The Scoring of Arousal in Sleep: Reliability, Validity, and Alternatives

Michael H. Bonnet, Ph.D.[1]; Karl Doghramji, M.D.[2]; Timothy Roehrs, Ph.D.[3]; Edward J. Stepanski, Ph.D.[4]; Stephen H. Sheldon, D.O.[5]; Arthur S. Walters, M.D.[6]; Merrill Wise, M.D.[7]; Andrew L. Chesson, Jr., M.D.[8]

[1]Dayton Department of Veterans Affairs Medical Center, Wright State University, and Kettering Medical Center, Dayton, OH; [2]Sleep Disorders Center, Department of Psychiatry and Human Behavior, Jefferson Medical College of Thomas Jefferson University, Philadelphia, PA; [3]Sleep Disorders & Research Center, Henry Ford Health System, Department of Psychiatry & Behavioral Neurosciences, School of Medicine, Wayne State University, Detroit, MI; [4]Rush University Medical Center, Chicago, IL; [5]Northwestern University, Feinberg School of Medicine, Chicago, IL; [6]Seton Hall University School of Graduate Medical Education, New Jersey Neuroscience Institute at JFK Medical Center, Edison, NJ; [7]Methodist Healthcare Sleep Disorders Center, Memphis, TN; [8]LSU Health Science Center – Shreveport, Department of Neurology and Office of the Dean, Shreveport, LA

Abstract: The reliability and validity of EEG arousals and other types of arousal are reviewed. Brief arousals during sleep had been observed for many years, but the evolution of sleep medicine in the 1980s directed new attention to these events. Early studies at that time in animals and humans linked brief EEG arousals and associated fragmentation of sleep to daytime sleepiness and degraded performance. Increasing interest in scoring of EEG arousals led the ASDA to publish a scoring manual in 1992. The current review summarizes numerous studies that have examined scoring reliability for these EEG arousals. Validity of EEG arousals was explored by review of studies that empirically varied arousals and found deficits similar to those found after total sleep deprivation depending upon the rate and extent of sleep fragmentation. Additional data from patients with clinical sleep disorders prior to and after effective treatment has also shown a continuing relationship between reduction in pathology-related arousals and improved sleep and daytime function.

Finally, many suggestions have been made to refine arousal scoring to include additional elements (e.g., CAP), change the time frame, or focus on other physiological responses such as heart rate or blood pressure changes. Evidence to support the reliability and validity of these measures is presented. It was concluded that the scoring of EEG arousals has added much to our understanding of the sleep process but that significant work on the neurophysiology of arousal needs to be done. Additional refinement of arousal scoring will provide improved insight into sleep pathology and recovery.

Keywords: EEG arousal, sleep, sleep fragmentation, reliability, sleep deprivation, CAP.

Citation: Bonnet MH; Doghramji K; Roehrs T et al. The scoring of arousal in sleep: reliability, validity, and alternatives. *J Clin Sleep Med* 2007:3(2);133-145

1.0 HISTORICAL PERSPECTIVE

The earliest systematic characterization of the electroencephalogram (EEG) of sleep, published by Loomis and colleagues in 1938, defined 5 distinct states of sleep, labeled A through E.[1] This system recognized that sleep was not a steady state but consisted of cyclic patterns that varied in depth. In addition, Loomis et al[1] described the presence of brief alpha arousals that were associated with increased efforts to breathe and body movements, but these events were not included with the state scoring. The discovery of REM sleep and the development of the Dement-Kleitman criteria, which incorporated REM sleep scoring rules, led to the formalized Rechtschaffen and Kales (R&K) manual for scoring human sleep.[2] The R&K manual made reference to movement arousals, which were intended to aid in scoring stages but were not an epoch score. However, movement arousal in R&K was defined as an increase in the electromyogram (EMG) with no reference to changes in the EEG, and no additional scoring criteria were provided. No other mention of brief EEG frequency changes was made in the R&K manual.

The emergence of sleep disorders medicine brought renewed attention to the significance of the brief arousal as a consequence of primary sleep disorders and a determinant of the associated daytime sleepiness. Brief, 3-10 sec increases in EEG frequency

were categorized by sleep stage and time of night by Halasz et al in 1979,[3] and empirical work linking brief arousals, sleepiness and sleep apnea in dogs was published in 1980.[4,5] In the next 5 years, research showed significant correlations between brief arousals during sleep and daytime sleepiness,[6] began to examine different arousal definitions and their relationship to level of daytime sleepiness in different sleep disorder groups,[7] and empirically confirmed that the production of brief arousals during sleep produced increased sleepiness in humans.[8]

2.0 METHODS

The activities of the Arousal Task Force (see page 144) included development of this evidence review paper and participation in a RAND/UCLA Appropriateness Method consensus process for the development of scoring rules as part of the AASM scoring manual development project.[9,10] The AASM scoring rules derived from this review paper and consensus will be published as a separate volume.

2.1 Timeline. An initial computer based PubMed literature search was performed on July 7, 2004, and included articles published since 1966. The Arousal Task Force met by conference call 7 times between September 2004 and September 2005 for evidence review and had one face-to-face meeting in June 2005 for consensus balloting.

2.2 Search terms and articles

The following string of search terms was used: Sleep and Arousals and EEG or "sleep fragmentation" or "sleep disruption." This generated 2415 references. The committee reviewed the abstracts of these papers for relevance to EEG arousal scoring, reliability, or validity. As a result, a total of 122 papers were selected, obtained, and made available to all committee members. Fourteen additional references were identified by pearling (identification of new references from the reference lists of already selected papers).

2.3 Evidence selection and grading

Task force members were assigned review topics (i.e., scoring reliability, arousal definition validity, clinical studies of arousal, and new measures of arousal), and committee members were free to review any of the papers included in their area of review. Papers reviewed for inclusion were required to present empirical data relevant to the section. Exclusion criteria included abstracts, reviews, theoretical papers, editorials, and case studies; these sources, however, were considered for the general introductory and discussion sections of the paper.

A data extraction sheet was developed prior to review of the articles. Each arousal paper was reviewed and summarized on the extraction sheet for the following information: study design, number and sex of subjects, types of arousal measures, study outcomes and outcome measures, significant results, evidence grade, and miscellaneous notes. Evidence grading was modified from Sackett[10] and is presented in Table 1. An evidence table which summarizes all the extraction sheets has been provided at the end of each section of the paper.

Table 1—Summary of Evidence Grading

Evidence Levels	Study Design
I	Randomized well-designed trials with low-alpha & low-beta errors
II	Randomized trials with high-beta errors
III	Nonrandomized controlled or concurrent cohort studies
IV	Nonrandomized historical cohort studies
V	Case series

3.0 BACKGROUND

As the role of arousals in sleep disorders and sleep restoration became more apparent, numerous studies were published. To encourage further research and allow consistency of definitions for arousals, in 1992 the American Sleep Disorders Association (ASDA, since named the American Academy of Sleep Medicine) published an arousal scoring manual for EEG arousals. [11] At that time, various methods of scoring arousals were discussed, and the 6 definitions of arousal that included both EEG and EMG changes and various duration criteria that had been published by Stepanski et al[7] were reviewed. It was concluded that a 3-second increase in EEG frequency, coupled with an amplitude increase in EMG, was the best predictor of level of daytime sleepiness. In the ASDA manual that followed, an EEG arousal was defined as "an abrupt shift in EEG frequency, which may include theta, alpha and /or frequencies greater than 16 Hz, but not spindles," which is at least 3 seconds in duration. The scoring rules required a concurrent increase in EMG to score arousals in REM sleep. These rules were consistent with scoring conventions in use by most labs prior to development of the manual, with the possible exception of the required EMG increase to score arousals in REM. As such, the use of the phrase "EEG arousal" in this paper should be considered similar to the ASDA definition unless otherwise specified.

Additional research has addressed the basic and clinical significance of EEG arousals. It is clear that EEG arousals occur as a typical feature of normal sleep. However, a good deal still needs to be known about arousals that may be pathological either by virtue of sheer number or by being markers for specific events. In addition, a number of definitional questions have arisen: It is known that EEG arousals are accompanied by a number of physiological changes. Some studies have found autonomic nervous system (ANS) signs of arousal (i.e., blood pressure changes, heart rate changes) in the absence of prominent EEG changes. [12] Do these ANS changes have functional significance by themselves? And if so, does that significance differ when an EEG arousal is present? Based upon scoring reliability issues, the ASDA arousal scoring criteria required that the EEG acceleration last for at least 3 seconds. It has been suggested that arousals of shorter duration may also be significant. In addition, computer scoring methods allow identification of arousals that are not visually scorable, and it is possible that these events may also have significance. [13] Finally, some have suggested that other events such as K-complexes or bursts of delta waves may also be forms of arousal that need to be considered, and there is continuing controversy concerning whether K-complexes and delta-bursts could reflect a "disruption" or a "protection" of sleep. [14, 15] If one considers that the R&K scoring system reflects the macrostructure of sleep, then the brief

EEG/EMG/ANS arousals, however defined and interpreted, reflect the microstructure of sleep. It is the microstructure of sleep that is the focus of this review. The review is divided into an initial section that considers reliability and validity data to support arousals as described by the ASDA in 1992.[11] The second section reviews other types of arousal scoring that have been suggested, along with reliability and validity data when available.

4.0 RELIABILITY

4.1 ASDA arousal reliability

Several studies have addressed the reliability of scoring arousals with fewer papers reporting arousal scoring reliability values. An important determinant of ASDA arousal scoring reliability was the event duration. The ASDA manual specified a 3-second duration, but this decision was made on the basis of personal experience of several of the task members and not on systematic data. Several more recent studies have verified that decision. As presented in Table 2 (which can be accessed on the web at www. aasmnet.org), six level III studies demonstrated reliability using the standardized ASDA definition.[16-21] and one study demonstrated reliability using the ASDA definition slightly modified for use with infants.[22] One additional level IV study performed prior to standardization of scoring rules but using similar rules was also included in the table.[7] Interscorer reliability for different definitions of arousals was assessed by Loredo et al in 20 subjects with and without obstructive sleep apnea.[17] Arousal scoring that used the ASDA definition had a 0.84 intraclass correlation coefficient (ICC) value, but the value dropped to 0.19 to 0.37 when shorter arousals were also scored. In another study of 20 children with obstructive sleep apnea syndrome and 16 healthy normals,[21] the ICC was 0.90 when arousals were scored based upon ASDA definitions and decreased to 0.42 when events of 2 seconds or longer were included. The ICC decreased further to 0.35 when one-second duration arousal events were also included.

Another factor evaluated was the addition of an EMG increase criterion during all sleep stages to the ASDA criteria. In the previously cited study by Loredo et al, the addition of the EMG criteria improved the ICC from 0.84 to 0.92.[17] Smurra et al assessed the addition of an EMG increase to ASDA scoring criteria in 20 patients with obstructive sleep apnea of varying severity.[18] They found the EMG criteria provided no further improvement. However, their reported ICC was already at 0.98 for the ASDA criteria. In a study that compared scoring between 14 different sleep centers on 90 events based upon sleep stage from which the arousal occurred, a kappa of 0.47 was reported.[16] On an event-by-event basis, this means agreement on 74 of 100 events. Comparing the REM kappa, which was higher, to the light sleep kappa led to the conclusion that EMG criteria could possibly provide better scoring reliability. Therefore, these studies generally support the use of an EMG criterion for a small increment in reliability.

The study by Drinnan et al[16] is the only study to have assessed the impact of sleep stage on arousal scoring; sleep stage might be expected to play a prominent role in reliability. Drinnan et al found that the highest kappa for arousal scoring was for arousals that occurred out of slow wave sleep.[16] In terms of signal-to-noise, the background EEG of slow wave sleep is much different from that of the low amplitude fast frequency of the arousal. In contrast, the background EEG in light sleep and REM sleep is of low amplitude with higher frequencies; these are more difficult to distinguish from lower amplitude and fast frequency shifts associated with arousals.

Another important factor, addressed by only one study, is the presence of additional cues on the polysomnogram (PSG) being scored. Many of the previously cited studies included respiratory tracings and, while not explicitly mentioned, probably also included EKGs. Thomas assessed arousal scoring reliability using ASDA criteria in 17 patients with obstructive sleep apnea syndrome.[19] The event-by-event scoring agreement between scorers was 91%. However, when the respiratory tracings were removed from the recording, the agreement dropped to 59%.

A final and obvious factor shown to affect scoring reliability is the experience of the scorers. A multicenter study assessed sleep stage (R&K) and arousal scoring (ASDA criteria) reliability of 30 unattended home polysomnograms.[20] An ICC of 0.54 was found for scoring arousals, and this figure improved to 0.76 when an inexperienced scorer was not included in the data set. In an infant study using slightly modified arousal scoring rules,[22] reliability, as calculated by kappa, was increased from 0.71 to 0.83 after additional training and from 0.83 to 0.86 after a month of additional experience.

4.2 Type of reliability

The majority of the studies discussed above assessed inter-rater reliability. One would expect intra-rater reliability to be higher than that of inter-rater reliability. The one study that compared intra- to inter-rater reliability found the event-by-event scoring agreement was 94% for the intra-rater scoring and 90% for the inter-rater scoring.[19] Finally, inter-rater reliability between centers is much lower than inter-rater reliability within centers. One study that compared arousal scoring from a set of 40-second samples sent to 14 laboratories found a kappa of 0.47 across laboratories, while another study across scorers in the same center found kappa in the range of 0.81 to 0.83.[16,20]

4.3 Summary

These studies indicate that EEG arousals can be scored with relatively high reliability. For ASDA arousals, a 3-second duration criterion is necessary because reliability drops greatly when trying to score shorter events. There is no clear evidence to suggest that adopting EMG criteria for NREM sleep would greatly improve scoring reliability. However, extensive training and experience is necessary to develop highly reliable arousal scoring. Further, evidence that scoring reliability was reduced when assessed between centers indicates that additional training experience (such as provision of a set of "gold standard" scored polysomnograms) might provide benefit. Finally, cues from available tracings, beyond those specified for arousal scoring, could be used to improve arousal scoring reliability.

5.0 VALIDITY

Although brief arousals had been noted as normal occurrences during sleep for some time, it is the potential relationship between increased number of nocturnal arousals and increased sleepiness that has justified the time-consuming scoring process. Empirical evidence that supports the relationship between sleep continuity and sleep restoration will be reviewed in this section as a means of addressing the validity of arousal scoring. The major means

of providing evidence for validity has been to vary arousals during sleep either empirically or through natural observation and to measure residual sleepiness. Change in function after sleep fragmentation can be compared with function after total sleep deprivation to determine the similarity between disturbance of sleep and complete loss of sleep. To the extent that these conditions are similar, i.e., to the extent that sleep fragmentation produces effects similar to that of sleep deprivation, measurement of sleep fragmentation parameters can provide an estimate of the means by which sleep restores normal alertness. The sections that follow will review evidence that arousals during sleep do reduce sleep restoration, the extent to which the reduction in function is similar to that seen after total sleep deprivation, and the extent to which similar arousal effects are seen in patients with sleep disorders such as sleep apnea syndrome. The extent to which clinical sleepiness resolves in patients when arousals are reduced will also be considered as evidence for validity.

The initial sections address validity as assessed from empirical animal and human studies. Twenty-one studies (8 level II [23-30] and 13 level III,[8,31,32,4,33-37,5,38-40]) provide evidence for the validity of arousals and are summarized in Table 3 (which can be accessed on the web at www.aasmnet.org).

5.1 Animal model

In dog studies published in 1980,[4,5] novel auditory stimulation produced for 30-sec periods every 2.5 minutes for 2-3 nights resulted in dogs that appeared behaviorally sleepy and fell asleep significantly faster after being aroused. The dogs also had decreased O_2 and CO_2 sensitivity during sleep, along with decreased arousal responses to laryngeal stimulation.

5.2 Human model

The animal observations corresponded with sleepiness found in patients with fragmented sleep associated with sleep apnea and led to a number of studies in humans with the specific hypothesis that fragmentation of sleep by the production of brief, periodic EEG arousals at a sufficient rate would impair the process of sleep restoration and produce increasing sleepiness. The empirical studies performed in this area will be reviewed in the context of validity; that is, to what extent do graded changes in brief arousals produce predictable changes in selected outcome variables? Selected outcome variables have typically been chosen based upon their previously demonstrated sensitivity to sleep deprivation. Therefore, it is possible to compare the effects of various types of sleep fragmentation to periods of total sleep deprivation using the same measures and to determine the extent to which outcomes are similar or different.

It is known that the effects of total sleep deprivation are cumulative – that deficits become more extreme as sleep loss continues. It could therefore be hypothesized that deficits should also become more extreme as sleep fragmentation continues. Sleep deprivation is typically an all or none event (one is being sleep deprived or one is not). Sleep fragmentation differs in that it is a graded event. Arousals occur as a normal part of the sleep process. One can imagine arousals increasing in frequency to the point that they occur immediately after each sleep onset, resulting in total sleep deprivation and producing the same effects as total sleep deprivation. Such reasoning would imply that arousals produced at frequencies between these extremes might produce a wide va-

riety of outcomes ranging from no effect to those consistent with no sleep. As such, it is possible to hypothesize that increases in the frequency of arousals should make effects more profound. The empirical literature was examined to determine the extent to which 1) deficits found after total sleep deprivation are also found in the same outcome measure or physiologic system after sleep fragmentation; 2) deficits found after sleep fragmentation accumulate with additional nights of fragmentation in the same way as they accumulate with additional nights of sleep deprivation; 3) deficits increase to approximate total sleep deprivation as the frequency of arousals within a night increase. To the extent that empirical studies support these contentions, arousals can be seen as meaningfully related to sleep restoration.

A number of outcome measures have been examined following both sleep fragmentation and sleep deprivation. Types of measures examined include psychomotor performance, sleepiness, hormones, respiratory function, metabolic rate, arousal threshold, EEG measures, and mood/subjective report. Sleep fragmentation studies examining these variables will be reviewed first, and 4 studies that directly compared sleep fragmentation schedules with sleep deprivation conditions within the same experiment will be reviewed second.

5.2.1 Psychomotor performance

Many empirical studies of sleep fragmentation have used psychomotor performance measures, often transferred directly from sleep deprivation studies, to document functional change after various schedules of fragmentation. Psychomotor performance changes following sleep fragmentation compared to baseline have consistently been in the same direction as seen after sleep deprivation, and significant impairment has been documented on a number of tasks including vigilance,[19,23-26,31,32,30] reaction time,[8,31,30] additions,[24,31] digit-symbol substitution,[8] divided attention,[27] and trail making.[35]

5.2.2 Sleepiness

Objective sleepiness following sleep fragmentation has been measured with various sleep latency tests including both the multiple sleep latency test (MSLT) and maintenance of wakefulness test (MWT). One sleep fragmentation review paper summarized changes in sleep latency as a function of different schedules of sleep fragmentation using data from 8 studies.[41,23,26,32,28,35,37,38,30] Four level II[23, 24, 28, 30] and 4 level III studies[35, 37, 38, 42] assessed the validity of EEG arousal in relation to objective sleepiness, though methods in all of these studies preceded the standardized ASDA arousal definition. Daytime sleepiness increased as the rate of sleep fragmentation increased, with a significant correlation of r = 0.775 across studies between the interval of undisturbed sleep at night and sleep latency the next day. In other words, as the period of sleep allowed between empirically-produced disturbances decreased, latencies on nap tests on the following day became shorter.

5.2.3 Hormones

Spath-Schwalbe[39] studied cortisol and ACTH throughout baseline and sleep fragmentation nights (fragmentation after each minute of sleep). Plasma cortisol increased significantly when the experimental sleep fragmentation began. However, this burst of secretion was inhibited after about 80 minutes, and cortisol secre-

tion dropped below baseline sleep levels before returning to normal. A similar but less pronounced effect was found for ACTH. Following total sleep deprivation, patterns of cortisol and ACTH resembled the pattern seen during sleep fragmentation.[39]

Other hormones have not been studied in the empirical sleep fragmentation design. However, one study has shown delay of the testosterone rhythm in a design that allowed 7-minute sleep periods followed by 13-minute wake periods around the clock.[43] Patients with sleep apnea show decreased growth hormone and prolactin secretion at night.[44,45] Although this decrease is probably secondary to sleep fragmentation and reverses after sleep is normalized, changes in these hormones have not been studied after simple sleep fragmentation.

5.2.4 Pulmonary Measures

Changes in pulmonary function related to sleep fragmentation were first directly studied in dogs.[4,5] The initial studies documented decreased O_2 and CO_2 sensitivity and decreased arousal responses to laryngeal stimulation. In a later study designed to study either sleep fragmentation or sleep apnea in dogs for months,[40] it was shown that both the sleep apnea condition and the simple sleep fragmentation condition produced similar increases in the length of the time to arousal in response to airway occlusion.[40,46] The study also reported greater tolerated oxygen desaturation, greater peak inspiratory pressure, and greater tolerated surges in blood pressure during airway occlusion following both sleep apnea and simple sleep fragmentation. Based upon these results, the authors concluded that "the changes in the acute responses to airway occlusion resulting from obstructive sleep apnea (OSA) are primarily the result of the associated sleep fragmentation."[40]

An increase in the number of apnea/hypopnea events has also been shown after sleep fragmentation[46] in humans. Another study reported an increase in upper airway collapsibility and apnea/hypopnea index in normal subjects after sleep fragmentation.[29] Studies have not found changes in hypercapnic ventilatory responsiveness or in arousal responses to external inspiratory resistive loading after sleep fragmentation.[46,47]

In comparison to sleep fragmentation, total sleep deprivation has also been shown to produce longer apnea events in both infants[48] and adults.[49] Series et al[29] showed that the increase in apnea/hypopnea index after sleep fragmentation was actually greater than the increase after a similar amount of total sleep loss.

5.2.5 Metabolic measures

One study recorded whole body metabolic rate (VO_2 and VCO_2) during baseline, sleep fragmentation, and recovery nights.[32] Significant elevation of VO_2 was found during the fragmentation night as compared to baseline. Significant decreases in VO_2 were found during the recovery night compared to baseline. Metabolic rate increases during wake compared to sleep and decreases as subjects move from stage 1 to stage 4,[50] but changes in metabolic rate have not not been reported during recovery from total sleep deprivation.

5.2.6 Arousal threshold

Many sleep fragmentation experiments have used auditory stimuli to produce fragmentation, and when graded stimuli have been used, auditory threshold to the production of an arousal response has been reported. Studies have shown significant increases in auditory arousal threshold during 2 nights with EEG arousals placed after each 1 minute or 2 minutes, but not after 60 minutes of sleep.[8,25,26] One study[8] used occasional novel tones to measure the extent to which the increase in auditory threshold was secondary to sleep fragmentation as compared to habituation and concluded that about 67% of the increase in arousal threshold was related to the fragmentation procedure. Increased arousal threshold of a similar magnitude was found during recovery sleep following total sleep deprivation[51] compared with baseline sleep.

5.2.7 EEG measures

Two studies reported that visual evoked responses had decreased amplitude at several sites (frontal, central, and temporal) after sleep fragmentation, but significant increases in latency were not found[34,33] In comparison, significantly decreased amplitude and increased latencies for visual evoked responses have been reported from a central site after one night of total sleep loss.[52]

Spectral analysis of EEG showed a decrease in alpha: theta ratio at central and frontal sites after sleep fragmentation.[33] This finding is an agreement with theta findings during total sleep deprivation.[53]

5.2.8 Mood/subjective measures

A number of measures, including visual activation scales, the Stanford Sleepiness Scale, and several mood scales such as the Clyde Mood Scale and the Profile of Mood States (POMS), have been used to assess sleepiness and mood in both sleep fragmentation and sleep deprivation studies. Significant increases in subjective sleepiness have been reported in several studies after sleep fragmentation.[8,24,25,31,26,34] Following sleep fragmentation, mood measures become more negative,[33,35] including increased irritability,[33] increased tension,[33] increased anger,[25] decreased clear thinking,[8,25,24] and decreased friendliness.[25] Increased subjective sleepiness is a universal finding in total sleep deprivation studies.[54,55] Other mood changes have been typically linked to the sleepiness dimension and include increased fatigue and decreased vigor.[55]

5.2.9 Accumulating deficits

One early study of sleep fragmentation showed that auditory thresholds and ratings of depth of sleep increased significantly both across nights of sleep fragmentation and overall from the first to second night of disturbance.[8] Several other studies have shown greater impairment following a second night of sleep fragmentation as compared to the first night on a number of measures including arousal threshold,[31,25] psychomotor performance,[24] spectral EEG parameters,[33] and mood.[8] Other studies that have not shown significant differences from the first to second disturbance night still typically have shown trend increases in sleepiness following the second night as compared to the first night.[30,38] Such findings confirm that sleep fragmentation produces cumulative effects across time in a manner that is similar to that seen after total sleep deprivation.

5.2.10 Sleep fragmentation versus sleep deprivation deficits

Four empirical studies directly compared total sleep deprivation and sleep fragmentation within the same experiment.[27,23,39,29]

The studies compared total sleep loss for one or two[23] nights with similar periods of sleep fragmentation. Sleep fragmentation was produced following each minute of sleep in 3 studies and at each onset of stage 2 in the fourth. Levine et al[27] reported that MSLT was reduced after both total sleep deprivation and 1-min sleep fragmentation, and these conditions did not differ statistically (the actual latencies were respectively 2.2 and 4.1min). Bonnet[23] reported that performance deficits were significantly greater following total sleep deprivation than following 1-min sleep fragmentation on number of addition problems completed and simple reaction time. However, deficits were similar, i.e., significantly different from baseline but not different from each other, on vigilance hit rate (25% vs. 25%) and on nap latency (2.2 vs. 3.7 min.) after total sleep deprivation as compared to sleep fragmentation. Spath-Schwalbe[39] showed that cortisol and ACTH peaked shortly after the initiation of sleep deprivation or sleep fragmentation and then showed a similar pattern of inhibition followed by a later peak in both sleep deprivation and sleep fragmentation conditions. This pattern was different from the normal sleep condition. Series et al[29] found a significant increase in apnea/hypopnea index after sleep fragmentation, and this increase was actually greater than that found after a similar period of total sleep deprivation. This finding is the single instance in which the effects of sleep fragmentation have been reported to be more extreme than a comparable period of total sleep loss.

5.2.11 Recovery

In early sleep fragmentation studies, the fragmentation procedure typically produced large decreases in SWS and REM.[8] Recovery sleep was also noted to have predominant SWS rebound that was similar in magnitude to that seen after total sleep deprivation.[8] Three more recent studies have sought to produce less change in sleep stages by more careful arousal technique.[33,36,30] Unfortunately, only one of these studies also recorded a recovery night following the fragmentation procedure. Cote et al[33] showed significant decreases in mood and in spectral EEG parameters after sleep fragmentation with no significant changes in TST, SWS, or REM from baseline to fragmentation nights. However, REM percent was decreased on the first fragmentation night as compared to the second fragmentation night, and this change was supported by a significant decrease in REM latency on the second fragmentation night as compared to baseline. The latency to stage 3 was also significantly decreased on the second fragmentation night and recovery night as compared to the first fragmentation night. Such changes indicate that the magnitude of sleep stage recovery values may be dependent upon the disturbance paradigm. However, as all 3 studies that have minimized sleep stage disturbance have nonetheless reported significant changes in sleepiness or mood after the disturbance nights, it is apparent that the sleep fragmentation continued to produce deficits despite lack of impact on traditionally scored EEG sleep stages. Unfortunately, such studies have not been done with control groups experiencing total sleep deprivation or more invasive fragmentation to be able to determine which part, if any, of the large deficits reported in earlier sleep fragmentation studies were related to decreases in sleep stage amounts and which were related to fragmentation itself. This issue of specific sleep stage effects, however, has also been examined in a design which sought to control the amount of sleep fragmentation, while at the same time maximizing or minimiz-

ing SWS. In that study, similar decrements were found after both (equivalent) fragmentation conditions even though the amount of SWS was experimentally varied to be significantly greater in one of the conditions, again supporting the contention that the fragmentation rather than sleep stage amounts was determining the deficits.[24]

5.3 Summary

Both animal and human studies have shown that empirically produced brief EEG arousals produce increased sleepiness and decreased responsiveness to respiratory stimuli. Studies in humans have reported increased objective and subjective sleepiness, decreased psychomotor performance on a number of tasks, changes in hormone secretion pattern, decreased upper airway function, increased whole body metabolic rate, increased sensory arousal thresholds, and decreased amplitude in evoked responses following various schedules of sleep fragmentation as compared to baseline. In all cases, these changes were in the same direction as those seen after total sleep deprivation. Several studies have shown that changes in mood, arousal threshold, psychomotor performance, and spectral EEG parameters, as with sleep deprivation, become significantly more extreme as a function of increasing nights of disturbance. Studies directly reporting outcome measures after high frequency sleep fragmentation as compared to total sleep deprivation have shown that some psychomotor performance measures are not affected as strongly by sleep fragmentation as by total sleep deprivation but that other measures of psychomotor performance and objective sleepiness are equally affected by sleep fragmentation and sleep deprivation. One study[29] has actually shown a greater increase in apnea/hypopnea index after sleep fragmentation as compared to total sleep deprivation. Sleep stage rebounds commonly found after total sleep deprivation are also seen after empirical sleep fragmentation. The extent of sleep stage rebounds seems to be related to the specific decreases in those sleep stages based upon the sleep fragmentation design. However, at least 3 experiments have shown that empiric sleep fragmentation continues to produce next-day deficits, even when total sleep time and sleep stage amounts are preserved in a sleep fragmentation design.

Taken as a whole, these studies indicate that brief, frequent, periodic disturbance of the ongoing EEG during sleep as described by the ASDA[11] produces numerous physiological and behavioral changes similar to those produced by total sleep deprivation.

5.4 Predicted changes in arousals in sleep disorders

Arousals play a prominent role in sleep disorders. A number of clinical studies have examined the direct relationship between arousals in patients and functional outcome variables including changes in function with treatments that decrease arousals. The 16 studies (4 level II,[56,57,42,58] 3 level III,[59-61] and 9 level IV[7,62-69]) that provide clinical evidence for the validity of arousals are summarized in Table 4 (which can be accessed on the web at www.aasmnet.org).

5.4.1 Arousals on baseline and recovery sleep nights

There are significant changes in sleep stages and arousal threshold during recovery sleep following sleep deprivation. If EEG arousals are related to the sleep restoration process, it could

be hypothesized that the number of brief arousals should also be reduced in recovery sleep after sleep deprivation. Two studies have examined this hypothesis. In one,[70] arousal index was significantly decreased from 16.2 to 8.3 during recovery sleep following 64 hours of total sleep deprivation in normal adults. In a second study,[71] arousal index was significantly reduced from 12.3 during baseline sleep to approximately 11 after 2 nights of SWS deprivation and to approximately 8.5 after one night of total sleep deprivation.

5.4.2 Arousals in patients with sleep associated respiratory problems

The above sections examined evidence regarding the impact of arousals on human and animal functioning following experimentally induced sleep fragmentation. In this section, the evidence for a relationship between arousals and other measures of impairment will be examined. The primary questions being asked are: 1) to what extent do arousals predict daytime impairment, and 2) to what extent would the dissipation of arousals following treatment result in a diminution of this impairment?

Stepanski and colleagues[7] performed the first systematic examination of the relationship between sleep fragmentation and daytime sleepiness in various conditions. They utilized correlational techniques to evaluate this relationship in patients complaining of excessive daytime somnolence (EDS) associated with sleep apnea (n = 15), periodic leg movements in sleep (PLMS) (n = 15), patients complaining of insomnia (n = 15), and healthy volunteers with no sleep complaints (n = 10). One night of PSG followed by an MSLT was obtained for each subject. Across all subjects, the total number of arousals was related to the MSLT-based sleepiness index (MSLT sum of latencies to stage 1 sleep on 4 nap tests multiplied by 1.25 and subtracted from 100; r = 0.48, p < 0.001). Martin et al[64] performed both objective (MSLT) and subjective (ESS) measures on 63 OSAS patients and examined arousals per hour of sleep by 3 different criteria (ASDA, mASDA [minimum duration of 1.5 sec], and Cheshire[62]) and R&K awakenings. They found significant relationships between arousals scored by any definition and mean sleep latency on MSLT. However, no relationship existed between arousals or awakenings and ESS. Therefore, the connection between arousals and objective measures of daytime somnolence appears to be supported.

In sleep related breathing disorders, an important question is the relative contribution of EEG arousals, as opposed to the variety of other measures of illness severity such as hypoxemia, AHI, etc. to daytime impairment. In the study by Martin et al,[64] arousals did predict the degree of objective daytime somnolence, and there was no relationship between R&K awakenings and MSLT mean sleep latency. However, the best predictor of MSLT sleep latency was the AHI. Roehrs et al[67] examined this question by performing PSG and MSLT in 466 patients with obstructive sleep apnea syndrome (OSAS). Multiple regression analyses indicated that the respiratory arousal index (RAI) produced higher correlation with MSLT scores than measures of hypoxemia. The best predictor of MSLT was RAI. However, many measures were interrelated. Cheshire et al[62] extended these findings by examining the relationship between arousals and psychomotor functioning and mood in 29 OSA patients (AHI >15). For the former, they administered the Paced Auditory Serial Addition Test, Trailmaking A and B, Inspection Time, Simple Reaction Time, Auditory Verbal Learning Test, National Adult Reading

Test, Weschler Adult Intelligence Scale; for the latter they utilized the Hospital Anxiety and Depression Scale. They noted that the arousal frequency was associated only with Block Design subtest of the WAIS but with no other measure of cognitive functioning or MSLT sleep latency. AHI and SpO_2 measures were associated with different cognitive variables. More recently, Sforza et al[61] compared 152 patients with OSA or primary snoring to 45 healthy controls by administering various measures of daytime sleepiness and psychomotor impairment including the Stanford Sleepiness Scale (SSS), Psychomotor Vigilance Task (PVT), MWT, and Epworth Sleepiness Scale (ESS). They noted that the number of arousals was lowest in snorers and increased with increasing AHI. There was no relationship between arousal frequency and reaction time (RT), yet there was a positive relationship between arousals and false responses, and arousals and lapses, on PVT. However, other measures of sleep related breathing disorder severity, such as AHI, SpO_2, and sleep architectural impairment, such as total sleep time (TST) were also related to measures of daytime impairment. Pitson et al[66] also examined 40 OSAS patients and noted a good correlation between EEG arousals and various other measures of apnea severity such as SpO_2 dips (r=0.69) and AHI (r=0.68). However, the relationship between all types of arousals (in addition to EEG arousals, the authors examined pulse transit time and heart rate arousals) and ESS was poor, being significant only for SpO_2 dips. Non-EEG arousals, such as autonomic arousals, pulse transit time swings, and heart rate rises/hour have been less promising predictors of EDS.[68]

Other investigators have extended these findings by examining the relationship between arousals and other disease-associated impairments. Lafoso et al,[63] for example, studied 105 subjects with the complaint of snoring but not meeting polysomnographic criteria for OSAS (AHI<10, AI<5, duration of episodes of SpO_2 <90% of less than 1 min). Nonapneic snorers were classified as sleep disrupted or not sleep disrupted based on arousal index threshold of 10. The group was assessed for the presence or absence of a history of hypertension (previous diagnosis of hypertension and treatment with antihypertensives, diastolic BP >95 and/or systolic BP >160 within 3 months of the onset of the study). They noted that hypertension was more prevalent in the group with sleep fragmentation and suggested that sleep fragmentation in snorers may play a role in genesis of systemic hypertension. In a separate investigation, Noda et al[65] examined 26 OSAS patients by separating them on the basis of presence or absence of hypertension (BP >140/90). They then assessed possible differences between the 2 groups on various polysomnographic variables such as AHI, total time with SpO_2 below 90%, SpO_2 nadir, EEG arousals, and movement arousals, as well as 24-hr systolic and diastolic BP measured at 30-min intervals. They noted that the frequency of movement arousals and decrements in SpO_2, in that order, made the most significant contributions to increased 24-hr systolic and diastolic BP. The hypertensive group had a higher frequency of EEG and movement arousals.

Taken collectively, the data suggest that sleep fragmentation is an important predictor of daytime impairment in areas such as alertness/wakefulness and psychomotor functioning, and possibly even hypertension. The relationship is better noted in objective, rather than subjective, measures of EDS. However, other measures of apnea severity, such as the extent of hypoxemia and AHI, are also related to daytime function, and the studies cannot conclusively answer the fundamental question of how these pro-

cesses are related to one another and to arousals. That is, they do not answer the question of whether these changes produce their effects on daytime function independently or through their effects on arousals. Also, although EEG arousals are important predictors of EDS in OSA and PLMS, their relative position in the hierarchy of all contributory variables is unclear.

If arousals cause daytime impairment, then their diminution should lead to an improvement in daytime function. Bennett et al[59] assessed daytime somnolence utilizing the ESS and OSLER test in 41 subjects who were comprised of nonsnorers, simple snorers, and OSA patients undergoing treatment with CPAP. In addition to ASDA arousals, they examined autonomic arousals and movement events. All markers of sleep fragmentation (arousals) had significant associations with baseline EDS and its improvement with nCPAP at similar magnitudes. AHI was also correlated closely with sleepiness and its response to treatment. However, no significant relationships could be found between any of the sleep fragmentation indices and post-nCPAP objective sleepiness measures. Multiple regression analysis indicated that the best predictor of the pretreatment objective sleepiness was the movement event index. In a separate investigation, Bennett et al[60] also examined SF-36 health status and noted that it improved following CPAP treatment in 51 OSA patients, and noted a significant, yet weak, relationship between change in health status following CPAP treatment and sleep fragmentation indices. Colt and colleagues[72] examined the relative contribution of arousals and other measures of apnea severity such as hypoxemia and AHI, to EDS in 7 OSA patients undergoing treatment with nCPAP. In this crossover design, patients were exposed to either nCPAP, or nCPAP plus induced hypoxemia for 2 nights. Both conditions resulted in improved MSLT sleep latency to the same extent. Therefore, the induction of hypoxemia did not diminish the improvement in objective measures of daytime somnolence, and this gavefurther support to the relationship between sleep fragmentation and EDS.

The presence of EEG arousals helps to identify upper airway resistance (UARS). One study has specifically examined differences in arousal variables in patients with mild sleep apnea as compared to patients with UARS.[69] The notable finding was that, in terms of EEG arousals, there were almost no differences between the groups. Significant differences in the study were primarily based upon group selection; that is, apnea patients had significantly increased respiratory events associated with visually scored EEG arousals as would be expected by selection.

The only investigation of the relationship between sleep fragmentation and EDS in central sleep apnea (CSA) syndrome was performed by Bonnet et al,[57] who treated 5 male CSA patients with placebo and triazolam (0.125 and 0.25 mg) in a double-blind crossover design. Daytime function was assessed by MSLT, POMS, short-term memory, Wilkinson Addition, Wilkinson Vigilance, and SSS. The triazolam condition was associated with a decrease in arousal and central apnea indices, as well as an improvement in next day psychomotor performance and a reduction in subjective sleepiness on SSS. MSLT sleep latency increased, but the change was not statistically significant. The medication was also associated with an increased total sleep time. The lack of improvement in objective EDS measures, despite improvement in subjective EDS measures and other performance measures in the face of diminished arousals may have been related to the small sample size.

5.4.3 Arousals in patients with periodic limb movements

A number of investigations have examined the relationship of arousals to EDS in patients suffering from periodic limb movement disorder (PLMD) prior to and following treatment. Doghramji et al,[58] in a placebo-controlled crossover design, assessed the effects of triazolam 0.25-0.5 mg in 15 PLMD patients with EDS. At end of the treatment period (4-7 days), triazolam was associated with a significant increase in MSLT mean sleep latency score, yet there was no change in SSS score. Treatment was associated with no change in the PLM index, despite a decrease in the PLM arousal index and total arousal index. Other measures of sleep architecture, such as TST, TWT, and percentage of stage 1 sleep, also revealed improvement. This study showed that improved sleep is associated with improved daytime function, yet it is unclear whether that benefit should be attributed to changes in the arousal frequency alone or to changes in other sleep architectural factors as well. The possibility that arousals are not important contributors to EDS was apparent in an investigation by Bonnet et al[56] in another placebo-controlled crossover investigation of the effects of triazolam 0.125 mg and 0.25 mg in 11 sleepy PLMD patients. They assessed daytime function via auditory vigilance tests, POMS, and short-term memory testing. They noted no change in total leg movements or arousals, yet there was an increase in TST and SE and a decrease in the number of stage changes. A parallel improvement in MSLT sleep latency and cognitive measures was also noted. Therefore, the improvement of daytime performance and alertness appeared to be independent of the medication's effects on arousals. In a similar investigation of 9 PLMD sufferers, which was followed by a 12-week open-label phase, Bonnet and Arand[42] noted that triazolam 0.25 mg did not decrease the arousal index; in fact, the medication was associated with an increase in EEG arousals in the late drug-treatment phase. Nevertheless, daytime function improved as assessed by vigilance performance, short-term memory, and some MSLT results. Sleep continuity variables and TST improved. When EEG arousals were corrected for both the increase in total sleep time and the associated increase in limb movements, there was a significant decrease during early medication administration. However, the fact that EDS and sleep variables improved more than arousals improved suggests that EDS may be based on a variety of nocturnal factors in addition to arousals.

These treatment data yield mixed results. On the one hand, they suggest that arousals are linked, at least at baseline, to severity of EDS and other psychomotor function indicators. However, whether the changes noted in daytime measures during and after treatment are due to changes in arousals is not altogether clear, possibly owing to the complex effects that medications have on a variety of measures of sleep architecture.

6.0 ONTOGENY

EEG arousals are ubiquitous throughout the human life span during sleep. Because failure of arousal has been considered an important issue in sudden infant death, a number of studies have examined the process of arousal in infants. These studies generally do not apply to the current examination of reliability and validity of arousal scoring in adults with the exception that the ASDA scoring criteria have been used in several studies. One study[22] has shown high arousal scoring reliability in infants with only minor scoring modifications, specified word for word in the publica-

tion. In a study of 5-year-old children, arousal reliability was also shown to be high using the standard ASDA criteria.[21]

Specific studies of the reliability or validity of arousals are also uncommon in older adults. Three papers have examined EEG arousals across broad age bands. One study found no significant difference in arousal index when comparing groups of young (20–35 years) and older (50–65 years) normal sleepers.[73] However, 2 other studies have both reported significant positive correlations of EEG arousals with age.[74,75] This correlation was associated with a significant increase in arousal index from 13.8 (SD 2.2) in teenage subjects to 27.1 (SD 3.3) in 60- to 80-year-old subjects in one study.[74]

7.0 OTHER MEASURES OF AROUSAL

It is recognized that visual scoring of EEG arousals according to ASDA criteria is just one approach to characterizing the degree of sleep fragmentation. Scoring EEG arousals is the best-validated approach and has replaced earlier measures used to quantify sleep fragmentation (e.g. amount of stage one sleep, number of awakenings). However, the number of EEG arousals does not always have a high correlation with subsequent daytime sleepiness. New measures may be developed that quantify sleep fragmentaion more precisely and provide more accurate prediction of subsequent daytime impairment. Attempts to improve this measure of sleep fragmentation have primarily involved other measures of EEG activity or measures of autonomic activity. Other approaches to quantifying sleep fragmentation are described below.

7.1 Cyclic alternating pattern

In 1985 the cyclic alternating pattern (CAP) was described as a physiological component of normal NREM sleep.[76] CAP also addressed periodic arousals occurring during sleep and described 3 different types of EEG arousal that occurred spontaneously in cycles as a normal characteristic of sleep. This system focused on NREM sleep alone, viewing these arousals as a characteristic of normal NREM sleep. The role of arousals, according to the theory underlying CAP, differs markedly from that underlying ASDA EEG arousals. CAP researchers believe that arousals occur either as increases or decreases in EEG frequency and that some of these events may protect sleep from disruption. Therefore, these arousals do not necessarily signal decreased sleep quality. As such, the use of the term "arousal" may be confusing.

A CAP scoring manual was published in 2002.[77] It defined 3 subtypes of arousal that constitute the A phase of the CAP pattern which is followed by the B phase of the pattern, a period without arousals. The EEG of the A1 subtype is characterized by a predominance of delta wave bursts, K-complex sequences, and vertex sharp transients. The A2 subtype is an admixture of slow and fast rhythms, while the A3 subtype is a predominance of fast rhythms.

Arousal-related phasic events occur in CAP, interrupting tonic NREM (or NCAP). It is hypothesized that understanding the interplay between CAP and NCAP states is important in appreciating the role of microarousals as normative phenomenon versus signs of pathology. However, the research in this area is at an early stage, and the significance of CAP is unclear. Investigations of CAP have primarily focused on viewing arousal as an oscillating phenomenon with an adaptive function, rather than a measure of the degree of sleep disturbance.

Three studies from those summarized in Table 5 (which can

be accessed on the web at www.aasmnet.org), have measured both ASDA arousals and CAP.[78,79,70] One evidence level IV study showed a statistically significant relation between the occurrence of ASDA arousals and certain subtypes of CAP activity.[78] In particular, subtypes A2 and A3 of CAP had a significant correlation of 0.84 (p<.0001) with ASDA-scored arousals. Two evidence level III studies examined the effect of sleep deprivation on CAP activity.[79,70] DeGennaro et al[79] found that CAP and ASDA arousals were both decreased during recovery sleep following a night of sleep deprivation.[79] Subtype A3 of CAP correlated 0.79 with ASDA arousals on the baseline night and 0.95 on the recovery night. Another study[70] reported a significant decrease in ASDA arousals during sleep following 64 hours of total sleep deprivation but did not find a significant change in either delta bursts or K-complexes (corresponding to the A1 phase of CAP) on the recovery night.

7.1.2 Evidence for validity/reliability of CAP arousals

One level III study has been done assessing CAP interrater reliability in a group of 11 normal adults, and it reported a Kendell W coefficient of concordance value of 0.90 for the various CAP parameters.[80] A standard approach to computing CAP with computer assistance has been recently published,[80] and this should lead to additional systematic studies.

7.1.3 Evidence for relation of CAP to sleep restoration independent of ASDA arousals

There are no data relating CAP measures to severity of daytime impairment. The studies comparing CAP to ASDA arousals have not included measures of daytime function. As a result, there are no validity data available.

7.1.4 Advantages/Disadvantages of CAP Compared with ASDA arousals

In summary, studies measuring CAP have not shown superiority of this approach in the measurement of sleep disruption or as a predictor of daytime impairment, compared with the ASDA arousal measure. Measurement of CAP is more time consuming than scoring of ASDA arousals. It can be argued that the high correlation between the A2 and A3 subtypes of CAP with ASDA arousals makes those measures similar (differing perhaps only in that ASDA arousals are scored during REM while CAP is not). The major theoretical question is whether the CAP A1 subtype, correlated with delta and K-complex activity, is actually an arousal event that might increase the predictive ability of ASDA arousals to identify nonrestorative sleep or a sleep protective event that does not inhibit sleep restoration. Direct tests of the role of K-complexes or evoked delta in limiting sleep restoration have not been published, but one study that actually produced consistent K-complex type responses during sleep at a periodicity of 22 sec for 30 consecutive nights did not report any signs of decreased function or increased sleepiness.[81] This finding, in conjunction with the finding of Sforza et al of no change in these delta events after sleep deprivation[70] supports the contention that these evoked delta and K-complex events are inherently different from ASDA arousals and should not be included with them in traditional arousal scoring. Overall, there is insufficient evidence at this time to use CAP in the routine assessment of arousals during sleep for the purpose of quantifying sleep fragmentation.

7.2 Measures of autonomic arousals

A number of studies have explored the utility of measuring autonomic events, usually blood pressure or heart rate changes, as a way to quantify arousals during sleep. This work is based on the observation that changes in measures of autonomic function, such as increases in arterial blood pressure, occur in response to acoustic stimuli[12] and at the conclusion of obstructive respiratory events.[82] Studies of autonomic arousals have generally addressed one of the following 3 areas: 1) the utility of using autonomic arousals as a screen for sleep disordered breathing, 2) use of autonomic arousals as a more easily scored proxy for EEG arousals, or 3) autonomic arousals as a predictor of daytime sleepiness as compared to EEG arousals. The first of these questions is beyond the scope of this paper and will not be addressed. The other 2 questions are related to the goals of this paper, and relevant studies are reviewed below. Twelve studies, including 6 level III studies,[32,83,12,36,82,81] 5 level IV studies,[84,85,66,69,68] and 1 level V[86] study are summarized in Table 6 (which can be accessed on the web at www.aasmnet.org).

7.2.1 Autonomic arousals and daytime sleepiness

Autonomic arousals occur with obstructive respiratory events and with snoring, even in the absence of EEG arousals.[82,86] Autonomic arousals have also been detected in normal sleepers in response to experimental stimuli when EEG arousals did not occur.[85,83] In contrast, EEG arousals were always accompanied by an autonomic arousal in those studies. These results suggest that autonomic arousals have a lower threshold and are more sensitive to perturbations in the CNS during sleep. Therefore, it is theoretically possible that these events may be a more accurate indicator than EEG arousals of CNS changes sufficient to cause daytime impairment. However, research investigating this topic has not shown a convincing relation between autonomic arousals and sleepiness that is independent of the co-occurrence of EEG arousals.

Pitson and Stradling[66] correlated standard EEG arousals with 2 measures of autonomic arousals (blood pressure arousals and heart rate arousals) in patients undergoing screening for OSA. They found significant correlations between all the arousal measures (0.51-0.67). All arousal measures were also significantly correlated with AHI and SpO_2 dips. No arousal measure was significantly correlated with the ESS, but there was a low significant correlation with SpO_2 dips (r=0.36). Another study investigated the importance of autonomic arousals associated with snoring or sleep-disordered breathing in causing sleepiness.[68] The investigators obtained a large sample (n=473) from the general community and found no relation between blood pressure changes using pulse transit time (PTT) and sleepiness measured by the ESS. One criticism of these studies is the use of a self-report measure of sleepiness that has a low correlation with objective measures of sleepiness.

Experimentally induced autonomic arousals administered to normal sleepers have been reported to increase daytime sleepiness.[36] In this study, tones were presented in a manner to produce increases in blood pressure or heart rate while attempting to avoid producing ASDA arousals; MSLT and MWT were used as outcome measures. It was concluded that events associated with the autonomic arousals, but not EEG arousals, were sufficient to decrease the restorative value of sleep (i.e., a significant decrease was found for MSLT and MWT). However, about 20% of the tones administered in this study caused EEG arousals, and these arousals may have played a role in the increased sleepiness observed in these subjects. The authors argue that because the overall number of EEG arousals did not differ between baseline and disrupted nights, the change in MSLT is solely attributable to the autonomic arousals. This rationale assumes that induced arousals and spontaneous arousals are equal in their distribution and likelihood of contributing to daytime sleepiness. It is commonly observed in studies utilizing experimental sleep fragmentation procedures that spontaneous arousals are decreased on nights when sleep is experimentally fragmented.[81] It has also been shown that longer arousals have a greater physiological impact than shorter arousals.[32] Therefore, the effects of the induced autonomic arousals are confounded by an increase in the number of longer experimentally induced EEG arousals in this study.

7.2.2 Predicting EEG arousals from autonomic arousals

Another area of interest related to autonomic arousals is the possibility that these events can be used to estimate the number of EEG arousals. Scoring EEG arousals is time-consuming and requires a well-trained staff to achieve acceptable reliability. Since autonomic arousals can be computed by online analysis, this approach to quantification of sleep fragmentation is potentially much faster. Catcheside et al[83] found that PTT and skin blood flow (SBF) are superior to HR in predicting EEG arousals. In that study, auditory stimuli administered to normal sleepers produced 85% agreement (measured as ROC area) between EEG arousals and PTT, as well as between blood pressure changes and EEG arousals. However, PTT has not been shown to be a reliable predictor of EEG arousals in patients with sleep disordered breathing.[69] Poyares et al[69] found that PTT predicted EEG arousals with a sensitivity of 90.4% but a specificity of only 16.8%. Adachi et al[84] used various autonomic arousal indices in an attempt to identify patients with frequent EEG arousals (>30 arousals per hour of sleep). They found that autonomic arousal measures tended to overestimate the frequency of EEG arousals, and the greatest mismatch occurred in mild cases with a low number of EEG arousals.

Another aspect of this area of research that requires further assessment is the association between autonomic arousals and EEG changes that do not meet traditional criteria for EEG arousal. The term "subcortical arousal" has been used to describe autonomic measures and differentiate them from traditional EEG-based measures of arousals. However, it may not be accurate to characterize these events as purely "subcortical" since investigators have noted changes in the power density of the EEG in conjunction with autonomic arousals even in the absence of traditionally defined EEG arousals.[82,36] Quantification of the changes in power density of the EEG may be useful in further understanding what types of events are predictive of sleepiness.[82,36]

7.2.3 Summary

Studies comparing autonomic measures with standard EEG measures of arousal consistently show that EEG arousals are associated with reliable changes in autonomic measures. Further, it has been shown that autonomic arousals may be caused by external stimuli as well as obstructive breathing events in the absence of EEG arousals. However, the significance of autonomic arousals that are not associated with EEG arousals in contributing to

daytime impairment is uncertain. Studies have not shown that autonomic arousals without EEG arousals are predictive of daytime sleepiness.

8.0 OVERALL SUMMARY

8.1 Utility of arousal scoring

Several studies have reported that arousals, as defined by the ASDA in 1992, can be scored reliably. A number of variables, such as training and additional recording channels, have been shown to increase scoring reliability. Other variables, such as decreasing the length of time required to score an arousal, have been shown to decrease measure reliability.

A number of empirical studies in animals, humans, and clinical populations have shown links between the frequency of EEG arousals and daytime sleepiness. Increasing the frequency of arousals and the number of nights during which arousals are induced increases residual sleepiness, and this effect is independent from reduction in EEG sleep stage amounts. In addition to sleepiness, sleep fragmentation produces numerous effects such as decreased psychomotor performance, degraded mood, altered hormone secretions, decreased pulmonary function, altered metabolic rate, increased arousal thresholds, and alteration in EEG evoked responses similar to that seen after total sleep deprivation. Sleep fragmentation may also result in sleep stage rebounds or decreased arousals during recovery sleep. Finally, patients with sleep fragmented by respiratory or movement disorders have been shown to have improved alertness and psychomotor function when arousals during sleep have been reduced.

These findings continue to support EEG arousals as an important component of the sleep process. As such, the scoring of EEG arousals is essential in clinical PSGs as a means of determining the extent to which daytime compromise is specifically related to sleep disturbance and to give treatment and treatment response guidelines to practitioners.

8.2 Potential for integration with new measures

Despite the empirical data that link arousals with daytime function, important questions remain. In clinical studies, the correlation between sleep fragmentation measures and daytime sleepiness frequently accounts for only a small part of the total variance. Reliable scoring of arousals is a difficult and time-consuming process. A number of new measures have been proposed to deal with these issues. Investigators have shown that several new measures are correlated at a relatively high level with ASDA arousals and with outcome variables such as sleepiness at a level similar to that seen with ASDA arousals. Unfortunately, because these are correlation studies, the independent effect of these new measures beyond the known effect of ASDA arousals on daytime outcome measures is unknown. For example, automated measures of heart rate or PTT are easy to make and are related to EEG arousals in patients. Unfortunately, no study has effectively produced periodic increases in heart rate or PTT without producing EEG arousals and shown an independent relationship to outcome variables. One study of "nonvisible" arousals[36] did show significant decreases in function on the following day, but 20% of the arousals actually produced traditional EEG arousals. Another study that used tones every 22 seconds for 30 days to produce consistent EEG, heart rate, and finger pulse volume responses did not find signs of sleepiness or performance

decline.[81] Additional empirical studies are needed to explore the relationship between repetitive physiological arousal without EEG arousal during sleep and residual function to understand if this is a graded effect (less sleep restoration as sleep disturbance increases), a threshold effect (physiological activation above a certain level produces decreased restoration), or an interaction between physiological systems. Such studies are also important because they will inform about both the process of sleep restoration and perhaps the key physiological requirements. Finally, this work is required to support a broader definition of arousal. There is currently no compelling evidence to support the scoring evoked delta or K-complex activity as being similar to ASDA arousals. There is insufficient evidence to support the scoring of increases in physiological measures such as heart rate in the absence of EEG changes as compromising restoration during sleep.

8.3 Practical considerations and technical implications

A number of suggestions have been made to improve the ability to score arousals. Perhaps the major issue is training and experience. Because EEG arousals are short, they can be easily confused with artifact or overlooked during scoring. Scorers need directed training on gold-standard sleep recordings; this training should be repeated periodically. Other important issues include whether an occipital channel is included in the recording and which EEG channels are used to score arousals (the number of arousals scored would be expected to be larger if 2 occipital and 2 central channels are recorded and events can be selected from any of the 4 channels compared to a single central EEG channel). Arousal reliability might be further increased if EEG changes were required to be accompanied by a heart rate change.

Another issue that might improve reliability is calibration measurements. Although it is not possible for patients to produce an arousal from sleep response during calibrations, knowing whether the patient produces alpha is important. Production of eyes-closed alpha at a given amplitude could be linked to specific alpha criteria to be used for arousals, while inability to produce such an alpha response could be used to justify less stringent requirements. Such rules might be particularly helpful in patients with very low levels of alpha or with alpha-delta patterns. More specific alpha and EMG requirements would also improve computer scoring of arousals. EMG increases scored for arousals related to REM sleep are usually derived from the submental lead. Typically, the duration of increased EMG is for a minimum of one second. Standardized submental derivations for EMG and specifications for duration of EMG increase would improve arousal scoring, especially in REM where EMG increase is required.

9.0 Unresolved issues and future research

Understanding of the role of brief arousal in the sleep process has the potential to provide much information about the time course and function of sleep. Despite a fairly clear behavioral understanding of the impact of these brief events on the sleep process, little neurophysiological data exist to help understand the mechanism of arousal. Such work may provide more specific suggestions for new measures of arousal or more precise definitions for arousal events.

More precise definition of arousal events and improved computer scoring algorithms will improve automated scoring of

arousals. However, to the extent that computer scored arousal events differ from visually scored events, it will again be important to repeat validity studies to ensure that such computer scored events are the same as visually scored events in their impact upon sleep restoration. In addition, it is likely that EEG arousals vary with some diseases and medications. Additional research will be required to fully define many arousal parameters.

AROUSAL TASK FORCE MEMBERS

The Arousal Task Force members participating in consensus decisions to derive arousal scoring rules included: Michael H. Bonnet (chair), Karl Doghramji, Timothy Roehrs, Steven Sheldon, Edward J. Stepanski, Arthur S. Walters, Merrill S. Wise, and Andrew L. Chesson.

REFERENCES

1. Loomis A, Harvey E, Hobart G. Distribution of disturbance patterns in the human electroencephalogram, with special reference to sleep. J Neurophysiol 1938;13(suppl):231-56.
2. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques, and scoring systems for sleep stages of human subjects. Washington, D.C.: Public Health Service, U.S. Government Printing Office, 1968.
3. Halasz P, Kundra O, Rajna P, et al. Micro-arousals during nocturnal sleep. Acta Physiol Acad Sci Hung 1979;54:1-12.
4. Bowes G, Woolf GM, Sullivan CE, et al. Effect of sleep fragmentation on ventilatory and arousal responses of sleeping dogs to respiratory stimuli. Am Rev Respir Dis 1980;122:899-908.
5. Phillipson EA, Bowes G, Sullivan CE, et al. The influence of sleep fragmentation on arousal and ventilatory responses to respiratory stimuli. Sleep 1980;3:281-8.
6. Carskadon MA, Brown ED, Dement WC. Sleep fragmentation in the elderly: relationship to daytime sleep tendency. Neurobiol Aging 1982;3:321-7.
7. Stepanski E, Lamphere J, Badia P, et al. Sleep fragmentation and daytime sleepiness. Sleep 1984;7:18-26.
8. Bonnet MH. Effect of sleep disruption on sleep, performance, and mood. Sleep 1985;8:11-19.
9. Iber C. Development of a new manual for characterizing sleep. Sleep 2004; 27:190-2.
10. Sackett DL. Rules of evidence and clinical recommendations for the management of patients. Can J Cardiol 1993;9:487-9.
11. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep 1992;15:173-84.
12. Davies RJ, Belt PJ, Roberts SJ, et al. Arterial blood pressure responses to graded transient arousal from sleep in normal humans. J Appl Physiol 1993;74:1123-30.
13. Black JE, Guilleminault C, Colrain IM, et al. Upper airway resistance syndrome. Central electroencephalographic power and changes in breathing effort. Am J Respir Crit Care Med 2000;162:406-11.
14. Halasz P, Terzano M, Parrino L, et al. The nature of arousal in sleep. J Sleep Res 2004;13:1-23.
15. Colrain I. The K-complex: A 7-decade history. Sleep 2005;255-273.
16. Drinnan MJ, Murray A, Griffiths CJ, et al. Interobserver variability in recognizing arousal in respiratory sleep disorders. Am J Respir Crit Care Med 1998;158:358-62.
17. Loredo JS, Clausen JL, Ancoli-Israel S, et al. Night-to-night arousal variability and interscorer reliability of arousal measurements. Sleep 1999;22:916-20.
18. Smurra MV, Dury M, Aubert G, et al. Sleep fragmentation: comparison of two definitions of short arousals during sleep in OSAS patients. Eur Respir J 2001;17:723-7.
19. Thomas RJ. Arousals in sleep-disordered breathing: patterns and implications. Sleep 2003;26:1042-7.
20. Whitney CW, Gottlieb DJ, Redline S, et al. Reliability of scoring respiratory disturbance indices and sleep staging. Sleep 1998;21:749-57.
21. Wong TK, Galster P, Lau TS, et al. Reliability of scoring arousals in normal children and children with obstructive sleep apnea syndrome. Sleep 2004;27:1139-45.
22. Crowell DH, Kulp TD, Kapuniai LE, et al. Infant polysomnography: reliability and validity of infant arousal assessment. J Clin Neurophysiol 2002;19:469-83.
23. Bonnet MH. Performance and sleepiness as a function of frequency and placement of sleep disruption. Psychophysiology 1986;23:263-71.
24. Bonnet MH. Performance and sleepiness following moderate sleep disruption and slow wave sleep deprivation. Physiol Behav 1986;37:915-8.
25. Bonnet MH. Sleep restoration as a function of periodic awakening, movement, or electroencephalographic change. Sleep 1987;10:364-73.
26. Bonnet MH. Infrequent periodic sleep disruption: effects on sleep, performance and mood. Physiol Behav 1989;45:1049-55.
27. Levine B, Roehrs T, Stepanski E, et al. Fragmenting sleep diminishes its recuperative value. Sleep 1987;10:590-9.
28. Magee J, Harsh J, Badia P. Effects of experimentally-induced sleep fragmentation on sleep and sleepiness. Psychophysiology 1987;24:528-34.
29. Series F, Roy N, Marc I. Effects of sleep deprivation and sleep fragmentation on upper airway collapsibility in normal subjects. Am J Respir Crit Care Med 1994;150:481-5.
30. Stepanski E, Lamphere J, Roehrs T, et al. Experimental sleep fragmentation in normal subjects. Int J Neurosci 1987;33:207-14.
31. Bonnet MH. The effect of sleep fragmentation on sleep and performance in younger and older subjects. Neurobiol Aging 1989;10:21-5.
32. Bonnet MH, Berry RB, Arand DL. Metabolism during normal, fragmented, and recovery sleep. J Appl Physiol 1991;71:1112-8.
33. Cote KA, Milner CE, Osip SL, et al. Waking quantitative Electro-encephalogram and auditory event-related potentials following experimentally induced sleep fragmentation. Sleep 2003;26:687-94.
34. Kingshott RN, Cosway RJ, Deary IJ, et al. The effect of sleep fragmentation on cognitive processing using computerized topographic brain mapping. J Sleep Res 2000;9:353-7.
35. Martin SE, Engleman HM, Deary IJ, et al. The effect of sleep fragmentation on daytime function. Am J Respir Crit Care Med 1996;153:1328-32.
36. Martin SE, Wraith PK, Deary IJ, et al. The effect of nonvisible sleep fragmentation on daytime function. Am J Respir Crit Care Med 1997;155:1596-601.
37. Philip P, Stoohs R, Guilleminault C. Sleep fragmentation in normals: a model for sleepiness associated with upper airway resistance syndrome. Sleep 1994;17:242-7.
38. Roehrs T, Merlotti L, Petrucelli N, et al. Experimental sleep fragmentation. Sleep 1994;17:438-43.
39. Spath-Schwalbe E, Gofferje M, Kern W, et al. Sleep disruption alters nocturnal ACTH and cortisol secretory patterns. Biol Psychiatry 1991;29:575-84.
40. Brooks D, Horner RL, Kimoff RJ, et al. Effect of obstructive sleep apnea versus sleep fragmentation on responses to airway occlusion. Am J Respir Crit Care Med 1997;155:1609-17.
41. Bonnet MH, Arand DL. Clinical effects of sleep fragmentation versus sleep deprivation. Sleep Med Rev 2003;7:297-310.
42. Bonnet MH, Arand DL. The chronic use of triazolam in older patients with periodic leg movements, fragmented sleep, and daytime sleepiness. Aging Clin Exp Res 1991;3:313-24.
43. Luboshitzky R, Zabari Z, Shen-Orr Z, et al. Disruption of the nocturnal testosterone rhythm by sleep fragmentation in normal men. J Clin Endocrinol Metab 2001;86:1134-9.
44. Cooper BC, White JES, Ashworth LA, et al. Hormonal and meta-

bolic profiles in subjects with obstructive sleep apnea syndrome and the effects of nasal continuous positive airway pressure (CPAP) treatment. Sleep 1995;18:172-9.

45. Spiegel K, Follenius M, Krieger J, et al. Prolactin secretion during sleep in obstructive sleep apnea patients. J Sleep Res 1995;4:56-62.

46. Espinoza H, Thornton AT, Sharp D, et al. Sleep fragmentation and ventilatory responsiveness to hypercapnia. Am Rev Respir Dis 1991;144:1121-4.

47. Gugger M, Keller U, Mathis J. Arousal responses to inspiratory resistive loading during REM and non-REM sleep in normal men after short-term sleep fragmentation / deprivation. Schweiz Med Wochenschr 1998;128:696-702.

48. Canet E, Gaultier C, D'Allest AM, et al. Effects of sleep deprivation on respiratory events during sleep in healthy infants. J Appl Physiol 1989;66:1158-63.

49. Persson HE, Svanborg E. Sleep deprivation worsens obstructive sleep apnea. Comparison between diurnal and nocturnal polysomnography. Chest 1996;109:645-50.

50. Ryan T, Mlynczak S, Erickson T, et al. Oxygen consumption during sleep: influence of sleep stage and time of night. Sleep 1989;12:201-10.

51. Williams HL, Hammack JT, Daly RL, et al. Responses to auditory stimulation, sleep loss, and the EEG stages of sleep. Electroencephalogr Clin Neurophysio 1964;16:269-79.

52. Corsi-Cabrera M, Arce C, Del Rio-Portilla IY, et al. Amplitude reduction in visual event-related potentials as a function of sleep deprivation. Sleep 1999;22:181-9.

53. Cajochen C, Brunner DP, Krauchi K, et al. Power density in theta/alpha frequencies of the waking EEG progressively increases during sustained wakefulness. Sleep 1995;18:890-4.

54. Johnson LC, Naitoh P, Moses JM, et al. Interaction of REM deprivation and stage 4 deprivation with total sleep loss: experiment 2. Psychophysiology 1974;11:147-59.

55. Bonnet MH, Gomez S, Wirth O, et al. The use of caffeine versus prophylactic naps in sustained performance. Sleep 1995;18:97-104.

56. Bonnet MH, Arand DL. The use of triazolam in older patients with periodic leg movements, fragmented sleep, and daytime sleepiness. J Gerontol 1990;45:M139-44.

57. Bonnet MH, Dexter JR, Arand DL. The effect of triazolam on arousal and respiration in central sleep apnea patients. Sleep 1990;13:31-41.

58. Doghramji K, Browman CP, Gaddy JR, et al. Triazolam diminishes daytime sleepiness and sleep fragmentation in patients with periodic leg movements in sleep. J Clin Psychopharmacol 1991;11:284-90.

59. Bennett LS, Langford BA, Stradling JR, et al. Sleep fragmentation indices as predictors of daytime sleepiness and nCPAP response in obstructive sleep apnea. Am J Respir Crit Care Med 1998;158:778-86.

60. Bennett LS, Barbour C, Langford B, et al. Health status in obstructive sleep apnea: relationship with sleep fragmentation and daytime sleepiness, and effects of continuous positive airway pressure treatment. Am J Respir Crit Care Med 1999;159:1884-90.

61. Sforza E, Haba-Rubio J, De Bilbao F, et al. Performance vigilance task and sleepiness in patients with sleep-disordered breathing. Eur Respir J 2004;24:279-85.

62. Cheshire K, Engleman H, Deary I, et al. Factors impairing daytime performance in patients with sleep apnea/hypopnea syndrome. Arch Intern Med 1992;152:538-41.

63. Lofaso F, Coste A, Gilain L, et al. Sleep fragmentation as a risk factor for hypertension in middle-aged nonapneic snorers. Chest 1996;109:896-900.

64. Martin SE, Engleman HM, Kingshott RN, et al. Microarousals in patients with sleep apnoea/hypopnoea syndrome. J Sleep Res 1997;6:276-80.

65. Noda A, Yasuma F, Okada T, et al. Influence of movement arousal on circadian rhythm of blood pressure in obstructive sleep apnea syndrome. J Hypertens 2000;18:539-44.

66. Pitson DJ, Stradling JR. Autonomic markers of arousal during sleep in patients undergoing investigation for obstructive sleep apnoea,

67. Roehrs T, Zorick F, Wittig R, et al. Predictors of objective level of daytime sleepiness in patients with sleep-related breathing disorders. Chest 1989;95:1202-6.

68. Stradling JR, Barbour C, Glennon J, et al. Prevalence of sleepiness and its relation to autonomic evidence of arousals and increased inspiratory effort in a community based population of men and women. J Sleep Res 2000;9:381-8.

69. Poyares D, Guilleminault C, Rosa A, et al. Arousal, EEG spectral power and pulse transit time in UARS and mild OSAS subjects. Clin Neurophysiol 2002;113:1598-1606.

70. Sforza E, Chapotot F, Pigeau R, et al. Effects of sleep deprivation on spontaneous arousals in humans. Sleep 2004;27:1068-75.

71. De Gennaro L, Ferrara M, Bertini M. EEG arousals in normal sleep: variations induced by total and selective slow-wave sleep deprivation. Sleep 2001;24:673-9.

72. Colt HG, Haas H, Rich GB. Hypoxemia vs sleep fragmentation as cause of excessive daytime sleepiness in obstructive sleep apnea. Chest 1991;100:1542-8.

73. Gosselin N, Michaud M, Carrier J, et al. Age difference in heart rate changes associated with micro-arousals in humans. Clin Neurophysiol 2002;113:1517-21.

74. Boselli M, Parrino L, Smerieri A, et al. Effect of age on EEG arousals in normal sleep. Sleep 1998;21:351-7.

75. Mathur R, Douglas NJ. Frequency of EEG arousals from nocturnal sleep in normal subjects. Sleep 1995;18:330-3.

76. Terzano M, Mancia D, Salati M, et al. The cyclic alternating pattern as a physiologic component of normal NREM sleep. Sleep 1985;8:137-145.

77. Terzano M, Parrino L, Smerieri A, et al. Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. Sleep Med 2002;3:187-99.

78. Parrino L, Smerieri A, Rossi M, et al. Relationship of slow and rapid EEG components of CAP to ASDA arousals in normal sleep. Sleep 2001;24:881-5.

79. De Gennaro L, Ferrara M, Spadini V, et al. The cyclic alternating pattern decreases as a consequence of total sleep deprivation and correlates with EEG arousals. Neuropsychobiology 2002;45:95-8.

80. Ferri R, Bruni O, Miano S, et al. Inter-rater reliability of sleep cyclic alternating pattern (CAP) scoring and validation of a new computer-assisted CAP scoring method. Clin Neurophysiol 2005;116:696-707.

81. Townsend RE, Johnson LC, Muzet A. Effects of long term exposure to tone pulse noise on human sleep. Psychophysiol 1973;10:369-75.

82. Rees K, Spence DP, Earis JE, et al. Arousal responses from apneic events during non-rapid-eye-movement sleep. Am J Respir Crit Care Med 1995;152:1016-21.

83. Catcheside PG, Chiong SC, Mercer J, et al. Noninvasive cardiovascular markers of acoustically induced arousal from non-rapid-eye-movement sleep. Sleep 2002;25:797-804.

84. Adachi H, Mikami A, Kumano-go T, et al. Clinical significance of pulse rate rise during sleep as a screening marker for the assessment of sleep fragmentation in sleep-disordered breathing. Sleep Med 2003;4:537-42.

85. Pitson D, Chhina N, Knijn S, et al. Changes in pulse transit time and pulse rate as markers of arousal from sleep in normal subjects. Clin Sci (Lond) 1994;87:269-73.

86. Lofaso F, Goldenberg F, d'Ortho MP, et al. Arterial blood pressure response to transient arousals from NREM sleep in nonapneic snorers with sleep fragmentation. Chest 1998;113:985-91.

Table 2—Evidence Table for Reliability Studies

Author/Year/ Citation #	Evidence Level	Study Design/ Type	Subjects (N/age /sex (mean, SD)	Outcome Measures	Conclusions
Crowell, D./ 2002/[22]	III	Observational	10 records from 415 infants	arousal scoring reliability improved from 0.71 to 0.83 (kappa) with training	arousal scoring in infants can be reliable
Drinnan, M. J./1998/ [9]	III	Observational	10/NS/8M2F (Excessive snorers and/or daytime sleepiness	Level of agreement showed no significant difference between sleep stage and only moderate agreement between scorers.	Only moderate agreement
Loredo, J. S./1999/ [10]	III	Observational	20/45.6±/16M4F	Interscorer reliability varied with the definition of arousal and ranged with an ICC 0.19 and 0.92. Greatest reliability associated with respiratory events ICC 0.76 to 0.92. ASDA arousal had high interscorer reliability of 0.84. Reliability lowest for arousals consisting of EEG changes lasting < 3 seconds ICC 0.19 to 0.37.	NS
Smurra, M. V/2001/ [11]	III	Observational	20/49±14/M17F3	There was no difference in the arousal indices compared between readers and between definitions by 2-way ANOVA. No difference in the relationship between readers and between definitions by 2-way ANOVA. No difference in the relationship between arousal index scored according to the 2 protocols.	NS
Thomas, R. J/2003/ [12]	III	Retrospective review of PSGs	17/46.3±6.2/17M	Event by event agreement 93.8% for the repeat scoring by the author and 90.5% for the technologist. Agreement dropped to 67.3% to 58.7% when EEG signal alone was used t score arousals as the "cue" effect of respiratory recovery was absent.	Scoring arousals with the 3-second rule may falsely minimize the apparent impact of abnormal breathing on sleep.
Whitney, C. W/1998/ [13]	III	Observational and blinded	NA	Moderately reliable with ICC 0.54% in arousal index and scorer experience ICC+ 0.76 with 2 experienced scorers	SHHS achieved high agreement of interscorer reliability for the scoring of sleep stage and RDI in unattended in home PSG studies
Wong, T. K./2004/ [14]	II	8 hour PSGs standard scoring by Rechtschaffen and Kales criteria	36/5.3±1.4/18M18F	ICC for scoring ASDA arousals was 0.90% indicating excellent interscorer agreement. The ICC for scoring method 1-sec and 2-seconds were 0.35 and 0.42 indicating poor to fair agreement.	Poorer agreement for scoring arousals shorter than 3 seconds when compared to the standard ASDA criteria.
Stepanski, E/1984/ [7]	IV	Patient chart review by sleep disorder group	15 apnea pts (43+10), 15 PLMS pts 55+8), 15 insomnia pts 38+12), 10 normals 49+14)	EEG arousals, MSLT	EEG arousals related to EDS

Table 3—Evidence Table for Validity Studies

Author/Year/ Citation #	Evidence Level	Study Design/Type	Subjects (N/age /sex (mean, SD)	Outcome Measures	Conclusions
Bonnet, M. H./ 1991/[25]	III	Repeated measures	12, 18-28	Nap Lat, additions, Vig, rt. , mood	Frequent fragmentation increases EDS, but infrequent does not
Bonnet, M. H./ 1986A/[16]	II	1-ADAPT, then 5 nt protocol, repeated x3: 1-BSL, 2-FRAG, 2-REC Repeated measures 20 lab nights/S (BSL, 2 SF or TSD nights, 2 recovery nights * 4 conditions). Conditions = TSD, SF 1 min, SF each 10 min, 2.5 hrs sleep followed by SF at each sleep onset. Ss awakened (verbal response) by audiometer	8 M Ss, 18-28	Perf battery, Single nap latency	Perf decreased after TSD and SF; in some cases SF reductions as great as TSD (Hit Rate, Nap Lat) and in some cases not (RT, Adds)
Bonnet, M. H./1987/[18]	II	1-ADAPT, then 4 nt protocol, repeated x3: 1-BSL, 2-FRAG, 1-REC	11 young adults	Perf battery, Single nap, Mood, SSS	EDS the same regardless of condition, but Ss felt sleepier after the awakening condition
Bonnet, M. H/1985/[8]	III	6 nt protocol: 1-ADAPT, 1-BSL, 2-FRAG, 2-REC	11 young adults	WA, DSST, SRT, SSS, Mood, threshold	Decreased perf and increased threshold after frag
Bonnet, M. H/1986B[17]	II	5 nt protocol: 1-ADAPT, 1-BSL, 2-FRAG, 1-REC	12 young adults	Perf battery, single nap latency, SSS	Both frag conditions showed significant and equal decreases in performance
Bonnet, M. H/1989A/[24]	III	5 nt protocol: 1-ADAPT, 1-BSL, 2-FRAG, 1-REC measures: Awakening: 10 mins solid sleep, then 20 mins FRAG (1 per min)	24/ 12 young (18-28 yrs); 12 older (55-70 yrs)	Perf battery, Single nap, Mood, SSS	Older subjects were less affected by frag
Bonnet, M. H/1989B[19]	II	Repeated measures; Study #1 & #2: 5 nt protocol: 1-ADAPT, 1-BSL, 2-FRAG, 1-REC	12 young adults; #2: 7 young adults 18-28	Nap lat, Adds, Vig, mood, metabolic rate	SF increased nocturnal metabolic rate and daytime EDS
Bowes, G./1980/[4]	III	Repeated measures	5 dogs	Return to sleep latency, respiratory variables	Frag affects sleepiness and respiratory parameters
Brooks, D/ 1997/ [33]	III	repeated measures	animal study does not meet usual (4 dogs)	airway pressure, SaO_2; BP, time to arousal	sleep apnea and control sleep fragmentation produced similar effects
Cote, K. A./2003/[26]	III	Repeated measure; bl, SF1, SF2, REC	8/33/4M4F	EEG spect, ERP, mood, 10 min RT, 5 min serial additions, 4 min discrimination	EEG measures more sensitive than the perf measures used
Kingshott, R. N./2000/[27]	III	1- ADAPT, 1-FRAG; 1-ADAPT, 1-BSL: 2 nt sequences counterbalanced	8 young adults	MSLT, Mood, SSS, Cognitive function, Evoked potentials	No change in MWT or P300 latency with FRAG. Decreased P300 amplitude. Increased sleepiness on SSS. Decreased mood with FRAG
Levine, B../1987/[20]	II	Single 3 hour nap given after night of sleep; between Ss, 8 Ss in each of 5 groups (no nap, or naps with frag each 1, 3 or 5 min or no frag) following a night of TSD. Frag started after 10 min of normal sleep. (1-sec tone every 10 sec at 90 dB)	40/18-35/20M20F	MSLT	Lat after frag1 similar to that seen after no sleep
Magee, J./1987/[21]	II	Parallel groups	24/22 yrs/24 M	MSLT BL4 = 9.1 se = 2.3, BL5 = 9.6; 4-min BL4 = 8.8 se = 2.2, BL5 = 9.4; 1-min BL4 = 4.0 se = 0.8, BL5 = 4.3	Frequent disturbance causes EDS
Martin, S. E., Engleman,/1996/[28]	III	Repeated measures; 1-ADPAT, 1-BSL, 1- ADAPT, 1-FRAG	16/ 24+3 yrs/ 8M8F Epworth = 4	MSLT; MWT, mood, dsst, trailmaking, steer clear, PASAT	SF increases EDS

Martin, S./1997/[29]	III	Prospective, repeated measures study	12, 25 + 6 yrs; 7M, 5F	MSLT, MWT, SSS	One night of induced autonomic arousals cause increased daytime sleepiness on objective testing
Philip, P./1994/[30]	III	Baseline and fragmentation in normals; 1-BSL, 1-Frag, not consecutive nts, counter-balanced	8 (24,5 + 3.5) Ss, BMI = 25.4	MSLT, tapping, dsst, adds, memory	MSLT reduced
Phillipson, E. A./1980/[5]	III	Repeated measures	5 dogs	Return to sleep latency, respiratory variables	Fag affects sleepiness and respiratory parameters
Roehrs, T./1994/[31]	III	3 nt protocol: 1-BSL, 2-FRAG	36/ NS/21-35	MSLT, divided attention	Significant impact on alertness
Series, F./1994/[22]	II	Repeated measures, emp sleep fragmentation and TSD; Base, SF, Sfrec, TSD, TSDrec; Frag: aud – 1 min after each appearance of K-complex or REM until EEG arousal	8/20-27 yrs/5M3f BMI = 24.1	AHI, Pcrit	Sleep fragmentation results in greater airway collapsibility than does TSD
Spath-Schwalbe, E.,/1991/[32]	III	Repeated measures: baseline, sleep deprivation (Ss awakened at onset of 2nd REM and kept awake), slp frag (Ss awakened at onset of 2nd REM and then after 'a few' min st 1 or onset of st 2) via a tone (intensity and other parameters not included)	10/23/28	ACTH, cortisol	Awakening (TSD) or SF produces a burst then inhibition of cortisol and ACTH not seen in normal sleep
Stepanski, E./1987/[23]	II	Repeated measures; 3 nt protocol: 1-BSL, 2-Frag	6/26 yrs/6M/ (18-32)	MSLT, aud threshold; RT, VIG	Differences for frag but could not differentiate conditions

Table 4—Evidence Table for Sleep Disorder Arousal Validity Studies

Author/Year	Evidence Level	Study Design/Type	Subjects (N/age /sex (mean, SD)	Outcome Measures	Conclusions
Bennett, L. S./1998/[52]	III	Uncontrolled treatment trial (active treatment = nCPAP, 4 weeks), comparison of pre- and post-treatment data	41/49/36M 5F	EDS on OSLER and ESS, ASDA arousals, neural network EEG analysis, sleep descent index, standard deviation of sleep depth, autonomic arousal index, movement event index	Sleep fragmentation indices are strongly related to sleepiness on OSAS, particularly movement number of body movements per hour.
Bennett, L. S./1999/[53]	III	Uncontrolled treatment trial (active treatment = CPAP), comparison of per- to post-treatment values	51/49/46M 5F	SF-36 dimensions, arousal index, respiratory arousal index, PTT (autonomic events), AHI	Arousals and other measures of OSA severity are related to health status, yet relationships are weak
Bonnet, M. H./1990/[49]	II	RCT/ DB, placebo controlled, repeated measures, crossover.	11/65 (57-74)/7M 4F	Sleep architecture, EEG arousals, subjective sleep measures, No. of leg movements, cognitive measures	Medication changed daytime performance and alertness measures despite lack of change in arousals
Bonnet, M. H./1991/[35]	II	RCT/ DB, placebo controlled, crossover, followed by 12-week open-label.	9/66(59-76)/7M 2F	Sleep architecture, EEG arousals, subjective sleep measures, No. of leg movements, cognitive measures	Triazolam improved sleep continuity variables with no tolerance evident at 12 weeks. However, given the negative effect on arousals, no conclusions can be made regarding relationship between arousals and EDS. The fact that EDS and sleep variables improved independently of arousals suggests that EDS is based on a variety of nocturnal factors in addition to arousals.
Bonnet, M. H/1990/[50]	II	Repeated Measures	5/70/5M	Sleep, performance and arousals	Triazolam consolidation sleep and decreased apnea and arousal
Cheshire, K./1992/[55]	IV	Study of a group of patients with OSAS	29/50/(35-74)/25M 4F	NS	Arousals are related to one measure of cognitive function in OSA patients, but not to many others. Significance of this finding is unclear 11. Biases: Psychometric tests performed on day other than MSLT. Were patients medication-free?
Doghramji, K/1991/[51]	II	RCT/ DB, placebo controlled, crossover. Multicenter	15/45.8/9M6F/10.7	Indexes and arousal indexes of PLMs, I (isolated) LMs, PSG sleep architectural variables, MSLT sleep latency, SSS score	Following treatment, reduction in arousal frequency is associated with a decrease in objective measures of daytime somnolence.
Lofaso, F/1996/[56]	IV	Study of a group of pts with snoring	105/unclear/82M 33F	Presence or absence of hypertension (previous dx of HTN and treatment with antihypertensives, diastolic BP > 95 and/or systolic BP > 160 within 3 months).	Sleep fragmentation in snorers may play a role in genesis of systemic hypertension
Martin, S/1997/[57]	IV	Survey of group of patients	63/49 SD 10/55M 8F	Microarousals per hour of sleep by 3 different criteria (ASDA, mASDA [minimum duration of 1.5 sec], Cheshire) and R&K awakenings, MSLT sleep latency, ESS score	Arousals are predictive of degree of objective, but not subjective, daytime somnolence. Awakenings are not.

Noda, A/2000/[58]	IV	Survey of group of patients separated into 2 groups on basis of presence or absence of hypertension (BP > 140/90, Gps I and II)	26/51.8 + 7.8/All male	AHI, Total time SpO$_2$ below 90%, SpO$_2$ nadir, EEG arousals, movement arousals, 24-hr systolic and diastolic BP measured at 30-min intervals	Movement arousals may play in important role in increase in 24-hr BP in OSAS patients
Pitson, D. J./1998/[59]	IV	Convenience sample; cross-sectional study	40/ 48 yrs/34M 6F (no SD; range 23-66)	ESS	Since BP arousals appear as reasonable estimate of EEG arousals, suggest that because these are easier to score they might be used in patients undergoing study for OSA.
Poyares, D/2002/[62]	IV	convenience sample; cross-sectional study	20, 15 male, 5 female, age range 18-50 10 with UARS, 10 with OSA	ASDA arousals, 1.5-3 sec 'micro'arousals, PTT changes	No diff between UARS and apnea pts in arousals except that apnea pts had more respiratory events with arousals
Roehrs, T/1989/[60]	IV	Survey of a group of patients with OSAS	466/52.5 +12.2/415M 51F	MSLT mean sleep latency, respiratory arousals, various PSG measures	Sleep fragmentation is an important predictor of daytime sleepiness in OSAS
Stepanski, E/1984/ [7]	IV	Patient chart review by sleep disorder group	15 apnea pts (43+10), 15 PLMS pts 55+8), 15 insomnia pts 38+12), 10 normals 49+14)	EEG arousals, MSLT	EEG arousals related to EDS
Sforza, E./2004/[54]	III	Correlational	control 35 M, 43+13; 152 apnea pts 99M 53+11	MWT, PVT	arousals, SpO$_2$, ESS, MSLT all correlated about the same with lapses
Stradling, J. R/2000[61]	IV	Cross-sectional	473 adults, age: median 52 for males, 48 for females	ESS	Autonomic arousals associated with snoring or SDB are not a major cause of sleepiness in the community

Table 5—Evidence Table for CAP Studies

Author/Year	Evidence Level	Study Design/Type	Subjects (N/age /sex (mean, SD)	Outcome Measures	Conclusions
De Gennaro, L/2002[72]	III	Repeated measures	9/23.3 (SEM: 0.4)	No measures of daytime function	CAP is sensitive to effects of SD; this measure may be redundant with ASDA arousals, or may be sensitive to different factors
Ferri, R/2005/[73]	III	Observational	11/20-32/4M7F	CAP quantified by Kendall with coefficient values above 0.9 for total CAP time; CAP rate showed high value 0.829	CAP scoring shows interrater reliability and might be compared in different labs
Parrino, L/2001/[71]	IV	Within subject	40/ 38 (20)	No measures of daytime function	CAP scoring encompasses both sleep fragmentation (A2/A3) and sleep maintenance components (A1) of NREM sleep
Sforza, E/2004[63]	III	64 hours TSD; repeated measures	12/27.4+7.9/12M	NS	SD reduces arousals

Table 6—Evidence Table for Autonomic Studies

Author/Year	Evidence Level	Study Design/Type	Subjects (N/age /sex (mean, SD)	Outcome Measures	Conclusions
Adachi, H./2003/[77]	IV	Survey of group of patients	33/49.1+ 13.1/29M 4F	AHI, breathing-related EEG arousals, pulse rate rise index (PRRI)	EEG arousals are closely associated with transient increases in pulse rate during sleep in suspected OSAHS patients
Bonnet, M./1991/[25]	III	repeated measures	12M, 18-28	Nap Lat, adds, vig, rt. , mood	frequent fragmentation increases EDS, but infrequent does not
Catcheside, P. G./2002/[76]	III	Convenience sample	11/20.3 + 0.5 SEM/11M	No measures of daytime function	Autonomic measures can be proxies for EEG arousals
Davies, R. J./1993/[12]	III	Cross-sectional study	5/18-29/3M2F	No measures of daytime function	There is a lower threshold for BP change (i.e. autonomic arousal) than for EEG arousal.
Lofaso, F./1998/[79]	V	Cross-sectional study	6/45 + 11/6M	ESS	Nonapneic snoring leads to BP increases, and these may occur without EEG arousal
Martin, S. E./1997/[29]	III	Prospective, repeated measures study	12/25 + 6 yrs/7M5F	MSLT, MWT, SSS	One night of induced autonomic arousals cause increased daytime sleepiness on objective testing
Pitson, D. J./1998/[59]	IV	Convenience sample; cross-sectional study	40/ 48 yrs/34M6F/ (no SD; range 23-66)	ESS	Since BP arousals appear as reasonable estimate of EEG arousals, suggest that because these are easier to score they might be used in patients undergoing study for OSA.
Pitson, D./1994/[78]	IV	Random sample; normal sleepers	8/19-30/8M/(mean: 22 yrs, no SD	No measures of daytime function	PTT changes may be more easily detected indicator of CNS arousal during sleep; authors acknowledge that utility of this depends on relation of symptomatic change to arousals as measured by PTT change
Poyares, D./2002/[62]	IV	Convenience sample; cross-sectional study	20/18-50/ 15M5F	No measures of daytime function	PTT alone is insufficient as a marker of arousals or respiratory events
Rees, K./1995/[75]	III	Cross-sectional study	15 /43.8 + 2.66 (SEM)/ 13M 2F	No measures of daytime function	The hierarchy of arousal is that events are seen in the autonomic system prior to being seen in the EEG
Stradling, J. R./2000/[61]	IV	Cross-sectional	473 adults/age: median/52 for Males, 48 for females	ESS	Autonomic arousals associated with snoring or SDB are not a major cause of sleepiness in the community
Townsend, R. E./1973/[74]	III	Repeated measures; Ss exposed to 80, 85, & 90 dB tones every 22 sec for 30 days	20/18-33 y/o/20M/ (physio monitoring only in 10)	Sleep stages, subjective report, heart rate, finger pulse, evoked responses	Individual physiologic response without adaptation and without residual consequences

Abbreviations: ADAPT=adaptation, BSL=baseline, BP=blood pressure, CAP=cyclic alternating pattern, dx=diagnosis, frag=fragmentation, GPs=groups, ICC= intraclass correlation coefficient, LMs=leg movement, MSLT=multiple sleep latency test, Nap Lat=nap latency, nt=night, perf=performance, PLMs=periodic leg movements, PSG=polysomnography, PTT=pulse transit time, RCT=randomized controlled trial, REC=recovery, SF=sleep fragmentation, Ss= subjects, SSS=Stanford Sleepiness Score, TSD=total sleep deprivation, UARS=upper airway resistance syndrome, Vig=vigilance