 2005c). Similar effects of the 5-minute walk were also found in patients with primary insomnia (Bonnet and Arand, 2000), and this suggests that sleep latency, as measured by the MSLT is a function of both trait level of physiologic arousal and state level of physiologic arousal (Bonnet and Arand, 2000) in addition to length of prior wakefulness and circadian time. These physiologic activity parameters were chosen because they were easy to operationalize and measure. However, a broad range of external events such as bright light, noise, social interaction, and temperature in addition to a broad range of internal events such as motivation, boredom, test anxiety, general anxiety, pain, or competing demands can produce physiologic arousal and therefore have a direct impact the MSLT. For example, subjects who were motivated to "fake" sleepiness on the MWT were able to reduce their sleep latency by 15 minutes to a level indicative of severe sleepiness on that test (Bonnet and Arand, 2005a).

Another simple example of an event that can impact sleep latency (i.e., modify the MSLT) is competing demands. The sleep latency on the final nap of either a four- or five-nap MSLT protocol is typically the longest observed in the testing day. This result is often explained as a circadian rhythm effect, but it is equally probable that the "last nap effect" could be due to end-of-testing stress/expectancy (i.e., the competing demands of falling asleep versus being able to leave the hospital as soon as the test is done). This effect can be tested by comparing sleep latency on the last nap of four-nap MSLT protocols to nap 4 in five-nap MSLT protocols. These naps occur at the same circadian time, but in one case nap 4 is the final nap and in one case it is not. Twenty studies (Alloway et al., 1997; Bonnet and Arand, 1995, 1996, 2003; Broughton et al., 1988; Carskadon, 1998; Carskadon and Dement, 1979; Hartse et al., 1982; Manni et al., 1991; Martin et al., 1996; Mattmann et al., 1982; Milder et al., 1987; Newman and Broughton, 1991; Reynolds et al., 1991; Richardson et al., 1978; Roth et al., 1987; Seidel et al., 1984, 1987; Zwyghuizen-Doorenbos et al., 1988) were found that reported individual nap mean and standard deviation for four or five nap MSLT protocols. Weighted means and standard deviations were calculated, and the MSL on the fourth nap in a four-nap MSLT protocol was significantly longer (11.3 minutes ± 6.4; N = 189) compared with a MSL of 9.4 minutes ± 5.8; (N = 231) on the fourth nap of a five-nap MSLT (t141 = 3.162; P < 0.01). One can conclude from these data that the arousal associated with the end of study demand produced a 1.9-minute increase in observed latency when nap 4 was the last nap of the day.

At the most trivial level, it is common that many of these potential sources of arousal (i.e., light level, noise, activity, social interaction, motivation) are poorly controlled or not controlled between nap attempts in many clinical or research settings where the MSLT is performed. Better recognition and control of sources of arousal may help reduce variability in reported results. At another level, however, the interrelationship of both sleep and arousal components in the determination of sleep latency implies that the MSLT is more than a simple measure of the sleep system. In studies where repeated measurements are made from the same subjects during sleep deprivation in a constant lab environment, sources of arousal are well controlled. However, this is much less likely to be the case in a clinical environment where patients with various disorders and sources of external and internal arousal will be compared with an arbitrary cut point. In such a situation, a large number of additional sources of variance exists.

TRAIT SLEEP LENGTH

Most sleep professionals will agree that each individual has his own sleep need (i.e., how many hours of sleep an individual regularly needs to feel well rested, awake, and alert) and that sleep need may be different from the number of hours that an individual usually sleeps at night (as nutritional needs differ from what people actually eat). However, almost no research is done with individuals who have a sleep need different from 7 to 8 hours per night (Bonnet, 2005). It is almost impossible to report studies from individuals who characteristically sleep less than 7 hours per night because all results are immediately criticized as being secondary to chronic partial sleep deprivation. There is no agreed on method of determining sleep need, and, as such, it is really not possible to tell if an individual who reports generally spending 8 hours per night in bed is chronically partially sleep deprived because his real sleep need is 10 hours per night or chronically sleep satiated ("overslept") because his real sleep need is only 7 hours per night. Because we are commonly unsure about actual sleep need, it is difficult to provide a true baseline sleep length before the MSLT, and this certainly adds variability to test results.

Currently, there is almost a circular definition where any sleep latency shorter than 20 minutes on the MSLT has been used as a de facto definition of less than full alertness (Kamdar et al., 2004), even though the MSLT is a measure of ability to fall asleep rather than to remain awake (Bonnet and Arand, 2005a). It has been speculated that full alertness, defined as being unable to fall asleep during an MSLT, is a favorable goal that could be achieved by simply continuing to increase nightly sleep time (Kamdar et al., 2004). However, this definition implies that falling asleep is never an appropriate response. Why should we consider this to be the case? The ability to fall asleep during a period of decreased performance demand may be extremely adaptable to provide improved alertness and performance when needed at a later time. Harrison and Horne (1996) described a group of young adults with borderline MSL between 6 to 8 minutes even after sleeping 10 hours per night for 2 weeks. They thought these
individuals who can fall asleep whenever convenient but deny excessive daytime sleepiness had “high sleepability without sleepiness.” Their ability to easily take “power naps” could be viewed as an adaptive advantage.

One method for identifying how much sleep an individual really needs is to track their sleep and wake times over several weeks by placing them in a time free environment that minimizes external demands. In such a “free-running” environment, young adults typically develop a circadian day that is longer than 24 hours, often in the range of 25.3 hours (Webb and Agnew, 1974). Study subjects who reported normally sleeping 7.5 hours per day before the study averaged 8.8 hours of sleep per day during a 14-day isolation stay. This increase of sleep by 1.3 hours per day implies that underlying sleep need in these young adults might be closer to 8.5 rather than 7.5 hours. However, one could just as well argue that their increased sleep time was determined by the decrease in arousing stimuli associated with the isolation environment and therefore might also be related to the level of arousal. The implication is that habitual sleep length is determined by the balance of the sleep and arousal systems. Falling asleep at an inappropriate time may be as much a failure of the arousal system as an imposition of the sleep system.

**CLINICAL UTILITY OF THE MSLT AND MWT AND LIMITATIONS**

The recent AASM review of the MSLT and MWT acknowledges many of the limitations inherent in these tests (Arand et al., 2005). Particularly important comments were: 1) “MSL (mean sleep latency) on both the MSLT and MWT does not discriminate well between patients with sleep disorders and normal populations...” (139); and 2) “MSL should not be the sole criterion for determining sleepiness or for certifying a diagnosis or response to treatment” (140). As indicated in this review, differentiating narcoleptics from normal young adults who may have a longer than 8-hour underlying sleep requirement or who may be partially sleep-deprived for other reasons is difficult. Some young adults may display rapid sleep onset secondary to parasympathetic dominance but may have normal alertness if evaluated in a different setting. Seventeen percent of normal young adults had two or more SOREMPs during a standard MSLT evaluation. Only further research will determine if this finding was based on subjects who were really sleep-deprived long sleepers showing REM rebound on naps, or whether a subpopulation of normal individuals without narcolepsy characteristically have short REM latencies. The data certainly imply that the incidence of SOREMPs in normal sleepers (17%) is much higher than the incidence of narcolepsy (0.05%) and leave substantial room for misdiagnosing normals as having narcolepsy. A confident diagnosis of narcolepsy in the absence of cataplexy is greatly lessened because of concern that the patient may only be an individual with a long sleep requirement.

Based on recently published practice parameters from the Standards of Practice Committee of the American Academy of Sleep Medicine (Littner et al., 2005), the MSLT is not recommended for patients with OSA. This lack of recommendation was based on the poor relationship between apnea severity and MSLT results. Clearly, there are a broad range of MSLT values in these patients (which may or may not reflect a broad range of underlying apnea and sleep disturbance) and a relative lack of improvement in sleep latency in the placebo-controlled CPAP treatment trials. Why OSA patients exhibit such variability in their sleep latencies will become more apparent with: a) increased sensitivity to the many arousal variables that may also help determine sleep latency in these patients; b) better understanding of the impact of the variable sleep fragmentation patterns seen in sleep apnea on sleep restoration, and c) better control of the testing environment.

The AASM practice parameters (Littner et al., 2005) also do not recommend a MSLT for patients with primary insomnia, in part because these patients do not appear to be sleepy, based on MSLT criteria. However, the real message is that additional research must be done to address the core causes and pathology in insomnia and to place that understanding within the context of sleepiness and sleep tendency. If the MSLT is going to be a useful clinical tool, we need to understand why certain populations and patients have longer sleep latencies on the MSLT despite reduced sleep at night because such test results are completely opposite from those seen in the partial or total sleep deprivation studies that were first used to validate the MSLT.

The data suggest that the MSLT and MWT are extremely sensitive. A large number of clinical and research investigations have used them. However, the data also suggest that these tests are sensitive to more than the sleep system and that we have much to learn about the determinants of sleep tendency in both normals and in patient groups. Effective clinical use and interpretation of the MSLT and MWT require a better understanding of sleep need and arousal system effects in normal subjects. We also need to better understand how age, body weight, activity patterns, motivation, and numerous clinical pathologies impact on the sleep and arousal systems.

We frequently use these tests in the real world to confirm the ability of individuals to remain awake and alert in medical, transportation, military, and industrial settings where sleepiness could compromise public safety. A better understanding of the variables that determine sleep onset is critical to provide the tools necessary for individuals to remain functional in our society. In addition, better understanding of the determinants of sleep onset can provide the framework for intelligent design of the workplace in a manner that maximizes performance by manipulation of sources of arousal.

So far, the MSLT has proved marginally successful for identifying narcolepsy patients. Although narcolepsy patients are important and deserving, they make up only about 0.05% of the general population. We currently do not have accurate tests to measure sleepiness and alertness in patients with other sleep disorders and the population as a whole. We need to refine our tests to better differentiate sleepiness and arousal and to give functional answers in applied settings.
REFERENCES


Bonnet MH, Anard DL. Arousal components which differentiate the MWT from the MSLT. Sleep. 2001;24:441–450.


Mitter MM, Waldehn IA, Sargent RB, Hirschowitz M. Sleep latency on the maintenance of wakefulness test (MWT) for 530 patients with narco-


