56 Appendix I: The subjective measurement of excessive daytime sleepiness

Murray W Johns

INTRODUCTION

All patients who present with a sleep disorder for investigation and treatment require some assessment of their daytime sleepiness. Excessive daytime sleepiness (EDS) is one of their most common and important symptoms, i.e. 'sleepiness in a situation when an individual would be expected to be awake and alert'.[1] Indeed, it may not be asking too much to include an assessment of EDS in all comprehensive medical examinations. However, the best way to make that assessment is still a matter for debate.[2–4] Many different methods have been proposed, some based on objective measurements, others on subjective reports.[5] Of course, objective methods are to be preferred if they give the same information as subjective reports but with greater accuracy and reliably. Subjective reports of all kinds have potential problems in terms of reporter bias due to differences in the awareness of one's behavioral states, poor language and reading skills, lack of motivation to answer questionnaires, concerns over the potential consequences of doing so, and occasionally because of deliberate deception.

Why, then, do we need to use subjective reports of sleepiness at all? The answer is becoming clearer. The currently used objective methods, including the Multiple Sleep Latency Test (MSLT) and the Maintenance of Wakefulness Test (MWT), cannot be relied on, in isolation, to be accurate methods for distinguishing EDS.[1,6,7] However, there is now an acknowledged role for other relevant clinical information and for subjective methods to be included when making such an assessment. There is another reason for using subjective methods under some circumstances. They are far quicker, cheaper, and simpler to use than objective methods, and can be used repeatedly and with large numbers of subjects in epidemiological studies.

Some people believe that the measurement of EDS is the Holy Grail of sleep medicine.[3] That belief seems to have arisen for several reasons, not least of which is the reluctance of most sleep clinicians and researchers to define what they mean by terms such as sleepiness, drowsiness, and fatigue. Several such definitions appeared two or three decades ago, at the beginnings of sleep medicine, with little discussion or analysis of the concepts involved. They have seldom been questioned since then, despite the fact that knowledge of the physiology of sleep and wakefulness has dramatically increased. Johns[6] has argued that confusion about these concepts and the inadequacy of currently accepted models of sleep and wakefulness underlies much of the current difficulty with the measurement of sleepiness.

The aims of this chapter are, first, to describe the methods that are currently used for subjective assessments of daytime sleepiness in clinical practice and research; secondly, to describe how the results from those different methods are related to each other; and, thirdly, to review the current concepts and definitions of sleepiness and to suggest some revisions that may help us understand what the various tests are measuring, and hence elucidate the nature of sleepiness and EDS.

CURRENT METHODS FOR MEASURING SLEEPINESS SUBJECTIVELY

There are several different methods that are currently used for measuring sleepiness subjectively. They are described here in an order that does not reflect their current importance or frequency of use. Several methods rely on introspection and self-reports of feelings and symptoms that would indicate the presence of the drowsy state (the intermediate state between alert wakefulness and sleep) at the time. Others rely on retrospective reports of *dozing behavior* in particular situations. Sometimes these tests are said collectively to measure *subjective sleepiness* as opposed to *objective sleepiness* measured by the MSLT or MWT, etc. However, as we shall see below, this is not a particularly useful distinction.

Stanford Sleepiness Scale

The Stanford Sleepiness Scale (SSS) was introduced in 1973.[8] It comprises a series of statements, numbered

Stanford Sleepiness Scale
Circle the ONE number that best describes your level of alertness or sleepiness right now.
1. Feeling active, vital, alert, wide awake
2. Functioning at a high level but not at peak, able to concentrate
3. Relaxed, awake but not fully alert, responsive
4. A little foggy, let down
5. Foggy, beginning to lose track, difficulty staying awake
6. Sleepy, prefer to lie down, woozy
7. Almost in reverie, cannot stay awake, sleep onset appears imminent

Figure 56.1 Stanford Sleepiness Scale. (After Hoddes et al.[8])

Karolinska Sleepiness Scale
Here are some descriptors about how alert or sleepy you might be feeling right now. Please read them carefully and CIRCLE the number that best corresponds to the statement describing how you feel at the moment.
1. Extremely alert
2. Very alert
3. Alert
4. Rather alert
5. Neither alert nor sleepy
6. Some sighs of sleepiness
7. Sleepy, but no difficulty remaining awake
8. Sleepy, some effort to keep alert
9. Extremely sleepy, fighting sleep

Figure 56.2 A modified version of the KSS. (After Reyner and Horne.[15])

1–7, that range from 'feeling active, vital, alert, wide awake' to 'almost in reverie, cannot stay awake, sleep onset appears imminent' (Figure 56.1). The different statements are presumed to represent an ordinal Likert scale that reflects different positions along a continuum of states between alert wakefulness, through progressively deeper levels of drowsiness, to sleep. The respondent is asked to choose which statement most accurately describes his feelings at the time.

The SSS has been widely used, particularly for studying the effects of sleep deprivation,[9] sleep fragmentation,[10] and circadian rhythms.[11] However, scores on the SSS are not closely related to sleep latencies measured a few minutes later in MSLTs.[12] Another problem with the SSS is that factor analysis suggests that it is not a unitary scale.[13] It seems to measure sleepiness, in the sense of drowsiness, and fatigue. This may arise because so many different poorly defined words are used in the SSS statements, such as 'responsive', 'foggy', 'vital', and 'woozy'. The SSS is best used to measure changes within subjects over time, particularly over periods of hours and days. Scores on the SSS often require standardization (e.g. to *z*-scores) to remove differences between subjects. The SSS cannot provide an overall measure of EDS in daily life.

Karolinska Sleepiness Scale

The Karolinska Sleepiness Scale (KSS) is a 9-point Likert scale somewhat similar to the SSS.[14] In its original format it had word descriptors only for scores of 1, 3, 5, 7, and 9. Those descriptors varied from 1 = 'very alert' to 9 = 'very sleepy, fighting sleep, an effort to keep awake'. However, additional descriptors were later added for all scores, as shown in Figure 56.2.[15] It is not always clear which version has been used in a particular study.

The changes observed in the EEG/EOG (electroencephalogram/electro-oculogram) with drowsiness do not usually appear until KSS scores reach ≥7.[16] The KSS is assumed to be an ordinal scale with a unitary structure, although that has not been confirmed. The KSS has been used widely, particularly for describing changes over time within subjects.[16–18] KSS scores may require standardization to control for differences between subjects.[16,17]

Visual analogue scales

A Visual Analogue Scale (VAS) is typically a horizontal line 100 mm long across a page, with a word at each end that represents the extremes, for example, between 'very sleepy' and 'very alert'.[19] The subject is asked to place a mark at that point on the line that represents his current state along that continuum. The VAS score is the distance (measured in mm) between the subject's mark and one or other end of the line. Scores on such a VAS respond to time-of-day effects and to sleep deprivation at least as well as, and perhaps better, than the SSS.[19] Several different VAS scores can be used in parallel, if required, with different pairs of words at the extremes for each scale.[20–22] Clearly, the choice of words is critical, so that each VAS represents a single dimension of variation. VAS scores often have to be standardized (e.g. to *z*-scores) to allow for changes within subjects while removing the differences between subjects. VAS scores have not been used to diagnose or quantify EDS in daily life, but presumably could be used, for example, to quantify the degree of 'difficulty' that EDS was causing.

Responses to individual questions about sleepiness

When inquiring about EDS for some purposes, it may be advisable to measure the subject's sleepiness in one particular context that is of special interest, such as the propensity to doze off while driving. For example, one question may ask: 'How many times have you dozed off at the wheel while driving during the last year?'. The answer could be selected from a range of frequencies, such as 0, 1–2, 3–5, 6–10, >10. Additional questions may be required, such as: 'How many times in the last year has the vehicle you were driving been involved in a crash caused, at least in part, by your drowsiness at the time?'. Answers to such questions are probably the best indicator yet of a driver's average sleep propensity while driving, and this could form one basis for deciding whether that person was fit or unfit to have a driving license. However, responses to such questions about drowsy driving may well be inaccurate if the granting of the driver's license depends on the answers given.

The Rotterdam Sleepiness Scale (RSS)[23] uses a series of such questions to make a 'global evaluation' of the impact of daytime sleepiness ('Are you troubled by extreme daytime sleepiness?') as well as the 'behavioral impact' and 'affected life domains'. The latter asks whether the respondent sometimes falls asleep 'when sitting quietly', 'during meetings', 'when you are driving a car', etc. The possible answers are either yes or no. This part of the RSS scale is somewhat similar to the ESS (see below) but has not been widely used.

Sleep–Wake Activity Inventory

The Sleep–Wake Activity Inventory (SWAI) was developed in 1991.[24] It has 35 items with several subscales, one of which is called 'daytime sleepiness'. This is based on five questions about how frequently the subject dozes off in different situations and how frequently they have 'difficulty staying alert throughout the day'. Scores on the 'daytime sleepiness' subscale of the SWAI are related to reduced hours of sleep, ease of falling asleep at night, snoring, and major depression.[25]

The SWAI has an advantage in being able to assess several different aspects of a subject's sleep and wake behavior at the same time, but it has not been widely used as a measure of daytime sleepiness or of EDS.

Adjective check lists

The Profile of Mood States (POMS) is a list of 65 adjectives to which subjects are asked to respond by either selecting or not selecting each word as a descriptor of their mood state at the time. There are six dimensions involving different combinations of adjectives. One of those dimensions relates to 'vigor–activity' and another to 'fatigue–inertia'. Scores on these two POMS scales have sometimes been used to study sleepiness.[22] In general, the use of POMS scales for this purpose is to be discouraged because other scales, such as a VAS of *alertness–drowsiness*, appear to be more robust and are better validated.

Epworth Sleepiness Scale

The Epworth Sleepiness Scale (ESS) was introduced in 1991[26] and is now the most commonly used method worldwide for assessing a person's sleep propensity subjectively. It is based on retrospective reports of dozing behavior in a variety of situations that are commonly experienced in daily life. The term dozing behavior requires some explanation. Drowsiness is characterized (among other things) by intermittent lack of awareness of the here-and-now.[27,28] There is also inhibition of tonic and phasic muscle activity.[29] This often becomes apparent first in the muscles of the eyelids, levator palpebrae superioris and orbicularis oculi. The velocity of eyelid movements during blinks is reduced by drowsiness, so the eyelids take longer to close and reopen during blinks, they tend to remain closed for longer, and the upper eyelids droop.[30–32] If we doze off while sitting with our head unsupported, the muscles that hold our head erect when we are awake begin to have their tonic activity inhibited, and this eventually allows the head to drop forward. That movement often rouses us briefly and makes us aware of having just dozed off, without us having been aware of the beginning of the nodding movement or the level of drowsiness at that time.

The ESS is a simple self-administered questionnaire that asks subjects to rate on a scale of 0–3 their usual chances of having dozed off when they have been in each of eight different situations *in recent times*

Epworth Sleepiness Scale

Name: ______________________________ Today's date: ______________________________

Your age (years): ______________ Your sex (Male = M, Female = F): ______________________

How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired?
This refers to your usual way of life in recent times.
Even if you haven't done some of these things recently try to work out how they would have affected you.
Use the following scale to choose the **most appropriate number** for each situation:

0 = would **never** doze
1 = **slight chance** of dozing
2 = **moderate chance** of dozing
3 = **high chance** of dozing

It is important that you answer each question as best you can.

Situation	Chance of dozing (0–3)
Sitting and reading	—
Watching TV	—
Sitting, inactive in a public place (e.g. a theater or a meeting)	—
As a passenger in a car for an hour without a break	—
Lying down to rest in the afternoon when circumstances permit	—
Sitting and talking to someone	—
Sitting quietly after a lunch without alcohol	—
In a car, while stopped for a few minutes in the traffic	—

THANK YOU FOR YOUR COOPERATION

Figure 56.3 Epworth Sleepiness Scale. (After Johns.[26])

(0 = 'would never doze' and 3 = 'a high chance of dozing') (Figure 56.3). The situations were chosen on a priori grounds to vary in what Johns initially called their soporific nature.[26] He later called this their somnificity[33,34] (see below).

The ESS score, which can range from 0 to 24, is the sum of 8 item-scores, each between 0 and 3. The ESS does not measure *subjective sleepiness* as some people have thought,[35] because it does not assess subjective feelings. It asks only for the subject's retrospective recall of dozing behavior in the specified situations. That the ESS refers to observable behavior rather than subjective feelings is supported by the close relationship ($r = 0.74$, $n = 50$, $p < 0.001$) between each subject's ESS score and that reported independently about the subject by his partner.[36] That is comparable to the test–retest reliability of ESS scores when repeated by the same subject after a few months ($r = 0.81$, $n = 87$, $p < 0.001$).[37]

The ESS score gives an estimate of what Johns calls the subject's *average sleep propensity* in daily life.[33] The higher the ESS score, the more likely the subject is to doze off in situations of low somnificity: i.e. the person has a higher sleep propensity in those situations than other people have. A large epidemiological study of identical and non-identical twins who answered the ESS independently found that about 40% of the variance in scores between subjects represents a genetically determined trait, reflected in long-term differences in *average sleep propensity* that are partially inherited.[38]

When choosing the items of the ESS, it was necessary to include only those situations that most people would encounter in their daily lives, not necessarily very often, but often enough for them to form an estimate of their dozing behavior in each. This precluded asking about dozing behavior in some situations that might otherwise be of special interest to investigators, such as dozing while driving a car. That is why the descriptor for Question 8 is 'In a car, while stopped for a few minutes in the traffic'. It does not specify as a driver or a passenger, because it must allow for those who do not drive. If one such question remained unanswered, the total ESS score could not be calculated. Interpolation of scores is not possible. The ESS has been translated into many languages, some published but many more unpublished. The original version of the ESS was modified in 1997 by the addition of an extra sentence of instructions ('It is important that you answer each question as best you can'). This greatly decreased the proportion of missing ESS scores, to the order of 1%.

Psychometric properties of the ESS

There is good evidence for the internal consistency, unitary structure, and consistent hierarchical item structure of the ESS as a sum-scale. That evidence has come from several quarters, including factor analysis, which has usually, but not always, found only one main factor that includes all ESS items, but with somewhat variable weightings in different groups of subjects.[36,39,40] Other evidence has come from item analysis and Cronbach's α statistic, α varying between about 0.73 and 0.88 in different groups of subjects.[37,40] More recently, there has been the application of Item Response Theory, e.g. with Rasch analysis, that might be considered to provide the most detailed evidence about the internal structure of the ESS. So far, the latter has been published only for a group of patients with Parkinson's disease.[40] However, those results are supported by other results of Rasch analysis from 990 subjects in Australia, including patients with a variety of sleep disorders and normal subjects (J Pallant, pers comm, 2007).

The external validity of the ESS has been demonstrated in several ways, e.g. by its high sensitivity (93.5–97%) and specificity (100%) for distinguishing narcoleptics (who have EDS) from normal subjects;[41,42] by the change in ESS scores after continuous positive airway pressure (CPAP) treatment for obstructive sleep apnea (OSA) syndrome, typically reducing scores by ≥5 units;[37,43] and by the change in ESS scores after treatment of narcoleptics with modafinil as compared with placebo.[44] ESS scores have a test–retest reliability over a period of 5 months of $r = 0.81$ ($n = 87$, $p < 0.001$).[26]

DIFFICULTIES ARISING FROM DIFFERENT DEFINITIONS OF SLEEPINESS

One important problem that has caused much confusion about daytime sleepiness and its measurement has been that of deciding what is being measured.[6] It is relevant to consider the common English dictionary definitions of the terms sleepiness, drowsiness, fatigue, and tiredness that have probably not changed for a century or more. This should provide a good basis from which to make any changes because of greater understanding in recent times. The *Shorter Oxford English Dictionary* defines the adjective *sleepy* as 'inclined to sleep, having difficulty in keeping awake, drowsy, somnolent'. The noun *sleepiness* therefore means the state of being *sleepy*. The definition of the adjective *drowsy* is virtually the same as for *sleepy* – 'inclined to sleep, heavy with sleepiness, half asleep, dozing'. By contrast, *fatigue* is defined as 'weariness resulting from bodily or mental exertion', and *tiredness* is defined as being 'fatigued, weary'. That is, *sleepiness* is synonymous with *drowsiness*, and *fatigue* is synonymous with *tiredness*.

There is no ambiguity from these dictionary definitions about the distinction between the state of sleepiness/drowsiness on the one hand and that of fatigue/tiredness on the other. That confusion evidently arose during the latter part of the 20th century when applied psychologists were trying to measure fatigue, particularly in people at work. For example, ID Brown,[45] who was very influential in the study of fatigue, defined it as 'the subjective experience of tiredness and a disinclination to continue performing the current task'. That was compatible with the common English dictionary definition of fatigue. However, Brown went on to say that 'the main effect of fatigue [in a driver] is a progressive withdrawal of attention from road and traffic demands' and 'probably the most frequent cause of general attentional impairment is the eye closure that accompanies sleepiness'. Without any explanation, and perhaps with little detailed knowledge of the drowsy state, Brown associated the state of fatigue/tiredness with the state of sleepiness/drowsiness. That confusion has persisted in many quarters, not only in applied psychology but also in sleep medicine,[46] and particularly in discussions about road safety.[47]

In the 1980s, a new meaning of the word sleepiness was widely accepted by sleep researchers as a 'physiological need state that leads to an increased tendency to sleep'.[48] Another was as 'a physiological drive usually resulting from sleep deprivation'.[49] These definitions loosely equated sleepiness with what might now be called the *sleep drive*, a measure of the coordinated activity of neuronal systems in the central nervous system which together promote the state of sleep as opposed to wakefulness. But this approach made it necessary for another concept to be introduced, that of *masking*. It was self-evident that we can usually avoid falling asleep by remaining active and, in particular, by not lying down. This led to the idea that 'physiological sleepiness may not necessarily be manifest'.[48] As a corollary, it was stated that 'heavy meals, warm rooms, boring lectures, and the monotony of long-distance driving unmask the presence of physiological sleepiness but do not cause it'.[48] That is, *masking* would not affect *physiological sleepiness* but it would prevent sleep onset as an expression of *manifest sleepiness*. This concept of *masking* was never developed and was not quantified. It is the present author's contention that those definitions, and the conceptual basis from which they arose, were never appropriate and have provided the basis for much of the confusion about sleepiness that is still evident today.[1] Sleepiness should not be equated with a *physiological sleep drive*.[6,42]

It has also become common in recent years for sleep researchers to use the word sleepiness to mean *sleep propensity*, or sometimes *physiological sleep tendency*.[7] By a circular argument, it is sometimes implied that sleepiness is that which is measured by sleepiness scales. Johns[33,50] has distinguished different categories of sleep propensity to help understand the nature of sleepiness and of EDS. These categories are based on different time-frames and on what the subjects are doing at the time.

Instantaneous sleep propensity

Instantaneous sleep propensity, a subject's sleep propensity at any particular time, whatever the circumstances, can vary rapidly with a change of posture and activity. However, it also reflects the subject's posture and activity during the preceding few minutes. Mild exertion, such as walking for 5 minutes instead of sitting and watching TV, reduces the *instantaneous sleep propensity* for the following few minutes.[51] Each nap in the MSLT and MWT provides a measure of the subject's *instantaneous sleep propensity* at the time, reflected in the sleep latency (SL). However, the *instantaneous sleep propensity* measured with each nap in the MSLT is very different from that measured in the MWT (approx. 11 and 18 minutes, respectively) and they are not highly correlated ($r = 0.41$).[52] The *instantaneous sleep propensity* is not only subject-specific but also posture and activity-specific and situation-specific.

The *instantaneous sleep propensity* increases with the level of drowsiness at the time – the drowsier you are, the closer you are to falling asleep. Neither the total ESS score nor the ESS item-scores measure the *instantaneous sleep propensity* directly. Measures of the *instantaneous sleep propensity*, *situational sleep propensity*, and *average sleep propensity* are only moderately intercorrelated. Similarly, the SSS, KSS, or VAS, which measure drowsiness rather than sleep propensity at a particular time, cannot measure the *instantaneous sleep propensity* directly. However, the presence of drowsiness at a particular time may provide an indirect measure of *instantaneous sleep propensity*, if that is the preferred parameter.

A subject's state of drowsiness can be measured objectively and continuously by monitoring the EEG and EOG, as for example while driving a truck.[53] However, this requires the attachment of multiple electrodes and, while this can be justified for research purposes, it is not a suitable method for routine use in daily life. Video camera images of a driver's face have been used for this purpose,[54] but there are technical problems maintaining the quality of data in real-life situations.[54,55] A new system of infrared (IR) reflectance oculography, using IR transducers embedded in a pair of glasses frames, has recently been introduced that can measure a person's state of drowsiness continuously and unobtrusively while they are driving or engaged in other activities.[56] How this can be used in clinical practice to help diagnose sleep disorders is yet to be established.

In summary, the *instantaneous sleep propensity* is a measure at some particular time of the subject's likelihood of making the transition from the waking state to the state of drowsiness and perhaps to sleep no matter what they happen to be doing at the time. In the MSLT and the MWT, that likelihood is assumed to be directly related to the SL measured for each nap. By contrast, scores on the SSS, KSS, and VAS of *alertness–drowsiness* (and the vigor and fatigue scales of the POMS) are all very similar in providing a measure of drowsiness/fatigue rather than *instantaneous sleep propensity* at the time, although they are moderately correlated.

Situational sleep propensity

There are as many different *situational sleep propensities*, a subject's usual sleep propensity when measured in the same situation repeatedly, as there are different situations in which to measure sleep propensity. Each ESS item-score gives a subjective measure of one particular *situational sleep propensity* (e.g. while 'sitting and reading'). Mean SLs in the MSLT and the MWT provide objective measures of two different *situational sleep propensities*. Different *situational sleep propensities* within the same subject are only moderately correlated (r = approx. 0.4) whether measured objectively or subjectively (see below).

Average sleep propensity

This is a hypothetical construct based on a subject's *average sleep propensity* when engaged in all the different activities of daily life.[36] The *average sleep propensity* usually remains fairly constant ($r = 0.81$), at least over periods of several months,[36] and is partly inherited. The *average sleep propensity* increases with the onset of a sleep disorder such as OSA or narcolepsy, and decreases again after successful treatment of such a disorder. The ESS score gives an estimate of a subject's *average sleep propensity*, but only in relation to the limited range of specified situations. Neither the MSLT nor the MWT gives an accurate measure of *average sleep propensity*, and their mean SLs are only moderately correlated with ESS scores, as expected (see below). Currently, there is no objective equivalent of the ESS for the measurement of *average sleep propensity*, which is what many clinicians want to measure in their patients.

RELATIONSHIPS BETWEEN DIFFERENT MEASURES OF SLEEPINESS

Much can be learned about the nature of sleepiness by comparing the results from the many different tests that are thought to measure it, whether by objective or subjective methods. Those comparisons can be extended to include the different categories of sleep propensity as defined above.

Comparisons between different subjective measures of sleepiness

When comparing the results of different subjective tests of sleepiness, we shall make two main types of comparison: first, comparing the results of different tests on the same subjects, at the same time, and in reference to the same set of circumstances (i.e. measuring one particular *situational sleep propensity* by different methods) and, secondly, on the same subjects, at the same time, but in different sets of circumstances (i.e. measuring several different *situational sleep propensities* by the same methods).

Comparing different subjective tests of sleepiness in the same circumstances

Several comparisons of this type have been made. For example, scores every 2 hours between 0700 and 2300 h reported by 40 healthy airmen were compared for the SSS, KSS, and two kinds of VAS, first as a single scale of *sleepiness* and secondly as the mean of 10 VAS scores, asking about 'tired eyes', 'heavy eyelids', etc.[17] Scores on each scale were converted to z-scores. All four scales then gave similar results and were sensitive to the time-of-day effect, but the effect was larger for the VAS scores than for either the SSS or KSS.

Pilcher and her colleagues[22] reported on the relationships between seven different subjective scales of sleepiness, all completed at essentially the same time by psychology students. The tests comprised the SSS, three different VAS scales (*sleepy–awake*, *active–lethargic*, and *alert–drowsy*), two subscales from the POMS adjective check list (*fatigue* and *vigor*), and the ESS. The ESS was used in two modes: first, in a modified mode in which the respondents were asked to rate their chances of dozing off in each situation 'now'; secondly, in the original mode of dozing off 'in recent times'. Raw scores from each pair of scales were intercorrelated significantly, but the highest correlations were between those scales other than the ESS (mean $r = 0.59$; range of $r = 0.32$–0.79). By contrast, the mean of correlations between the ESS and all other scales was $r = 0.26$ with the range of $r = 0.17$–0.32. Factor analysis clearly showed that ESS scores, whether reported in the modified or the original mode, formed quite a different factor from all the other scales, which together formed a single factor.

This suggests several conclusions. Likert scales, such as the SSS and KSS, as well as VAS scores and adjective check lists asking about feelings and symptoms related to the presence of drowsiness and/or fatigue at a particular time, all provide similar measures that are highly intercorrelated ($r = 0.59$). For many people, the state of drowsiness may be difficult to distinguish subjectively from the state of fatigue, and these scales may draw on a combination of feelings and symptoms of both drowsiness and fatigue. By contrast, ESS scores measure the subject's *average sleep propensity*, which is an average of several different *situational sleep propensities* reported subjectively about different circumstances and other times. Measurements of *average sleep propensity* (ESS scores) are not highly correlated with the subjective measures of drowsiness/fatigue at one particular time derived from the SSS, KSS, or VAS scores ($r = 0.26$). Nor are ESS scores highly correlated with scores on other scales measuring longer-term fatigue, such as the Fatigue Severity Scale (FSS) ($r = 0.33$, $n = 489$, $p < 0.001$).[57]

Comparing different tests of subjective sleepiness in different circumstances

Kim and Young[58] described a study in which they performed factor analysis of subjective measures of sleepiness using responses to individual questions which, in the light of the categories described here, included measures of different *situational sleep propensities*. They used the eight individual ESS item-scores as well as responses to six other questions asking about the frequency of occurrence of particular feelings and problems, such as 'not feeling rested during the day', the 'need for coffee or other stimulants to stay awake during the day', and finding it 'very difficult to get up in the morning'. The latter responses were given as estimates of the frequency of occurrence, with a range of times per month on a 5-point Likert scale. Factor analysis was performed on raw scores for 13 variables, followed by oblique rotation so that highly correlated factors could be retained.

They reported three factors. The first factor they called *perceived daytime sleepiness*, which was based on reports of frequently having a 'feeling of excessive daytime sleepiness', and a 'need for coffee, etc.'. This factor could be interpreted as the frequency of occurrence of feelings of drowsiness during the day. The authors called their second factor *subjective sleep propensity in active situations* and their third factor *subjective sleep propensity in passive situations*. Unfortunately, their distinction between *active* and *passive* situations was poorly based. They described 'sitting quietly after lunch without alcohol' and 'sitting inactive in a public place (e.g. a theater or a meeting)' as active situations without an explanation to justify that distinction (see below about the somnificity of ESS activities). Their second and third factors were highly correlated oblique factors ($r = 0.59$, $p < 0.05$), so a clinically significant distinction between them may be difficult to establish.

In the light of our earlier discussions about the hierarchical structure of ESS items, it seems reasonable to conclude that Kim and Young's analysis confirmed that the frequency of feelings of drowsiness during the day (which they called *perceived daytime sleepiness*) represents a different aspect of sleepiness from that of a *situational sleep propensity*, which is what each ESS item measures for a particular situation. Drowsiness and sleep propensity, in its various categories, could be considered to be two different dimensions of sleepiness, although their measures are usually weakly correlated (r = approx. 0.2–0.3).

Analysis of ESS item-scores has enabled two different kinds of comparison to be made:

- by intercorrelating the 8 ESS item-scores reported by each subject, thereby demonstrating the relationships between different *situational sleep propensities* within subjects
- by ranking the item-scores within each subject, to investigate the effects that different situations/activities have on sleep propensity, regardless of differences in overall sleepiness between subjects.

Johns[33] introduced the term *somnificity* to describe the latter. Somnificity is not a characteristic of people or of their EDS. It is the characteristic of a set of circumstances, including the subject's posture and level of activity, both physical and mental, and the level of environmental stimulation at the time, that facilitates or inhibits sleep onset in the majority of subjects. The brief descriptor for each ESS item gives no more than an outline of such circumstances.

A set of ESS item-scores has been analyzed for 990 adult subjects, aged between 17 and 78 years old, that included groups of patients from the Epworth Sleep Centre (Melbourne, Australia) with a variety of sleep disorders such as OSA and narcolepsy, as well as a group of male and female industrial workers, and groups of ostensibly healthy university students.[34] Their total ESS scores varied between 0 and 24. Because the item-scores were not always normally distributed, Spearman's non-parametric correlations were calculated for each pair of item-scores from the 8 items for each subject. All 28 correlation coefficients were highly significant ($p < 0.0001$). Their mean was 0.45 and their range was 0.31–0.57. Principal components analysis revealed a single factor that included all items, which is consistent with previous results.[36] This is good evidence to suggest that the ESS item-scores are each tapping a common source of variance that represents the subject's overall level of sleepiness, or *average sleep propensity*, in different situations. However, each of the 28 correlation coefficients accounted for only about 20% of the variance between different *situational sleep propensities* within the same subject (range 10–33% for different items). Much of that variance (80%) was unaccounted for.

The same data were also transformed into ranks of item-scores, from 8 = highest item-score to 1 = lowest item-score within each subject. Tied ranks were each assigned to their mean.[34] This procedure enabled differences between the overall levels of sleepiness (*average sleep propensity*) between subjects to be eliminated. Similar analysis was also performed on data from another 614 subjects, aged 36–48 years old, who had taken part in the CARDIA study in the USA, answering the ESS on two separate occasions 1 year apart.[59] The mean rank of ESS item-scores within subjects was calculated separately for the Australian and US subjects (Figure 56.4). Differences were tested by Wilcoxon's matched pairs *t*-tests.

The item-ranks were very similar in these groups of subjects from different countries, and were also very similar a year later when repeated by the US subjects. Their item-ranks were taken to represent different somnificities of the respective sets of circumstances described in the ESS items. Combining the results from Australia and the USA led to an ordinal scale of somnificities, with six significantly different levels (p <0.001). The order of items, from highest to lowest somnificity, was 5, 2, 1, 4, 7 and 3, 8 and 6. Items 7 and 3 did not differ significantly, nor did items 8 and 6. Some of the differences in somnificity can be ascribed to differences of posture (lying down in item 5 vs sitting in all other items). However, more subtle influences must be involved in the differences in somnificity between items 1 and 2, on the one hand, and items 6 and 8, on the other hand, such as the interaction and emotional engagement with other people. We need far more experimental evidence about such influences.

These results provide evidence for the role of sensory inputs in the control of *situational sleep propensities*. Whether or not someone dozes off or not during the day will depend not only on their overall level of daytime sleepiness (their *average sleep propensity*) but also on what they are doing at the time. We cannot measure a subject's sleep propensity accurately without reference to the set of circumstances within which the measurements are made. By measuring one particular *situational sleep propensity* (e.g. when sitting and

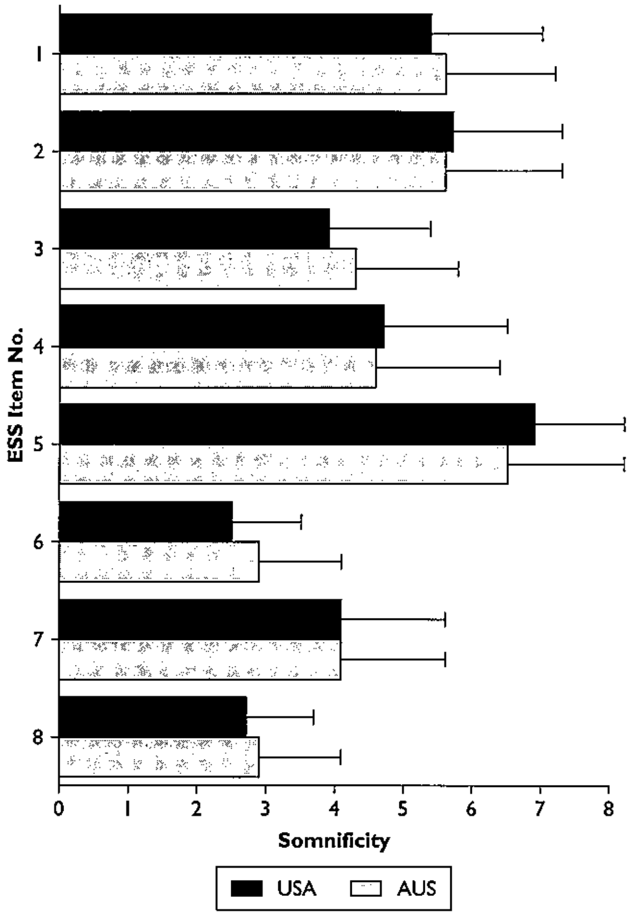

Figure 56.4 The somnificity of different items described in the Epworth Sleepiness Scale (mean + SD) for 614 US subjects and 990 Australian subjects.

reading), we are not measuring the subject's *average sleep propensity* directly, nor are we making a direct measurement of their *physiological need state* or *sleep drive* (see below). It appears that these limitations also apply to objective measures of different *situational sleep propensities*, including those from the MSLT and MWT.

Comparisons between subjective and objective measures of sleepiness

Objective and subjective measurements of drowsiness and sleep propensity cannot usually be made at precisely the same time. The fact that a person's *instantaneous sleep propensity* can change in a matter of seconds, and that all subjective reports of drowsiness are relatively inaccurate, can help explain why SSS scores and SLs in each MSLT nap, when measured within the same subject and within a few minutes of each other, are not highly correlated.[12,60]

Many researchers have measured the correlation between ESS scores and mean SLs in the MSLT.[26,61] Most, but not all such correlations have been statistically significant, but none has indicated a very close relationship. The mean of Pearson's correlation coefficients from nine separate studies, some involving as many as 522 subjects, was –0.30.[42] The reasons for this are a matter of contention.[35] However, those results certainly indicate that the ESS and the MSLT are measuring different aspects of sleep propensity. Similar findings have been reported in relation to the ESS and MWT.

Comparisons between different objective measures of sleepiness

To put the above comparisons involving different subjective measures of sleepiness into a broader perspective, it is helpful to consider comparisons between different objective measures also. Different *situational sleep propensities* in the same subject can be measured objectively from the mean SLs in MSLTs and MWTs performed on the same day. The mean SLs in the two tests are quite different (approx. 11 and 18 minutes, respectively) and they are moderately correlated ($r = 0.41$, $n = 258$, $p < 0.001$;[53] $r = 0.52$, $n = 522$, $p < 0.001$[52]). However, some subjects fall asleep more quickly in the MWT, when instructed to stay awake, than in the MSLT, when instructed to fall asleep.

These results have led to the suggestion that the two tests were measuring two different *abilities*: the MSLT, the *ability to fall asleep*, and the MWT, the *ability to stay awake*.[52] Harrison and Horne[62] introduced another term, *sleepability*, for the former. The implication was that these *abilities* were general, rather than situation-specific characteristics of each subject: i.e. they were thought to represent general *dimensions of sleepiness*. This is an example of generalizations about sleepiness, in the many different senses of the word, that have not been adequately discussed or questioned and which have created confusion as a result. Within the context of the definitions used here, the MSLT and the MWT each measure a different *situational sleep propensity*, and we should not be surprised that their results are only moderately correlated, as appears to be true for all *situational sleep propensities*.

What the term *sleepiness* currently means in the practice of sleep medicine is evidently quite confused. It may be appropriate to restrict the meaning of the word *sleepiness* to the state of *drowsiness*, as defined in common English usage, whether it is measured subjectively or objectively. The term *sleep propensity*, with its separate categories as defined here, would then be different from drowsiness and would have both subjective and objective methods for its measurement.

NORMAL DAYTIME SLEEPINESS

It is not that easy to decide what normal daytime sleepiness is, whether measured by the MSLT, the ESS, or any other method. It took more than 20 years for the reference range of normal mean SLs in the MSLT to be defined adequately.[1,42] That normal range is not 5–20 minutes, as was stated for many years on the basis of a 'rule of thumb'.[63] By that criterion, anyone with a mean SL <5 minutes was said to have *pathological sleepiness*, which was quite misleading. Many normal subjects, without evidence or complaints of EDS in other situations, fall asleep that quickly in the MSLT.[12,60] Johns[42] calculated the reference range for the MSLT to be 3–20 minutes, and others, using a broader database, reported it to be 2–19 minutes.[1] A mean SL <3 minutes would be required in the MSLT before making a diagnosis of EDS on that basis alone. Residual confusion about such matters is not helping the progress of sleep medicine. The reference range of normal values for the MWT has been defined, but there are different numbers for different test modes, depending on whether the test lasts for 20 or 40 minutes.[1]

Reference range of normal values for the ESS

When attempting to define normal values for the ESS, fairly strict criteria were used.[64] Among the 331 ostensibly healthy industrial workers that were studied, only those subjects were included who did not have a sleep disorder that could influence their daytime sleepiness, whether or not they complained about it. Those who snored frequently, whether or not they had significant OSAs, and others who had restless legs syndrome or insufficient sleep syndrome, etc., according to their subjective reports, were excluded. Only 22% ($n = 72$) of the whole group were included as 'normal' sleepers. Their mean ESS score was 4.6 ± 2.8 (SD). The reference range of normal ESS scores was therefore defined as 0–10, which is the mean ± 2 SD, and which also coincides with the 2.5th and 97.5th percentiles. By this definition, ESS scores >10 (i.e. between 11 and

24) would represent some degree of EDS, particularly for ESS scores >14. Similar results have been reported for healthy control subjects without a sleep disorder from Italy (4.4 ± 2.8)[65] and from England (4.5 ± 3.3).[41] However, more data of this type are required from other populations to confirm this. It is very unlikely that a patient with narcolepsy would have an ESS <11.[42] Unfortunately, some websites that refer to the ESS have suggested that the normal range is 0–9, without any evidence to support that.

WHAT IS EXCESSIVE DAYTIME SLEEPINESS?

One definition of EDS is 'sleepiness in a situation when an individual would be expected to be awake and alert'.[1] Other terms such as *pathological sleepiness* are sometimes used in relation to sleep disorders such as narcolepsy.[7] Problems with the definitions of drowsiness and sleep propensity that we have addressed here become very relevant again when attempting to define and measure EDS. The term *daytime* in EDS may no longer be appropriate because difficulties can arise because of being too drowsy or having too high a sleep propensity at any time of day or night under different circumstances. However, we probably need different definitions of what we might call *excessive sleep propensity* in terms of different time frames and circumstances:

1. **Excessive instantaneous sleep propensity**: when individuals' sleep propensity at any particular time is so high that it is likely to affect adversely the safety and efficiency of their performance of the intended task at hand.
2. **Excessive situational sleep propensity in particular situations**: when individuals have a higher-than-normal *situational sleep propensity*, sufficient to impair their performance significantly whenever they are in the particular circumstances of concern, e.g. whenever they are driving at night after being awake for 18+ hours.
3. **Excessive average sleep propensity**: when individuals' sleep propensity in most situations is so high that they are likely to doze off inadvertently in many different circumstances in daily life.

Scores on the SSS, KSS, and VAS can give subjective measures of drowsiness (and indirectly of *instantaneous sleep propensity*) at a particular time, but the significance of particular scores has not been validated as a measure of *excessive sleep propensity*. When drowsy subjects are trying to remain awake, their drowsiness usually fluctuates between different levels, as described by Ian Oswald[66] many years ago. At deeper levels of drowsiness, there is a lack of awareness of the here-and-now[27,28] and errors of omission in the performance of a task are likely to occur, whether or not the eyelids are open at the time.[67] In between, there will be periods of lesser drowsiness, when fewer errors of omission are made and there is greater awareness of the here-and-now, including some awareness of having just been more drowsy. However, when a drowsy driver makes an error of omission (e.g. fails to see or respond to something of particular importance), it is unlikely that he would be aware of it at the time. This may explain the apparent inconsistency between a general awareness of some level of drowsiness, e.g. reported subjectively in the KSS, and the lack of awareness of impairment in the performance of a task as a result of that drowsiness.[68,69] In a more practical context, a driver may be generally aware of becoming drowsy, on the basis of self-awareness during less drowsy periods, but not be aware of driving off the road a few seconds later during a more drowsy period. The awareness of having been drowsy may return when he rouses after the event. His level of drowsiness may have increased to a dangerous level and then decreased again to a less dangerous level within a matter of seconds.

Currently, the best method for monitoring *excessive instantaneous sleep propensity* is a combination of EEG/EOG recordings with video camera images of the person's face and eyes, detecting microsleeps objectively by the presence of theta waves in the EEG, and any associated changes in blinks and eyelid closures from the EOG and video.[70] A new method for deciding who is too drowsy to drive at a particular time has recently been proposed, based on continuous monitoring of drivers' drowsiness by IR reflectance oculography. Different levels of drowsiness, on a scale from 0 to 10, were calibrated against the likelihood of making errors of omission each minute in a visual psychomotor vigilance test and also during simulated driving in a car.[31,71]

For many years the MSLT was widely regarded as the gold standard for measuring EDS in all situations, and in 1992 it received the official blessing of the American Academy of Sleep Medicine (AASM).[72] That position has now changed officially. The MSLT is now officially regarded as no more than a 'de facto standard',[7] with the recommendation that the mean sleep latency in the MSLT and MWT 'should not be the sole criterion for determining sleepiness or for certifying a diagnosis or response to treatment'.[7] This is in accord with the conclusion, reached independently by Johns,[6,33] that we do not have a single gold standard method for measuring excessive sleep propensity. A person's *situational sleep propensity* measured in one particular situation, such as in the MSLT or MWT, cannot be relied on to give an accurate measure of their *situational sleep propensity* in other situations, or of their *average sleep propensity* in daily life. Nevertheless, measurements made by the MSLT and MWT are moderately correlated with *average sleep propensity*, as measured by the ESS, and may contribute objective evidence to confirm or deny the presence of excessive sleep propensity in a particular subject.

The most commonly used method for measuring excessive sleep propensity is the ESS, using a score of >10 as the indicator of excessive sleep propensity. The ESS has been used in many surveys, often in languages other than English, to establish the proportions of various groups that have *excessive sleep propensity*. For example, 17.7% of a sample of 2301 adults selected from the general adult population in Norway had ESS scores >10.[73] So did 18% of 526 consecutive drivers aged between 16 and 86 years old who presented to the Department of Motor Vehicles in Madison, WI.[74] Among 3871 school students in Seoul, Korea, 15.9% had ESS scores >10, and their school performance declined as their ESS scores increased.[75] In the group of 331 male and female industrial workers in Australia, 10.9% had an excessive sleep propensity, and in a Japanese population of 3909 industrial workers who did not do shift work the proportion was 7.2%.[76] A random, population-based survey among car drivers in New Zealand revealed 9.2% with ESS scores >10.[77] It seems reasonable to conclude that, in different countries, between about 7 and 18% of the general community, selected without reference to their sleep habits, have *excessive sleep propensity*.

Excessive sleep propensity and obstructive sleep apnea

In a recent survey in the USA, 46% of patients with moderate to severe OSA had *excessive sleep propensity*.[78] However, the relationship between *excessive sleep propensity* and the severity of OSA has been a matter of contention. There have been many reports of statistically significant but relatively weak correlations between different measures of daytime sleepiness and the frequency of apneas and hypopneas,[79] and sometimes also the levels of oxygen desaturation during sleep apneas.[80] Some studies have failed to find a significant relationship at all.[81] That has been equally true for all measures of *excessive sleep propensity*, whether objective or subjective. What is now clear is that there are some patients with OSA, occasionally with 'severe' OSA, who do not have *excessive sleep propensity*.[78] The relationship between the severity of OSA and sleep propensity (however they are measured) can no longer be used as the major evidence for or against the validity of any measure of *excessive sleep propensity*, whether objective or subjective in nature.

THE NEED FOR A NEW CONCEPTUAL MODEL OF SLEEP AND WAKEFULNESS

Apart from difficulties with the measurement of sleepiness because of differences in definitions and in peoples' understanding of what the various tests are measuring, there is another major problem that has impeded progress in this field. That is the lack of an adequate model to describe the processes that control sleep and wakefulness. Borbély[82] and his colleagues have clearly established that sleep and wakefulness are heavily influenced by two processes, Process C and Process S: the former provides a circadian influence over sleep propensity, and the latter provides a homeostatic influence that increases progressively during wakefulness and is discharged during sleep. Variations of this basic model have included the addition of a third process (Process W) that influences sleep inertia after waking up.[83] There is much evidence to suggest that these processes should be part of any such model, but Johns has argued that they are not sufficient.[6,80] Knowledge about the neurophysiological processes involved in sleep and wakefulness has increased dramatically in recent years, particularly in relation to opponent processes that together form a sleep–wake switch in the hypothalamus.[84] Another whole new system has been discovered[84] that promotes wakefulness, the orexin/hypocretin system. Yet this and other information has not been integrated into our understanding of the measurement of sleepiness, particularly in terms of the major influences that need to be explained.

Johns[6] has proposed a new conceptual model of sleep and wakefulness that specifically addresses the role of inputs to the central nervous system (CNS) from both the exteroceptive and enteroceptive sensory systems as major determinants of sleep and wakefulness, in addition to Processes C and S. Whether we are asleep or awake at any particular time depends on the relative strengths (not the absolute strengths) of both the *sleep drive* and the *wake drive* that are to some extent independent of each other, and which are mutually inhibitory to each other.[85] Johns[6] goes further in suggesting that the overall *wake drive* has more than one component. The *primary wake drive* would result mainly from the spontaneous activity of the suprachiasmatic nucleus (SCN), with a circadian rhythm reaching its peak in the evening (about 7 p.m.) and its nadir in the early hours of the morning (at about 4–5 a.m.). The phase of this circadian rhythm would be set mainly by the timing of periods of daylight and dark. This would be the equivalent of Process C. However, a *secondary wake drive* would provide additional wake-promoting activity derived from inputs from all sensory systems, filtered, integrated, and relayed via the thalamus. It is assumed that this *secondary wake drive* involves the orexin/hypocretin system in the hypothalamus which would be separate from the *primary wake drive* with input from the SCN.[84] There is experimental evidence from cats that orexin/hypocretin neurons are activated during wakefulness, but mainly when the animals are moving, when enteroceptive sensory inputs would be high, rather than when the animals are at rest.[86]

Other authors have pointed out the role that environmental stimulation plays in maintaining alertness at any particular time.[87] Having the bedroom comfortably warm, dark, and quiet would reduce the inputs to the CNS from the subject's exteroceptive sensory systems, including the visual, auditory, hot and cold temperature sensing systems, etc. That is what usually happens when we lie down in bed at night to sleep, and this reduced environmental stimulation facilitates sleep onset. However, it is self-evident that we would be very unlikely to fall asleep, regardless of how quiet and dark the room was, if we remained standing up rather than lying down. This emphasizes the role f posture in the control of sleep propensity, a role that has been almost completely ignored by sleep researchers.

Johns[6] has postulated that inputs to the CNS from all the enteroceptive sensory systems would also contribute to the *secondary wake drive*, perhaps even more than inputs from the exteroceptive sensory systems as a result of environmental stimulation. The enteroceptive inputs to the CNS would be derived from many sources, including the stretch receptors in all skeletal muscles that are active tonically when we maintain a particular posture (e.g. standing up) and active phasically when we move. Posture also influences input to the CNS from the carotid baroreceptors,[88] and from the labyrinthine system in the ear that detects the position and motion of the head. Ongoing cognitive processes and the associated emotional activity would also contribute to the 'secondary wake drive'.[89] The fact that it is very difficult to measure these enteroceptive sensory inputs may be one reason why they have been neglected by all but a few researchers.[51]

The existence of a *secondary wake drive* helps to explain the very important role of body posture and movement in the control of sleep and wakefulness. It would also explain the influence of environmental stimulation on sleep propensity. Most people can fall asleep within 20 minutes at any time of the day simply by lying down, ceasing voluntary movements, and relaxing most skeletal muscles apart from orbicularis oculi, which they usually contract voluntarily to close the eyelids actively and reduce visual input to the retina. This gives significant control over the *secondary wake drive* and hence voluntary control over sleep onset under most circumstances. This helps to explain why most people do not usually fall asleep at inappropriate times in their daily lives, and how they can extend periods of wakefulness overnight without much difficulty.

CONCLUSIONS

In the past it has been assumed that there must be some measure of sleepiness, preferably an objective measure, which would be a general characteristic of each subject and which would enable that person's tendency to become drowsy at some future time and place to be predicted accurately. Consequently, the problem was seen by some as a search for an elusive new test. Unfortunately, things are much more complicated than that. Direct measures of sleep propensity, objectively or subjectively, can be made only with reference to the subject's posture, activity, and environmental situation at the time. Sleep propensity appears to be influenced by the *wake drive*, particularly the *secondary wake drive* as described by Johns, as much or even more than by the *sleep drive* under normal circumstances. Extrapolation of a subject's sleep propensity from one set of circumstances to another has caused much confusion in the past. We need far more research in the future about drowsiness and the multiple influences over it. Some new definitions of different categories of sleep propensity may help in this regard. Our currently accepted models of sleep and wakefulness are inadequate and need revision.

ACKNOWLEDGMENT

Dr Kate Crowley helped with the preparation of this chapter.

REFERENCES

1. Arand D, Bonnet M, Hurwitz T et al. The clinical use of the MSLT and MWT. Sleep 2005; 28(1): 123–44.
2. Krieger J. Clinical approaches to excessive daytime sleepiness. Sleep 2000; 23(Suppl 4): S95–8.
3. Bliwise DL. Is the measurement of sleepiness the Holy Grail of sleep medicine? Am J Respir Crit Care Med 2001; 163(7): 1517–19.
4. George CF. Staying awake for safety sake: a dream not yet realized. Sleep 2005; 28(11): 1360–1.
5. Weaver TE. Outcome measurement in sleep medicine practice and research. Part 1: assessment of symptoms, subjective and objective daytime sleepiness, health-related quality of life and functional status. Sleep Med Rev 2001; 5(2): 103–28.
6. Johns M. Rethinking the assessment of sleepiness. Sleep Med Rev 1998; 2(1): 3–15.
7. Littner MR, Kushida C, Wise M et al. Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. Sleep 2005; 28(1): 113–21.
8. Hoddes E, Zarcone V, Smythe H et al. Quantification of sleepiness: a new approach. Psychophysiology 1973; 10(4): 431–6.
9. Dinges DF, Orne MT, Whitehouse WG et al. Temporal placement of a nap for alertness: contributions of circadian phase and prior wakefulness. Sleep 1987; 10(4): 313–29.
10. Stepanski EJ. The effect of sleep fragmentation on daytime function. Sleep 2002; 25(3): 268–76.
11. Monk TH, Weitzman ED, Fookson JE et al. Task variables determine which biological clock controls circadian rhythms in human performance. Nature 1983; 304(5926): 543–5.
12. Pressman MR, Fry JM. Relationship of autonomic nervous system activity to daytime sleepiness and prior sleep. Sleep 1989; 12(3): 239–45.
13. MacLean AW, Fekken GC, Saskin P et al. Psychometric evaluation of the Stanford Sleepiness Scale. J Sleep Res 1992; 1(1): 35–9.
14. Åkerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci 1990; 52(1–2): 29–37.
15. Reyner LA, Horne JA. Falling asleep whilst driving: are drivers aware of prior sleepiness? Int J Legal Med 1998; 111(3): 120–3.
16. Gillberg M, Kecklund G, Akerstedt T. Relations between performance and subjective ratings of sleepiness during a night awake. Sleep 1994; 17(3): 236–41.
17. Casagrande M, Curio G, Ferrara M et al. Subjective ratings of sleepiness: sensitivity of different measures in revealing time of day effects. Sleep Res Online 1999; 2(Suppl 1): 226.
18. Ingre M, Akerstedt T, Peters B et al. Subjective sleepiness and accident risk avoiding the ecological fallacy. J Sleep Res 2006; 15(2): 142–8.
19. Babkoff H, Caspy T, Mikulincer M. Subjective sleepiness ratings: the effects of sleep deprivation, circadian rhythmicity and cognitive performance. Sleep 1991; 14(6): 534–9.
20. Herbert M, Johns MW, Doré C. Factor analysis of analogue scales measuring subjective feelings before and after sleep. Br J Med Psychol 1976; 49(4): 373–9.
21. Sauter C, Asenbaum S, Popovic R et al. Excessive daytime sleepiness in patients suffering from different levels of obstructive sleep apnoea syndrome. J Sleep Res 2000; 9(3): 293–301.
22. Pilcher JJ, Pury CL, Muth ER. Assessing subjective daytime sleepiness: an internal state versus behavior approach. Behav Med 2003; 29(2): 60–7.
23. van Knippenberg FC, Passchier J, Heysteck D et al. The Rotterdam Daytime Sleepiness Scale: a new daytime sleepiness scale. Psychol Rep 1995; 76(1): 83–7.
24. Rosenthal L, Roehrs TA, Roth T. The Sleep-Wake Activity Inventory: a self-report measure of daytime sleepiness. Biol Psychiatry 1993; 34(11): 810–20.
25. Breslau N, Roth T, Rosenthal L et al. Daytime sleepiness: an epidemiological study of young adults. Am J Public Health 1997; 87(10): 1649–53.
26. Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 1991; 14(6): 540–5.
27. Foulkes D, Vogel G. Mental activity at sleep onset. J Abnorm Psychol 1965; 70: 231–43.
28. Ogilvie RD. The process of falling asleep. Sleep Med Rev 2001; 5(3): 247–70.
29. Martin CM, Ogilivie RD. A microanalysis of EMG and EEG activity during the sleep onset period (SOP). Sleep 2002; 25: A31.
30. Van Orden KF, Jung TP, Makeig S. Combined eye activity measures accurately estimate changes in sustained visual task performance. Biol Psychol 2000; 52(3): 221–40.
31. Johns MW, Chapman R, Crowley K et al. A new method for determining when a driver is too drowsy to drive. In: SENSATION 2nd International Conference: monitoring sleep and sleepiness with new sensors within medical and industrial applications, Chania, Greece, 2007.
32. Caffier PP, Erdmann U, Ullsperger P. Experimental evaluation of eye-blink parameters as a drowsiness measure. Eur J Appl Physiol 2003; 89(3–4): 319–25.
33. Johns MW. Sleep propensity varies with behaviour and the situation in which it is measured: the concept of somnificity. J Sleep Res 2002; 11(1): 61–7.
34. Johns MW, Knutson KL. The somnificity of different activities described in the Epworth Sleepiness Scale. Sleep 2006; 29: A30.
35. Chervin RD, Aldrich MS. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. Neurology 1999; 52(1): 125–31.
36. Johns MW. Sleepiness in different situations measured by the Epworth Sleepiness Scale. Sleep 1994; 17(8): 703–10.
37. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 1992; 15(4): 376–81.
38. Carmelli D, Bliwise DL, Swan GE et al. Genetic factors in self-reported snoring and excessive daytime sleepiness: a twin study. Am J Respir Crit Care Med 2001; 164(6): 949–52.
39. Broman JE, Bengtson H, Hetta J. Psychometric properties of a Swedish version of the Epworth Sleepiness Scale. J Sleep Res 2000; 9(Supp 1): 27.
40. Hagell P, Broman JE. Measurement properties and hierarchical item structure of the Epworth Sleepiness Scale in Parkinson's disease. J Sleep Res 2007; 16(1): 102–9.
41. Parkes JD, Chen SY, Clift SJ et al. The clinical diagnosis of the narcoleptic syndrome. J Sleep Res 1998; 7(1): 41–52.
42. Johns MW. Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the Epworth sleepiness scale: failure of the MSLT as a gold standard. J Sleep Res 2000; 9(1): 5–11.
43. Engleman HM, Kingshott RN, Wraith PK et al. Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med 1999; 159(2): 461–7.
44. Mitler MM, Harsh J, Hirshkowitz M et al. Long-term efficacy and safety of modafinil (PROVIGIL((R))) for the treatment of excessive daytime sleepiness associated with narcolepsy. Sleep Med 2000; 1(3): 231–43.
45. Brown ID. Driver fatigue and road safety. Alcohol Drugs Driving 1993; 9(3–4): 239–52.
46. Martikainen K, Hasan J, Urponen H et al. Daytime sleepiness: a risk factor in community life. Acta Neurol Scand 1992; 86(4): 337–41.

47. Dobbie K. Fatigue-related crashes: an analysis of fatigue-related crashes on Australian roads using an operational definition of fatigue. Commonwealth Department of Transport and Regional Services Road Safety Research Report; 2002.
48. Roth T, Roehrs T, Carskadon MA et al. Daytime sleepiness and alertness. In: Kryger MH, Roth T, Dement WC (eds). Principles and Practice of Sleep Medicine. Philadelphia, PA: WB Saunders, 1989: 14–23.
49. Aldrich MS. Automobile accidents in patients with sleep disorders. Sleep 1989; 12(6): 487–94.
50. Johns MW. A sleep physiologist's view of the drowsy driver. Transportation Res 2000; Part F 3: 241–9.
51. Bonnet MH, Arand DL. Level of arousal and the ability to maintain wakefulness. J Sleep Res 1999; 8(4): 247–54.
52. Sangal RB, Thomas L, Mitler MM. Maintenance of wakefulness test and multiple sleep latency test. Measurement of different abilities in patients with sleep disorders. Chest 1992; 101(4): 898–902.
53. Mitler MM, Gujavarty KS, Browman CP. Maintenance of wakefulness test: a polysomnographic technique for evaluation treatment efficacy in patients with excessive somnolence. Electroencephalogr Clin Neurophysiol 1982; 53(6): 658–61.
54. Dinges DF, Mallis MM, Maslin G et al. Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management. U.S. Department of Transportation National Highway Traffic Safety Administration DOT HS 808 762, 1998.
55. Hartley LR, Horberry T, Moabbott NA et al. Review of fatigue detection and prediction technologies. National Road Transport Commission Victoria; 2000.
56. Johns MW, Tucker AJ, Chapman R et al. A new scale of drowsiness based on multiple characteristics of blinks: The Johns Drowsiness Scale. Sleep 2006; 29: A365.
57. Shen J, Botly LC, Chung SA et al. Fatigue and shift work. J Sleep Res 2006; 15(1): 1–5.
58. Kim H, Young T. Subjective daytime sleepiness: dimensions and correlates in the general population. Sleep 2005; 28(5): 625–34.
59. Knutson KL, Rathouz PJ, Yan LL et al. Stability of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Questionnaires over 1 year in early middle-aged adults: the CARDIA study. Sleep 2006; 29(11): 1503–6.
60. Johnson LC, Spinweber CL, Gomez SA et al. Daytime sleepiness, performance, mood, nocturnal sleep: the effect of benzodiazepine and caffeine on their relationship. Sleep 1990; 13(2): 121–35.
61. Mitler MM, Walsleben J, Sangal RB et al. Sleep latency on the maintenance of wakefulness test (MWT) for 530 patients with narcolepsy while free of psychoactive drugs. Electroencephalogr Clin Neurophysiol 1998; 107(1): 33–8.
62. Harrison Y, Horne JA. "High sleepability without sleepiness". The ability to fall asleep rapidly without other signs of sleepiness. Neurophysiol Clin 1996; 26(1): 15–20.
63. Richardson GS, Carskadon MA, Flagg W et al. Excessive daytime sleepiness in man: multiple sleep latency measurement in narcoleptic and control subjects. Electroencephalogr Clin Neurophysiol 1978; 45(5): 621–7.
64. Johns M, Hocking B. Daytime sleepiness and sleep habits of Australian workers. Sleep 1997; 20(10): 844–9.
65. Manni R, Politini L, Ratti MT et al. Sleepiness in obstructive sleep apnea syndrome and simple snoring evaluated by the Epworth Sleepiness Scale. J Sleep Res 1999; 8(4): 319–20.
66. Oswald I. Sleep and Waking. Amsterdam: Elsevier, 1962.
67. Chapman R, Johns MW, Crowley K. In the drowsy state, errors of omission in a visual reaction-time test occur with eyes open or closed. Sleep 2006; 29: A364.
68. Dorrian J, Lamond N, Dawson D. The ability to self-monitor performance when fatigued. J Sleep Res 2000; 9(2): 137–44.
69. Atzram M, Chow C, Price NJ et al. Can sleep attacks occur without feeling sleepy? Sleep 2001; 24: A428.
70. Mitler MM, Miller JC, Lipsitz JJ et al. The sleep of long-haul truck drivers. N Engl J Med 1997; 337(11): 755–61.
71. Johns MW, Tucker A, Chapman R et al. Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. Somnologie 2007; 11: 234–42.
72. Thorpy MJ. The clinical use of the Multiple Sleep Latency Test. The Standards of Practice Committee of the American Sleep Disorders Association. Sleep 1992; 15(3): 268–76.
73. Pallesen S, Nordhus IH, Omvik S et al. Prevalence and risk factors of subjective sleepiness in the general adult population. Sleep 2007; 30(5): 619–24.
74. Benbadis SR, Perry MC, Sundstad LS et al. Prevalence of daytime sleepiness in a population of drivers. Neurology 1999; 52(1): 209–10.
75. Shin C, Kim J, Lee S et al. Sleep habits, excessive daytime sleepiness and school performance in high school students. Psychiatry Clin Neurosci 2003; 57(4): 451–3.
76. Doi Y, Minowa M, Fujita T. Excessive daytime sleepiness and its associated factors among male non-shift white-collar workers. J Occup Health 2002; 44: 145–50.
77. Connor J, Norton R, Ameratunga S et al. Prevalence of driver sleepiness in a random population-based sample of car driving. Sleep 2001; 24(6): 688–94.
78. Kapur VK, Baldwin CM, Resnick HE et al. Sleepiness in patients with moderate to severe sleep-disordered breathing. Sleep 2005; 28(4): 472–7.
79. Poceta JS, Ho S-L, Jeong D-U. The maintenance of wakefulness test in obstructive sleep apnea. Sleep Res 1990; 19: 268.
80. Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea. The Epworth Sleepiness Scale. Chest 1993; 103(1): 30–6.
81. Chervin RD, Aldrich MS, Pickett R et al. Comparison of the results of the Epworth Sleepiness Scale and the Multiple Sleep Latency Test. J Psychosom Res 1997; 42(2): 145–55.
82. Borbély AA. A two process model of sleep regulation. Hum Neurobiol 1982; 1(3): 195–204.
83. Åkerstedt T, Folkard S. Prediction of intentional and unintentional sleep onset. In: Ogilivie RD, Harsh JR (eds). Sleep Onset: Normal and Abnormal Processes. Washington, DC: American Psychological Association, 1994: 73–87.
84. Saper CB, Chou TC, Scammell TE. The sleep switch: hypothalamic control of sleep and wakefulness. Trends Neurosci 2001; 24(12): 726–31.
85. Edgar DM, Dement WC, Fuller CA. Effect of SCN lesions on sleep in squirrel monkeys: evidence for opponent processes in sleep-wake regulation. J Neurosci 1993; 13(3): 1065–79.
86. Torterolo P, Yamuy J, Sampogna S et al. Hypocretinergic neurons are primarily involved in activation of the somatomotor system. Sleep 2003; 26(1): 25–8.
87. Dinger DF. The nature of sleepiness: causes, context and consequences. In: Stunkard AJ, Baum A (eds). Perspectives in Behavioural Medicine: Eating, Sleeping and Sex. Hillsdale, NJ: Lawrence Erlbaum, 1989: 147–79.
88. Cole RJ. Postural baroreflex stimuli may affect EEG arousal and sleep in humans. J Appl Physiol 1989; 67(6): 2369–75.
89. De Valck E, Cluydts R, Pirrera S. Effect of cognitive arousal on sleep latency, somatic and cortical arousal following partial sleep deprivation. J Sleep Res 2004; 13(4): 295–304.