JCSM
Journal of Clinical Sleep Medicine

The Visual Scoring of Sleep in Adults

Michael H. Silber, M.B.Ch.B.[1]; Sonia Ancoli-Israel, Ph.D.[2]; Michael H. Bonnet, Ph.D.[3]; Sudhansu Chokroverty, M.D.[4]; Madeleine M. Grigg-Damberger, M.D.[5]; Max Hirshkowitz, Ph.D[6]; Sheldon Kapen, M.D.[7]; Sharon A. Keenan, Ph.D.[8]; Meir H. Kryger, M.D.[9];Thomas Penzel, Ph.D.[10]; Mark R. Pressman, Ph.D.[11]; Conrad Iber, M.D.[12]

[1]Mayo Clinic College of Medicine, Rochester, MN; [2]University of California, San Diego,CA; [3]Dayton Department of Veterans Affairs Medical Center, Wright State University, and Kettering Medical Center, Dayton, OH; [4]New Jersey Neuroscience Institute at JFK Medical Center, Edison, NJ; [5]University of New Mexico School of Medicine, Albequerque, NM; [6]Baylor College of Medicine & VAMC, Houston, TX; [7]Wayne State University Medical School and VAMC, Detroit, MI; [8]The School of Sleep Medicine, Inc., Palo Alto, CA; [9]University of Manitoba, Winnipeg, Canada; [10]University of Marburg, Marburg, Germany; [11]Lankenau and Paoli Hospitals, Wynnewood, PA; [12]University of Minnesota, Minneapolis, MN

Abstract: The 1968 Rechtschaffen and Kales (R & K) sleep scoring manual was published 15 years after REM sleep was discovered. Advances in the ensuing 28 years warranted a re-look at visual scoring of sleep stages. This paper describes the work of the AASM Visual Scoring Task Force, including methodology, a literature review and the rationale behind the new rules. Reliability studies of R & K scoring were reviewed; reliability was low for stage one and moderate for slow wave sleep. Evidence indicated that K complexes and slow waves are expressed maximal frontally, spindles centrally and alpha rhythm over the occipital region. Three derivations of EEG, two of electro-oculography, and one of chin EMG were recommended. Scoring by 30-second epochs was retained. New terminology for sleep stages was proposed. Attenuation of alpha rhythm was determined to be the most valid electrophysiological marker of sleep onset. Alternative measures were proposed for non-alpha generating subjects. K complexes associated with arousals were determined to be insufficient alone to define the new stage N2. No evidence was found to justify dividing slow wave sleep into two stages. No reasons were found to alter the current slow wave amplitude criteria at any age. The phenomena of REM sleep were defined. The rules for defining onset and termination of REM sleep periods were simplified. Movement time was eliminated and major body movements defined. Studies are needed to test the reliability of the new rules. Future advances in technology may require modification of these rules with time.

Keywords: Sleep scoring; sleep stages; Non-REM sleep; REM sleep; eye movements.

Citation: Silber MH; Ancoli-Israel S; Bonnet MH et al. The visual scoring of sleep in adults. *J Clin Sleep Med* 2007:3(2);121-131

Disclosure Statement

This is was not an industry supported study. Dr. Ancoli-Israel has received research support from Takada; is a consultant, on the scientific advisory board, and/or speakers bureau for Ferring Pharmaceuticals, King Pharmaceuticals, Merck, Neurocrine Biosciences, Neurogen, Sanofi-Aventis, Sepracor, Somaxon, and Takeda. Dr. Bonnet is on the advisory board for Jazz Pharmaceuticals; has participated in speaking engagements for Sanofi-Aventis and Takeda; and has received research support from Cephalon. Dr. Chokroverty is a member of the advisory board of Sanofi-Aventis and has participated in speaking engagements for Boehringer-Ingelheim. Dr. Grigg-Damberger has financial interests in GlaxoSmithKline and Sanofi-Aventis. Dr. Hirshkowitz is a member of the speakers bureaus for Sanofi-Aventis, Takeda, and Cephalon; has participated in speaking engagements for Exxon Mobile, Atlanta School of Sleep Medicine, Palo Alto School of Sleep Medicine, and the American Academy of Sleep Medicine; and is the principal investigator on research managed by Baylor College of Medicine including contract research from Evotec, Neurogen, Sanofi-Aventis, GSK, Merck, Takeda, NBI, Respironics, and Organon. Dr. Keenan is the president and CEO of the School of Sleep Medicine Inc; has financial interests in SSM, Inc; and has participated in speaking engagements for the Montana Regional Sleep Society, Sydney University of Dept of Medicine, Mount Hospital, and Perth Sleep Clinic. Dr. Penzel is on the board of directors of Advanced Sleep Research Berlin and has participated in speaking engagements for and has received research support from Respironics and Weinmann. Drs. Silber, Kapen, Kryger, Pressman, and Iber have reported no financial conflicts of interest.

Submitted for publication February 1, 2007
Accepted for publication March 15, 2007

Address correspondence to: Michael H. Silber, M.B.Ch.B., Mayo Sleep Disorders Center, Mayo Clinic College of Medicine, 200 First St. SW, Rochester, MN 55902, E-mail: msilber@mayo.edu

1.0 BACKGROUND

In the more than a third of a century since the Rechtschaffen and Kales manual on sleep stage scoring was published, much has changed in sleep science and medicine. Digital polygraphs capable of recording 16 or more channels of data have replaced paper based analog four channel studies. Researchers into sleep science have been outnumbered by clinicians studying patients with a wide spectrum of sleep disorders. What was predominantly a fascinating and untapped field of human physiology has developed into a scientific discipline and a recognized medical specialty with accredited centers and training programs and increasing public recognition. Despite this, the scoring of sleep stages has largely followed the rules set out by Rechtschaffen and Kales in 1968.[1]

In 2004, the American Academy of Sleep Medicine commissioned a revision of sleep scoring rules, covering not only sleep stages but also the scoring of arousals, respiratory events, sleep related movement disorders and cardiac abnormalities. Under the overall supervision of a steering committee,[2] six task forces were established. Sleep stage scoring was assigned to two task

Table 1—Evidence Levels Adapted from Sackett[4]

1. Well-controlled studies (including blinded analysis and randomization, when relevant); low alpha* error (adequate statistical analysis and level of significance of result) and low beta* error (adequate sample size); comparison to a reference standard; well-defined, homogeneous samples
2. Well controlled studies as in 1 but with higher beta error (smaller sample size)
3. Less-well controlled studies; do not use reference standards; have less well-defined or non-homogeneous samples
4. Uncontrolled and observational studies with reasonably well-defined samples, adequate sample size, and standardized techniques
5. Case reports, case series, or observational studies not fulfilling the criteria of 4

*Alpha (type I error) refers to the probability that the null hypothesis is rejected when in fact it is true (generally acceptable at 5% or less, or $p<0.05$). Beta (type II error) refers to the probability that the null hypothesis is mistakenly accepted when in fact it is false (generally trials accept a beta error of 0.20). The estimation of type II error is generally the result of a power analysis. The power analysis takes into account the variability and the effect size to determine if sample size is adequate to find a difference in means when it is present (power generally acceptable at 80-90%).

forces, one to examine visual scoring rules and the other digital analysis. In addition the Geriatric and Pediatric Task Forces had liaison representatives on each task force. This paper summarizes the work of the Visual Scoring Task Force (see page 129), with added insights from the Geriatric and Pediatric Task Forces. It will cover the methods used and the history of sleep staging. The reasoning behind the new proposals will be discussed, first in terms of evidence gleaned from the literature and then in terms of the consensus based decision making that resulted in the new proposed rules.

2.0 METHODS

The Visual Scoring Task Force, consisting initially of 10 members, was appointed by the AASM Scoring Manual Steering Committee in 2004. Two additional members were added in June 2005. The task force met by conference call on 8 occasions between October 2004 and September 2005 and face-to-face once in June 2005. Discussions totaled approximately 12 hours. A computer-based PubMed literature search was performed for all human studies in English published between 1968 and September 2004 using the following key words: REM sleep, stage 4 sleep, stage 3 sleep, stage 2 sleep, stage 1 sleep, sleep onset, alpha AND sleep, delta AND sleep, drowsiness AND normal, eye blinks, eye movements AND sleep, K complex, spindles, sleep staging. Well over 1,000 articles were identified, 26 of which were considered relevant to the topic after initial review. Only papers deemed relevant to the definition and scoring of sleep stages and phenomena were included. Standardized evidence tables (which can be accessed on the web at www.aasmnet.org) were prepared and evidence levels assigned to each study (see Table 1). Approximately 93 additional background articles were also identified and reviewed. As most articles provided only indirect evidence pertinent to sleep stage scoring, the differentiation of evidence-based articles from background material was sometimes arbitrary, and both groups are summarized in this paper. The task force followed a modified RAND/UCLA Appropriateness Method.[3] Eventually, 71 questions were formulated covering 32 major topics, including montage options, analysis domains, waveform definitions, and multiple practical scenarios, especially sleep state boundary issues. After the face-to-face meeting in June 2005, the task force voted by secret ballot on each question at least once. In contrast to the usual RAND/UCLA method, extensive discussion of the issues occurred prior to any voting. If the results showed an unequivocal preference of the group for an outcome based on standard RAND/UCLA methodology, a second ballot was not taken. If no clear choice emerged from the first vote, further discussions were held and a second round of voting taken, often with a modified question. New scoring rules were proposed in September 2005 based on the results of these votes. In the discussions leading up to the ballots, the task force followed certain guiding principles: proposed rules should be compatible with published evidence; they should be based on biologic principles; they should be applicable to clinical disorders; and they should be easily applied by practicing sleep medicine specialists, scientists, and technologists. Between October and December 2005, a subgroup of the task force with input from the Geriatric and Pediatric Task Forces crafted this paper, covering the background to the topic, a summary of the evidence based material, and the reasoning behind the choice of scoring methods and criteria. The manuscript was circulated to the entire task force and submitted for outside review prior to approval by the AASM Board of Directors.

3.0 HISTORICAL PERSPECTIVE

A 19th century Liverpool physician, Richard Caton was first to demonstrate how "feeble currents of varying direction" could be recorded from the exposed brain surface of every dog and rabbit he studied.[5] A half century later in his seminal 1929 paper, the German psychiatrist Hans Berger reported recording electrical activity from the scalp of human subjects through an intact skull.[6] Berger was the first to identify and name the "alpha rhythm" (an electrical rhythm which oscillates at a frequency of 8-13 Hz generated over the occipital scalp regions in humans during a state of relaxed wakefulness with eyes closed). Berger was also the first to report that the patterns of EEG activity in humans changed with sleep. His work was almost entirely disregarded and even ridiculed until Adrian and Matthews (1934) confirmed the validity of "Berger's rhythms."[7]

Loomis et al in 1937 were the first to describe and classify NREM sleep,[8-10] recognizing 5 different stages. *Stage A* was characterized by the first fragmentation of awake alpha EEG activity, while alpha activity disappeared in *stage B*. They reported recording slow rolling eye movements during stages A and B with an electrode over the left eyebrow. Sleep spindles first appeared in *stage C*, and sleep spindles were intermixed with slower waves in *stage D*. Increasing amounts of large amplitude delta slowing with or without sleep spindles appeared during *stage E*.

Blake et al,[11] studying EEG and behavior in 8 normal subjects and 11 patients with narcolepsy, suggested a different method for classifying sleep by averaging consecutive 5 minute periods "through many nights of sleep." They defined the following stages of sleep: *awake* by 10 Hz alpha activity; *light sleep* by alpha, delta, and 14 Hz activity (sleep spindles), with alpha persisting after the appearance of delta; *deep sleep* by 0.5-5 Hz delta and 14-Hz (sleep spindle) activity; *null* by low voltage EEG without

Table 2—Rechtschaffen and Kales Sleep Staging Criteria

Sleep Stage	Scoring Criteria
Waking	>50% of the page (epoch) consists of alpha (8-13 Hz) activity or low voltage, mixed (2-7 Hz) frequency activity.
Stage 1	50% of the epoch consists of relatively low voltage mixed (2-7 Hz) activity, and <50% of the epoch contains alpha activity. Slow rolling eye movements lasting several seconds often seen in early stage 1.
Stage 2	Appearance of sleep spindles and/or K complexes and <20% of the epoch may contain high voltage (>75 μV, <2 Hz) activity. Sleep spindles and K complexes each must last >0.5 seconds.
Stage 3	20%-50% of the epoch consists of high voltage (>75 μV), low frequency (<2 Hz) activity.
Stage 4	>50% of the epoch consists of high voltage (>75 μV) <2 Hz delta activity.
Stage REM	Relatively low voltage mixed (2-7 Hz) frequency EEG with episodic rapid eye movements and absent or reduced chin EMG activity.

delta activity; *sleep to wake* by intermittent alpha; *awake* by alpha again but lower intensity.

Gibbs and Gibbs[12] were the first to emphasize age-related differences in sleep onset EEG in adults and children. They identified vertex waves (calling them biparietal humps) and positive occipital sharp transients of sleep (POSTS). They devised their own classification of sleep: vertex waves and POSTS characterized v*ery light sleep;* increase in slow wave activity and decreased numbers of sleep spindles were seen in *moderately deep sleep*; diffuse high voltage delta activity was seen in *very deep sleep;* and *early morning sleep* was characterized by "an EEG pattern difficult to distinguish from normal waking," probably representing as yet unnamed REM sleep.

Emmons and Simon[13] provided a modification of the classification of Loomis et al,[9] based on a study of 21 normal adults with well-defined occipital alpha activity while awake. They renamed wakefulness as stage 0. When alpha activity occupied less than 50% of a 10-second epoch of EEG, they named this *stage A+* and thought it represented a light drowsy state; *stage A-* represented a deep drowsy state characterized by either continuous or discontinuous slower alpha (approximately 2 Hz slower than in level 0). Roth[14] subdivided Loomis stage B into *Stage 2a* when EEG activity flattened; *Stage 2b* characterized by 10-40 μV amplitude 5-6 Hz activity; and *Stage 2c* characterized by 50-80 μV amplitude 3-4 Hz activity.

In 1949 Eugene Aserinsky, then a graduate student in physiology at the University of Chicago, was encouraged by his faculty advisor Nathaniel Kleitman to study eye movements in sleep. By 1952, he[15] had recorded more than 50 studies on about two dozen infants. He discovered that "jerky eye movements" occurred during sleep and were associated with 10% increases in average heart rates and 20% increases in respiratory rates. In 1953, Aserinsky and Kleitman reported how rapid eye movements occurred in regular cycles across a night in sleeping adult subjects.[16]

In 1957, Dement and Kleitman published a landmark paper,[17] identifying REM sleep by the presence of rapid eye movements and low voltage fast EEG activity, and observing that NREM and REM sleep alternate cyclically across a night of sleep. In the same paper, they proposed the first classification of sleep stages that included REM sleep, defining four stages. *Stage 1* was characterized by an absolute lack of spindle activity, generally with a low voltage, relatively fast pattern corresponding to the A and B stages of Loomis et al. *Stage 2* was characterized by spindle activity on a low voltage background and included segments with K complexes. *Stage 3* was scored when two or more waves over 100 μV amplitude and less than 2 Hz frequency occurred in 10 seconds, while stage 4 was scored when such slow waves constituted at least half the record. They noted that stage 1 sleep was associated with rapid eye movements except at the start of the night and clearly differentiated the dreams of REM sleep from the hypnagogic imagery of sleep onset stage 1 sleep.

In 1968, a group of sleep researchers under the chairmanship of Allan Rechtschaffen and Anthony Kales met to develop the first consensus based guidelines for staging and scoring sleep in normal human subjects. The subsequently published manual summarizing these rules has formed the basis for sleep staging ever since.[1] The committee recommended using a minimum of one channel of central EEG (either C_3 or C_4 to the opposite ear or mastoid), chin EMG, and two channels of EOG (electrodes placed below and lateral to one eye and above and lateral to the other eye, both referenced to the same ear or mastoid. They recommended an epoch-by-epoch approach to scoring, using epochs of 20 or 30 seconds. Table 2 summarizes the Rechtschaffen and Kales (R & K) scoring criteria. It should be noted that a somewhat similar scoring system had been developed at the University of Florida by Williams, Karacan, and Hursch in the early 1960s, based on one-minute epochs with slightly different frequency and amplitude criteria and a different EEG montage. Extensive normative data of human sleep was published by this group in 1974.[18]

Limitations of the R & K system have been noted.[19] These include the reliance on an epoch-based system that is often dependent for scoring on the preceding and following epochs and does not assign a stage shift at the actual point it occurs. The method is often difficult to apply to abnormal, heavily fragmented sleep. The single EEG derivation has been criticized, as well as the choice of EOG derivations. Some of the rules, especially with respect to stage boundaries, are hard to follow. Difficulties arise in subjects who do not generate alpha rhythm. The amplitude criteria for the scoring of slow wave sleep have been criticized. A committee of the Japanese Society of Sleep Research has published proposed supplements and amendments to the system[20,21] and modifications in montages have been suggested.[22,23] An additional stage of transitional sleep (T-sleep) has been proposed for sleep with multiple shifts between NREM sleep and wakefulness due to sleep disordered breathing.[24] A technique for scoring the cyclic alternating pattern as a measure of sleep microarchitecture has been published.[25] However, since 1968 no systematic attempt has been made to develop a new scoring system.

4.0 INTER-RATER AND INTRA-RATER RELIABILITY IN SCORING SLEEP

The task force reviewed 7 studies examining inter-rater reliability of different human scorers.[26-32] Two studies were graded at evidence Level 1,[26,32] two at evidence Level 2,[30,31] two at evidence Level 3,[27,28] and one at evidence Level 4.[29] Three examined the PSGs of normal subjects,[27-29] two examined the PSGs of patients with various sleep disorders,[26,31] and two included PSGs of both normal subjects and patients.[27,30] Methodologies differed; the num-

ber of PSGs scored varied between 1 and 98 (median =19.5), and the number of scorers per record varied between 2 and 27 (median = 2.5). Four studies used standard R & K montages,[26,27,31,32] two used in addition the O_1-M_2 derivation,[28,30] two used modified EOG derivations,[26,28] and one (from before the publication of the R & K manual) used F_4-A_2 and P_4-A_1 without a submental EMG.[29]

Three studies (two of patients with sleep disorders and one of patients and normal subjects), all using a standard R & K montage, utilized the kappa statistic to measure overall inter-rater reliability for all epochs.[26,31,32] (Kappa values of 0.21-0.4 indicate fair agreement, 0.41-0.6 moderate agreement, 0.61-0.8 substantial agreement, and 0.81-0.99 almost perfect agreement.[33]) The resulting values of 0.68, 0.79, and 0.82 indicated at least substantial agreement. In 4 studies (two with occipital in addition to central electrodes), inter-rater agreement between different stages was calculated.[27,28,30,32] These showed best agreement for REM (78%-94%) and for stage 2 sleep (79%-90%). Agreement for wake was 68%-89% (three studies), for stage 1 was 23-74% (three studies), for stage 3 was 44%-60% (two studies), for stage 4 was 45%-80% (two studies), and for slow wave sleep (SWS-combined stages 3 and 4) was 69% (two studies). Kappa statistics for individual stages were calculated in one study.[26] These fell in the perfect range for REM, the substantial range for wake, stage 2, and SWS (combined stages 3 and 4), and in the fair range for stage 1.

One Level 4 study examined intra-rater reliability with each of 3 scorers rescoring 20 studies after a median of 6.5 months.[32] Agreement for wake was 89%-93%, for stage 1 was 18%-42%, for stage 2 was 76%-85%, for SWS was 55%-75% and for REM sleep was 72%-88%. Overall reliability measured with the kappa statistic was 0.79-0.87.

The task force concluded that inter-rater and intra-rater reliability were substantial for staging of records as a whole with the use of the R & K montages. This suggested that the R & K method could be used as a basis for a revised scoring system. Greatest inter-rater accuracy was achieved for REM sleep followed by stage 2 sleep. Lowest reliability was found for stage 1 sleep, while reliability for wake and SWS was moderate. Similar results were noted in the one study examining intra-rater reliability. It was concluded that scoring rules for stage 1 and SWS sleep needed reassessment.

5.0 VALIDITY AND RATIONALE FOR SPECIFICATIONS AND STAGING

5.1 EEG derivations

The R & K manual specified a single optimal central EEG derivation for scoring sleep (C_4-A_1 or C_3-A_2). They noted that "regional differences are not critical for the scoring of sleep stages except insofar as certain critical types of activity, i.e., alpha, vertex sharp waves, sleep spindles, K complexes, and delta waves are adequately registered. Furthermore, the wide use of eight channel electroencephalographs and a desire to maximize data acquisition by running two subjects on each machine often limits the channels of information from each subject to four."[1] In a different era, when a minimum of 16 channels of data are easily recordable on each subject, the question arises whether a single EEG derivation "adequately registers" the characteristic wave forms of sleep. It should be noted that the alternative EEG scoring system in use in the 1960s (Williams, Karacan, Hursch) used a minimum of 3 EEG derivations (F_1-F_z, P_1-T_5, O_3-O_ZP_Z).[18]

The task force reviewed 4 studies which examined the localization of scalp EEG activity during NREM sleep (evidence Level 4).[34-37] All four involved normal subjects with mean ages of 23-26 years. Three studies utilized various digital analyses,[34,35,37] while the fourth relied on visual scoring alone.[36] These studies suggested that K complexes were maximally represented with frontal electrodes (2 studies), sleep spindles maximally with central electrodes (2 studies), and delta activity maximally with frontal electrodes (2 studies). In the one study with visual scoring,[36] K complexes were required to have a minimum amplitude of 75 µV, a criterion not stated in the R and K rules. Defined in this way, 77% of K complexes were identified using the Fp1 electrode, 71% with the F_3 electrode, and 37% with the C_3 electrode. Thus, fewer than half the K complexes would have been identified with the C_3 electrode alone. One study reviewed the effects of aging on the generation of sleep spindles and K complexes (evidence Level 2),[38] comparing the EEG of 14 young subjects (mean age 21 years) and 20 older subjects (mean age 75 years). Spindle density fell with age by 69% in men and 79% in women, and K complex density fell by 30% in the older men and 12% in the older women.

Thus, evidence from these studies suggests that sleep spindle activity is optimally recorded with central electrodes, while K complexes and delta activity are optimally recorded with frontal electrodes. The predominant localization of alpha rhythm over the posterior head regions, and especially the occipital cortex, has been unchallenged since the days of the early EEG pioneers.[7] Failure to record frontal activity may result in reduced identification of K complexes and thus inaccurate scoring of stage 2 sleep, especially in older subjects, and the absence of an occipital derivation may hamper the determination of sleep onset. As a result, the task force determined through the consensus process that a minimum of three EEG derivations will be required, sampling activity from the frontal, central, and occipital regions.

The task force considered whether EEG should be recorded using referential or bipolar derivations. A study of 10 patients with epoch-to-epoch comparisons of scoring using the traditional R & K montage versus a montage consisting of Fpz-C_Z and P_Z-O_Z was reviewed. Agreement between the two montages using kappa statistics was fair to good or excellent, with a tendency to score more slow wave sleep with the bipolar montage (evidence Level 4).[22] The advantages of using a reference electrode on the opposite ear or mastoid included the familiarity of polysomnographers with these derivations and the ease with which the origin of a potential can be recognized by the maximal amplitude of the wave. However, the ear or mastoid is an active reference which, in addition to recording EEG activity, also may record EMG and EKG artifact. In contrast, the use of midline bipolar derivations allows for localization of potential origin by the site of phase reversal and reduced EMG and EKG artifact. While these derivations are in use in some laboratories, they are less familiar to most polysomnographers.

The consensus voting process showed that both approaches were considered acceptable to the task force, although the use of referential derivations was slightly favored. The task force proposed recommended derivations of F_4-M_1, C_4-M_1, and O_2-M_1 (right sided active electrodes and a reference electrode over the left mastoid, rather than the ear). (The choice of the mastoid was dictated by technical considerations: electrodes attached to the ear are more likely to become detached during the night.) However, equally acceptable alternative derivations were F_Z-C_Z, C_Z-O_Z, and C_4-M_1. It was also recommended that appropriate backup elec-

trodes for each standard electrode be applied in case of electrode malfunction during the sleep study.

5.2 EOG derivations

The R & K manual recommended at least two electro-oculogram (EOG) derivations to record eye movements during sleep. The first active electrode is placed one centimeter above and slightly lateral to the outer canthus of one eye and referred to the ipsilateral ear or mastoid, while the second active electrode is placed one centimeter below and slightly lateral to the outer canthus of the other eye and referenced to the contralateral ear or mastoid. Thus the reference electrode is the same for both derivations.[1] The term "slightly lateral" is often interpreted to mean one centimeter lateral. This arrangement results in almost all eye movements producing out-of-phase deflections on the two derivations, while electrode artifacts produce in-phase or single channel deflections. On the other hand, these derivations do not allow the direction of eye movements to be determined and may miss some low amplitude oblique eye movements. The R & K manual notes that a supranasion reference electrode would allow direction of eye movements to be determined.

The task force reviewed a study of 8 subjects (aged 14-50 years) with EOG recorded with two derivations, one lateral to each eye and the other above and below the right eye (evidence Level 4).[39] Five percent to 15% of REM sleep eye movements were horizontal, 25%-35% vertical, and 55%-65% oblique. A study of 15 subjects (nine 18-36 years, six 64-87 years) comparing linked bilateral mastoid with midline forehead EOG reference electrodes was reviewed.[38] In the younger, but not the older group, significantly fewer REMs were recorded with the mastoid electrodes (11%) and this resulted in significantly shorter REM periods being scored. However, it should be noted that the definition of REM sleep was based on pre-R & K criteria (evidence Level 4). In another study,[41] (evidence Level 4) seven normal male subjects were asked to voluntarily move their eyes in different angular directions, with eye movements recorded by an array of active electrodes referenced to the left mastoid. The study suggested that highest amplitude signals were obtained with electrodes one centimeter above and below the eyes at the level of the outer canthi rather than 1 cm lateral to them.

The task force considered the standard R & K derivations as well as derivations consisting of active electrodes placed one centimeter below and one centimeter lateral to the outer canthi of both eyes referenced to Fpz.[42] The task force found both approaches acceptable using the consensus voting process, although the use of the R & K derivations was slightly favored. The group recommended that the standard EOG derivations be E1-M2 and E2-M2, with the E1 electrode placed one centimeter below the left outer canthus and the E2 electrode one centimeter above the right outer canthus. (Based on available evidence discussed above,[41] the R & K recommendation placing the electrodes "slightly lateral" to the outer canthi was modified.) However, equally acceptable alternative derivations were E1-Fpz and E2-Fpz with the E1 electrode placed one centimeter below and one centimeter lateral to the left outer canthus and the E2 electrode one centimeter below and one centimeter lateral to the right outer canthus. The latter derivations are especially helpful if it is important to record the direction of eye movements, as vertical eye movements will show in-phase deflections and horizontal eye movements out-of-phase deflections.

5.3 Chin EMG derivation

Given the uniformly accepted practice of using EMG derivations specified in the R & K manual, the task force recommended continuation of the standard bipolar single chin surface EMG derivation. Chin EMG should be recorded from electrodes placed above and below the chin with a backup electrode placed below the chin close to the primary electrode. The group has proposed specific landmarks for the placement of these electrodes, described in detail in the scoring manual.

5.4. Epoch-based versus continuous scoring

The task force considered at length whether the new sleep stage scoring system should continue to be based on discrete epochs or should use a system of visual adaptive scoring.[16] This method defines the start and end of each stage and does not rely on epochs. Advantages of such a method might include a closer representation of the continuous nature of sleep, greater speed of scoring long periods of unchanging stage, and a more accurate measure of the microstructure of fragmented sleep. However, using such a system would make the visual scoring of highly fragmented sleep extremely laborious and time consuming, and the arousal index could give a more easily derivable measure of sleep continuity. Also, definitions of most sleep stages, such as slow wave sleep, depend on the use of discrete epochs. The task force was unable to find literature that directly addressed this issue. Using the consensus method, a recommendation was made to retain the traditional epoch based scoring method. The group also voted to recommend that an epoch length of 30 seconds continue to be used for stage scoring, finding no compelling evidence to change it.

5.5 Terminology

As will be discussed below, the task force elected to retain the division into wakefulness, NREM, and REM sleep, with 3 stages of NREM sleep. The group considered many names for the new stages and voted to recommend the following terminology:

Stage W (Wakefulness)
Stage N1 (NREM 1 sleep)
Stage N2 (NREM 2 sleep)
Stage N3 (NREM 3 sleep)
Stage R (REM sleep)

5.6 Defining the wake-sleep boundary

Attempts to define the wake-sleep transition date back to 1935 when Loomis et al described fragmentation of alpha rhythm followed by attenuation of alpha around sleep onset.[8,9] In 1938 Davis et al described detailed changes in alpha frequency and amplitude prior to loss of alpha rhythm and also noted that repetitive periods of interruption of alpha rhythm by theta activity preceded the complete disappearance of alpha rhythm.[41] Slow eye movements around sleep onset were first identified by Aserinsky and Kleitman in 1953 and 1955[16] and were noted to be often the first indicator of drowsiness.[44-46] Based on these and other early observations, several authors suggested criteria for the first stage of sleep[9,13,14,17] culminating in the R & K definition of stage 1 NREM sleep in 1967.[14] Subsequently, a classification of drowsiness based on nine stages from wakefulness to stage 2 NREM sleep has been proposed.[47]

The task force reviewed a large study of 55 normal young subjects which aimed at systematically identifying the EEG and EOG features of wake-sleep transition (evidence Level 4).[48] Ten percent of subjects had no alpha rhythm while awake with eyes closed, and another 10% had only brief runs of low amplitude alpha mixed with beta rhythms. In subjects with well-defined alpha rhythm, fallout of alpha was preceded by either decrease or increase in alpha amplitude, centrofrontal spread of alpha, and theta and delta activity intermixed with alpha. The earliest sign of drowsiness was disappearance of mini-blinks (best detected by eyelid movement transducers, but also visible with EOG), while slow eye movements usually appeared before loss of alpha rhythm.

Changes in electrophysiology present only one view of the wake-sleep transition, with psychophysiological studies providing a different perspective. An early study showed that normal healthy adults provided increasing numbers of wrong answers to questions or were unable to provide any answer as drowsiness deepened. During periods of alpha slowing or discontinuous alpha, subjects answered correctly or even reported having heard the question less than a fourth of the time.[9] We reviewed a study in which tones 5-10 dB above background noise were administered multiple times to 11 healthy subjects during the course of a night's sleep (evidence Level 3).[49] Subjects were requested to respond to the tones by pressing a button taped to the palm. Using 40-second epochs scored according to Rechtschaffen and Kales criteria, the probability of responding was 0.88 in wake, 0.39 in stage 1 sleep, and 0.03 in stage 2 sleep. If sleep was scored in 1-5 second epochs at the time the stimulus was applied, then the probability of responding was 0.94 in wake, 0.24 in stage 1 sleep, and 0.02 in stage 2 sleep.

Other physiological changes occur at wake-sleep transition, including alterations in respiration, cardiac function, cerebral blood flow and evoked potentials. Minute ventilation falls abruptly within 1-2 breaths of a transition from alpha to theta frequencies, with little further change at the transition between stage 1 and stage 2 sleep.[50,51] Similarly, phasic EMG activity of the diaphragm, intercostals, genioglossus, and tensor palatine falls at the start of alpha-theta transitions.[52,53] Upper airway resistance also increases within one breath after transition to theta[49,51] but continues to rise as NREM sleep progresses.[54] Heart rate decreases within five beats, and ECG T-wave amplitude increases within one beat of transition from alpha to theta activity.[55] Near infrared spectroscopy of the frontal lobe reveals a fall in oxyhemoglobin level, a surrogate marker of cerebral blood flow, within about 5 seconds of alpha-theta transition.[56] A positron emission tomography study using ^{15}O-labelled water revealed a relative increase in cerebral blood flow in the occipital lobes and a decrease in the bilateral cerebellum, bilateral posterior parietal cortices, right premotor cortex, and left thalamus during stage 1 sleep compared to wakefulness.[57] Event related potentials change at sleep onset. The P1 and P2 wave amplitudes increase in stage 1 sleep compared to relaxed wakefulness and the N1 wave amplitude attenuates.[58] The P300 wave attenuates frontally in stage 1 sleep and the maximum voltage is recorded over the lateral parietal and occipital regions, compared with the parietal vertex during wakefulness. [59]

Aspects of the available data summarized above suggests that the process of transition from wakefulness to sleep is a continuum and may be best visualized as a sleep onset period lasting from the reduction in the rate of mini-blinks to the development of sustained theta activity with K complexes and sleep spindles.[49] In support of this concept are the EEG and EOG changes that precede the loss of alpha activity and the variable points on the spectrum between wakefulness and stage 2 sleep at which different subjects fail to respond to auditory tones. On the other hand, a number of lines of evidence provide validity to the concept that sleep onset can be practically defined at the point at which EEG theta activity predominates over alpha. There is a 49%-70% drop in responses to auditory stimuli in stage 1 sleep compared to wakefulness, depending on the length of epochs used. Significant changes in respiration, heart rate, cerebral blood flow, and event related evoked potentials occur in stage 1 sleep compared with wakefulness, often commencing within seconds of alpha-theta transition.

The task force considered whether to establish a new stage of drowsiness, which would overlap parts of the currently defined stages of wakefulness and stage 1 sleep. However, through the consensus voting process, the task force chose not to follow this approach, instead electing to retain the traditional distinction between wakefulness and stage 1 sleep. Proposed criteria for stages W and N1 in subjects who are good alpha rhythm generators (80%-90% of the population) are largely unchanged, and sleep onset is defined as the start of the first epoch of sleep other than stage W.

The issue of definition of stage N1 and sleep onset in subjects who generate little or no alpha rhythm (10%-20% of the population) was carefully considered. This group most probably accounts for the low inter- and intra-rater reliability for scoring of stage 1 sleep. Through the consensus voting process, the task force chose to concentrate on the development of slow eye movements in the EOG as the best measure of early sleep in the absence of any visually discernable alterations in EEG. While the reduction of mini-blink rate is the first change in eye movements in drowsiness, this is best measured by eyelid movement transducers, rather than the EOG. Thus the proposed criteria for the start of stage N1 in the absence of adequate alpha rhythm are the observation of the earliest of any of the following phenomena: 4-7 Hz activity with slowing of background frequencies by ≥1 Hz from those of stage W; vertex sharp waves; and slow eye movements. Based on limited data[60,61] and clinical experience, the group defined slow eye movements as conjugate, reasonably regular, sinusoidal eye movements with an initial deflection usually lasting >500 msecs.

5.7 Defining sleep spindles and K complexes

After the transition phase from wakefulness, NREM sleep is characterized by low amplitude mixed frequency background with two morphologically distinct waveforms superimposed: K complexes and sleep spindles. The presence of these wave forms defines stage N2 sleep. The task force examined their morphology, amplitude, frequency, and distribution, as well as considering the rules governing the start and end of the stage.

The first description of K complexes is attributed to Loomis et al in 1937.[9] K complexes can be divided into spontaneous (occur without a particular identifiable cause) or evoked (related to a known sensory stimulus) although both have the same appearance.[43,62,63] The K complex consists of a bi- or triphasic sharp wave complex[63] with average duration 0.63 seconds (range 0.5-1.0 seconds).[64] The initial component is surface negative[65] with the 2nd wave reversed in phase compared to the first.[63] Peak-to-peak voltage, using automated detection software, has been calculated as 100-400 microvolts.[64] K complexes tend to recur at a frequency of 1.0 to 1.7 per minute (range 0.7 to 3.76).[66] The K complex may be followed by a sleep spindle[63] or a run of alpha rhythm.[67]

The K complex induced by sensory stimuli is actually a long latency evoked potential recorded after 550 msec (N550).[38,65,68] The K complex is maximal in amplitude over the frontal regions and failure to record with a frontal electrode results in many K complexes (at least those with amplitude <75 μV) being overlooked (see earlier under EEG activity).[35,36]

The task force voted to define a K complex as a well delineated negative sharp wave immediately followed by a positive component standing out from the background EEG with total duration ≥0.5 seconds. It is usually maximal in amplitude over the frontal regions. After careful consideration of available data and personal experience, the group voted not to apply an amplitude criterion, as the morphology and duration are sufficiently distinctive that additional characterization would unnecessarily complicate visual analysis. The task force considered the special situation in which sleep is highly fragmented by arousals associated with K complexes, often induced by upper airway obstruction. The group voted to require either sleep spindles or spontaneous K complexes to define stage N2 sleep; if only K complexes associated with arousals are present, then the stage will be scored as stage N1.

Sleep spindles were first described by Berger[69] in 1933 but Loomis et al[8] were first to use the term for waves of 20-40 μV amplitude and 14-15 Hz frequency. Sleep spindles are a comprised of a group of rhythmic waves which progressively increase and then gradually decrease in amplitude. A number of studies have examined their frequency range, using either visual or computerized analysis. While the R & K manual[1] specified their frequency as between 12 to 14 Hz, most studies have suggested a wider range with lower limit of about 11 Hz[19,70,71] and an upper limit of about 15-16 Hz.[19,70,72,73] Median spindle frequencies have been reported to form a U-shape within a single episode of NREM sleep, with faster frequencies at the beginning and again at the end of the episode, and slower frequencies in between.[16] Spindle duration in young adults ranges from 0.5 to 1.2 seconds.[72,73] Spindle frequency can vary about 2 Hz and spindle duration from 0.5 to 0.8 seconds between different spindle trains.[74] The mean density of spindles in young adults has been reported as 2.3-3.6 per minute, but with considerable variation between subjects.[36,72,73] Sleep spindles are most frequently detected over the central head regions. [36]

Following review of the literature, the task force voted to define a sleep spindle as a train of distinct waves with frequency 11-16 Hz (most commonly 12-14 Hz) with a duration ≥0.5 seconds, usually maximal in amplitude over the central regions.

The task force also considered rules regarding the termination of a period of stage N2 sleep. The "3-minute rule" of the R & K manual states that a maximum of 3 minutes between K complexes and/or sleep spindles can be scored as stage 2 sleep, provided there are no movement arousals or pronounced increases in muscle tone.[1] The Visual Task Force could find no evidence-based justification for the 3-minute time interval and voted to discard it. Instead it was decided that stage N2 sleep should continue to be scored in the absence of K complexes or spindles until a transition to stage W, N3, or R, or the occurrence of a major body movement followed by slow eye movements or an arousal.

5.8 Defining slow wave sleep

The task force considered several aspects of SWS, specifically evidence for the validity of scoring SWS as an independent stage, frequency criteria, amplitude criteria including the effects of aging, and whether the current subdivision of SWS into stages 3 and 4 sleep should be retained.

A study of 10 normal young men revealed increased frequency of K complexes in stage 2 sleep preceding SWS compared with stage 2 sleep preceding REM sleep, suggesting that K complexes are forerunners of delta waves in SWS (evidence Level 4).[75] While this might suggest that SWS is not fundamentally different from stage 2 sleep in terms of scalp EEG, intracortical recording of electrical neuronal activity suggests otherwise. Slow oscillations (<1 Hz), recorded during NREM sleep and largely generated by cortical neurons, underlie K complexes. In contrast, delta oscillations (1-4 Hz), generated by both thalamocortical pathways and cortical networks, are responsible for the delta frequencies of slow wave sleep.[76]

Other physiologic phenomena provide further validation for classifying SWS as a separate stage of NREM sleep. The onset of SWS is tightly linked to the release of growth hormone,[77] with maximal secretion occurring within minutes of the onset of the change in stage.[78] Drugs that stimulate SWS increase growth hormone secretion.[79,80] Passive body heating and exercise[81,82] in the evening increase slow wave sleep but have little effect on stage 2 NREM sleep. Similarly, warming the preoptic/anterior hypothalamic regions causes increased delta activity but no change in theta or sigma frequency bands during NREM sleep.[83] Total sleep deprivation results in rebound slow wave sleep during the first recovery night, but no increase in stage 2 sleep.[84,85] In summary, there are sufficient distinctive physiologic features of SWS to warrant its retention as a separate stage of sleep.

In electroencephalography, the delta frequency band has traditionally been defined as frequencies <4 Hz. However, both the Dement and Kleitman and the R & K scoring systems used a cut off of 2 Hz to specify slow wave sleep. No data could be found to provide any validity for different definitions of frequency. Based on consensus experience that most slow activity scored visually appears to be <2 Hz in frequency and the established use of the 2 Hz cut off in almost every study of SWS over more than 40 years, the task force voted to retain defining SWS in terms of frequencies <2 Hz. However, the group endorsed the R & K manual recommendation that the use of the term "delta sleep" should be strongly discouraged to avoid confusion. The task force also voted to set a lower limit of 0.5 Hz, as lower frequency activity is usually related to artifact. Although slow neuronal oscillations of <1 Hz may have important physiological implications, they cannot be detected visually in scalp EEG.

Stages 3 and 4 sleep are the only stages in which the R & K manual specifies amplitude criteria. This was felt to be important, as there were no distinctive morphological features of slow waves in contrast to K complexes and sleep spindles. The amplitude of EEG waves measured in microvolts is dependent on filter settings and derivations used. The R & K manual defines the low frequency filter setting as a time constant not shorter than 0.3 seconds[1] and specifies the derivation for measuring slow wave amplitude as C_4-A_1 or C_3-A_2. The criterion of peak-to peak slow wave amplitude >75 μV was reached empirically, and the authors state that the selection "should not deter investigators from using measures of slow wave activity other than the one suggested."[1]

We reviewed a study of 25 subjects, divided into 5 groups with ages from 3-79 years (evidence Level 4).[86] Analysis of all delta activity >5 μV amplitude (F_1-F_7 derivation) showed that abundance did not fall with age, but amplitude dropped significantly

after about age 40 years. However, similar age-related drops in amplitude or spectral power have been noted in the theta, alpha, and spindle frequency bands,[87,89] suggesting that this is a nonspecific phenomenon. We reviewed a study of 40 subjects, mean age 55 years (range 50-60) in which stages 3 and 4 sleep were scored by 2 raters without amplitude criteria. using one-minute epochs (evidence Level 4).[90] Frontal, parietal, and temporal EEG derivations were available in addition to central derivations. A mean of 72-83 minutes SWS was scored. The inter-rater correlation was 0.72 for stage 3, 0.82 for stage 4, and 0.8 for slow wave sleep as a whole. We were unable to find any studies comparing the validity of scoring SWS with different amplitude criteria at different ages. In contrast, there are extensive data showing that the abundance of SWS defined in terms of >75μV amplitude significantly changes in both the young and the elderly in response to sleep deprivation experiments,[91-93] induced sleep fragmentation,[93,94] and forced desynchronization protocols.[97,98]

In summary, SWS amplitude falls with age, but in parallel to EEG waves of other frequencies. SWS can be reliably measured without amplitude criteria, and thus presumably also with amplitude criteria lower than 75 μV. However, there is no evidence to suggest that there would be any benefit in using a criterion different from 75 μV, and there is considerable evidence indicating that changes in SWS defined in the conventional way can be detected at all ages in response to a variety of experimental interventions. Thus both the Visual Scoring Task Force and the Geriatric Task Force voted to retain the 75 μV criterion for all ages. Data discussed earlier in the EEG derivation section indicates that slow wave amplitude is higher when recorded by frontal compared to central derivations. As a result, the task force voted to recommend the use of a frontal rather than a central derivation to measure slow wave amplitude, understanding that most of the studies discussed above were performed using a central derivation.

No evidence could be found to indicate validity or biological significance in the subdivision of SWS into stages 3 and 4 based on the percentage slow waves in each epoch. The task force felt that this was an arbitrary distinction which many laboratories have already discarded and voted not to subdivide SWS. The group considered various cutoffs for the minimum percentage of slow waves necessary in each epoch, but found no reasons to change from the current definition of stage 3 sleep. It was therefore recommended that stage N3 sleep be scored when 20% or more of an epoch consists of waves of 0.5-2 Hz frequencies with peak-to-peak amplitude of >75 μV in the frontal derivation.

5.9 Defining REM sleep

The task force felt that there was overwhelming, well-accepted evidence in support of the validity of REM sleep as a state separate from NREM sleep. As discussed earlier, inter- and intra-rater reliability for scoring according to R & K rules is highest for REM sleep compared to other stages. Thus the group concentrated predominantly on refining and simplifying REM sleep scoring rules, especially with regard to the start and end of periods of REM sleep, an area of considerable complexity in the R & K manual. We also examined the nomenclature of REM sleep phenomena and their definitions.

We reviewed a few studies that examined less commonly described aspects of REM sleep. One study of 12 epileptic patients identified mu-like 7-10 Hz central EEG activity in about 50% of patients (evidence Level 4).[99] A study of ventilation during REM in 6 patients using a pneumotachometer showed reduction of tidal volume and increase in respiratory frequency during phasic but not tonic REM (evidence Level 4).[100] A study of the cortical N300 auditory evoked potential in 12 subjects showed amplitude attenuation 0.5-2.5 seconds before loss of submental muscle tone (evidence Level 4).[62]

The Visual Task Force resolved that stage R sleep should be scored in the presence of the 3 phenomena of low amplitude mixed frequency EEG background, rapid eye movements, and low chin EMG tone. Some epochs of stage R sleep do not contain rapid eye movements, and rules for their scoring are discussed later. The presence of sawtooth waves or transient muscle activity ("phasic twitches") is strongly supportive evidence, especially if one or more of the basic features are equivocal. Consideration was given to including variations in respiratory amplitude and frequency, but it was decided that these changes were often too hard to separate from pathologic breathing to be helpful as diagnostic criteria. Although there are physiologic differences between periods of tonic and phasic REM sleep, the group felt that defining these as separate substages would unnecessarily complicate the scoring process.

Neuro-ophthalmologists have given little or no attention to the rapid eye movements of REM sleep. Eye movements during wakefulness are generally divided into two categories, saccades and pursuit movements, each of which has somewhat different neuroanatomic and neurophysiologic mechanisms. Pursuit movements have a maximum velocity of about 60 degrees/sec while saccades are much faster, with the usual maximum for humans considered 300 degrees/sec. The majority of waking saccades are relatively small in magnitude, generally 15 degrees or less.[101] Sleep neurophysiologists have taken the known data on eye movements in awake primates as their starting point and have tried to fit the results of their observations into the conceptual framework of ocular motor neurophysiology. Some investigators have suggested that rapid eye movements of REM sleep have similar characteristics to waking saccades[102-104] but the most definitive study[16] found major differences. REM sleep rapid eye movements are slower than waking saccades of the same size and have a longer duration. Peak velocities of small saccades (5.5 degrees) were 201.4 ±4.6 degrees/sec compared to 149.8±6.6 degrees/sec for rapid eye movements. Peak velocities of large saccades (11 degrees) were 328.4±6.1 degrees/sec compared to 180.2±8.1 degrees/sec for rapid eye movements. The mean duration of the small rapid eye movements was 74.3±2.5 msec (compared to 52.8±2 msec for waking saccades) and the mean duration of large rapid eye movements 121.6±5.6. msec (compared to 66.3±2 msec for waking saccades). It is not possible using AC recordings to determine accurately the velocity or size of eye movements. Thus on the basis of the above data and clinical experience, the task force determined by consensus that a rapid eye movements during sleep should be defined as conjugate, irregular, sharply peaked eye movements with an initial deflection usually lasting <500 msec.

Low amplitude, mixed frequency activity in stage R sleep is similar to that seen in stage N1 sleep, but in some subjects alpha frequencies are more abundant and widely distributed, often at a slightly slower frequency of alpha that is observed during wakefulness. The term "low chin tone" was preferred to chin atonia and the R & K definition of EMG activity no higher than in any sleep stage and usually at its lowest level was retained. Based on

descriptions in the literature[105,106] and the experience of task force members, sawtooth waves were defined by consensus as trains of sharply contoured or triangular, often serrated, 2-6 Hz waves maximal in amplitude over the central head regions and often, but not always, preceding a burst of rapid eye movements. The group recommended that the term "phasic muscle twitches" be replaced by "transient muscle activity." Dictionary definitions of "phase" do not include short-lived transient activity and the term is used in physics for "a particular appearance or state in a regularly recurring cycle of changes."[107] Transient muscle activity was defined by consensus as short irregular bursts of EMG activity usually with duration <0.25 seconds superimposed on a background of low EMG activity. It was noted that this activity may be seen in the chin or anterior tibial derivations, as well as in EEG or EOG derivations, the latter indicating activity of cranial nerve innervated muscles. It is maximal in association with rapid eye movements.

The task force addressed scoring of the onset and termination of periods of stage R sleep by consensus voting, with the aim of simplifying the current rules and giving clear guidelines for most circumstances. The detailed rules will appear in the manual, but in summary, stage R sleep commences when chin EMG tone falls, unless K complexes or spindles persist, in which case stage N2 persists until rapid eye movements develop. If chin EMG tone is low in stage N2 as well as REM sleep, the transition to Stage R occurs after the last K complex or spindle. If K complexes or sleep spindles are interspersed among what are otherwise epochs of unequivocal REM sleep (as may especially occur in the first REM period of the night), then stage R should be scored if rapid eye movements are present and stage N2 if eye movements are absent. If epochs of REM sleep are followed by epochs with low amplitude mixed frequency EEG and persistently low chin EMG tone, but no rapid eye movements, K complexes, or spindles, then they should continue to be scored as stage R until there is a transition to stage W, N2, or N3, an arousal or major body movement followed by slow eye movements, or an increase in chin EMG tone signifying a change to stage N1.

5.10 Major body movements

The R & K manual defines movement time as epochs during which the polygraph record is obscured by movements of the subject for more than 50% of each epoch.[1] The Visual Task Force decided by consensus voting to eliminate this stage, as a movement of this duration and magnitude commonly results in a transition to wakefulness. Instead such an epoch with a major body movement will be scored as wake if any part of the epoch shows alpha rhythm, or if a wake epoch either precedes or follows the epoch in question. Otherwise the epoch is scored as the same stage as the epoch that follows it.

6.0 UNRESOLVED ISSUES AND FUTURE RESEARCH

The introduction of a new sleep scoring system with even modest changes from a long accepted method requires considerable adaptation by physicians, scientists, technologists, and the manufacturers of polysomnography systems. The advantage of change needs to be self-evident, and users must easily understand the benefits of the new system. We believe that the proposals discussed in this paper have the advantages of simplicity, ease of application, faithfulness to current biological knowledge. and compatibility with the limited evidence available in the literature.

However, future studies of inter- and intra-rater reliability for visual scoring of the new system are essential. It will also be necessary to test the method not just in normal subjects of all ages, including children and older adults, but also in patients with abnormal sleep due to sleep disordered breathing and other pathologies. Modifications may be needed to deal with certain abnormalities, such as REM sleep behavior disorder in which increased muscle activity occurs during REM sleep. New methodologies for the detection of drowsiness will need to be assessed. These include DC eye movement monitors and techniques which assess eyelid movements.[108] Adaptation of the new rules to digital scoring algorithms will also need to be tested for reliability compared to visual scoring.

No visual based scoring system will ever be perfect, as all methods are limited by the physiology of the human eye and visual cortex, individual differences in scoring experience, and the ability to detect events viewed using a 30-second epoch. Nevertheless, we believe it is possible to develop a rigorous, biologically valid scoring system that can be applied meaningfully in clinical and research settings. The new scoring system is presented as a step forward along this path.

VISUAL TASK FORCE MEMBERS

The Visual Task Force members participating in evidence review and/or consensus decisions to derive visual scoring rules and specifications included: Michael H. Silber, Chair, Sonia Ancoli-Israel, Michael H. Bonnet, Sudhansu Chokroverty, Madeleine M. Grigg-Damberger, Max Hirshkowitz, Sheldon Kapen, Sharon A. Keenan, Meir H. Kryger, Thomas Penzel, Mark R. Pressman, and Conrad Iber.

REFERENCES

1. Rechtschaffen A, Kales A, eds. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Los Angeles, CA: BI/BR, 1968.
2. Iber C. Development of a new manual for characterizing sleep. Sleep 2004;27:1-3.
3. Fitch K, Bernstein S, Aguilar M, et al. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA: RAND Corporation, 2001.
4. Sackett DL. Rules of evidence and clinical recommendations. Can J Cardiol 1993;9:487-489.
5. Caton R. The electric currents of the brain. Br Med J 1875;1875:278.
6. Berger H. Über das Elektroenkelphalogramm des Mensen. Arch Pyschiatr Nervenkr 1929;87:527-70.
7. Adrian ED, Matthews BHC. Berger rhythms: potential changes from occipital lobes in man. Brain 1934;57:355.
8. Loomis AL, Harvey EN, Hobart GA. Further observations on the potential rhythms of the cerebral cortex during sleep. Science 1935;82:198-200.
9. Loomis AL, Harvey EN, Hobart GA. Cerebral states during sleep, as studied by human brain potentials. J Exp Psychol 1937;21:127–44.
10. Davis H, Davis PA, Loomis AL, Harvey EN, Hobart GA. Human brain potentials during onset of sleep. J Neurophysiol 1938;1:24-38
11. Blake H, Gerard R, Kleitman N. Factors including brain potentials during sleep. J Neurophysiol 1939;2:48-60.
12. Gibbs E, Lorimer F, Gibbs F, eds. Atlas of electroencephalography, vol. 1, methodology and controls. 2nd ed. Reading, MA: Addison-Wesley Publishing Company, 1950.
13. Simon EW, Emmons WH. EEG, consciousness, and sleep. Science 1956;124:1066-9.
14. Roth B. The clinical and theoretical importance of EEG rhythms corresponding to states of lowered vigilance. Electroencephalogr Clin Neurophysiol 1961;13:395-9.

15. Brown C. The stubborn scientist who unraveled a mystery of the night. Smithsonian 2003;October:92-9.
16. Aserinsky E, Kleitman N. Regularly occurring periods of eye motility and concomitant phenomena during sleep. Science 1953;118:273-4.
17. Dement W, Kleitman N. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. Clin Neurophysiol 1957;9:673-90.
18. Williams RL, Karacan I, Hursch, CJ, eds. Electroencephalography (EEG) of human sleep: clinical applications. New York: John Wiley & Sons, 1974.
19. Himanen S, Hasan J. Limitations of Rechtschaffen and Kales. Sleep Med Rev 2000;4:149-67.
20. Hirshkowitz M. Standing on the shoulders of giants: the Standardized Sleep Manual after 30 years. Sleep Med Rev 2000;4:169–79
21. Hori T, Sugita Y, Koga E, Shirakawa S. Proposed supplements and amendments to "A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects", the Rechtschaffen and Kales (1968) standard. Psychiatr Clin Neurosci 2001;55:305-10.
22. van Sweden B, Kemp B, Kamphuisen HA, Van der Velde EA. Alternative electrode placement in (automatic) sleep scoring (Fpz-Cz/Pz-Oz versus C4-A1). Sleep 1990;13:279-83.
23. Shepard JW, ed. Atlas of sleep medicine. Armonk, New York: Futura Publishing Company, 1991.
24. McGregor P, Thorpy MJ, Schmidt-Nowara WW, Ledereich PS, Snyder M. T-sleep: an improved method for scoring breathing-disordered sleep. Sleep 1992;15:359-63.
25. Parrino L, Smerieri A, Rossi M, Terzano MG. Relationship of slow and EEG components of CAP to ASDA arousals in normal sleep. Sleep 2001;24:881-5.
26. Danker-Hopke H, Kunz D, Gruber G, et al. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. J Sleep Res 2004;13:63-9.
27. Kim Y, Kurachi M, Horita M, Matsuura K, Kamikawa Y. Agreement of visual scoring of sleep stages among many laboratories in Japan: effect of a supplementary definition of slow wave on scoring of slow wave sleep. Jpn J Psychiatry Neurol 1993;47:91-7.
28. Martin W, Johnson LC, Viglione SS, et al. Pattern recognition of EEG–EOG as a technique for all-night sleep stage scoring. Electroencephalogr Clin Neurophysiol 1972;32:417–27.
29. Monroe L. Inter-rater reliability and the role of experience in scoring EEG sleep. Psychophysiology 1967;5:376-84.
30. Norman RG, Pal I, Stewart C, Walsleben JA, Rapoport DM. Interobserver agreement among sleep scorers from different centers in a large dataset. Sleep 2000;23:901-8.
31. Sangal RB, Semery JP, Belisle CL. Computerized scoring of abnormal human sleep: a validation. Clin Electroencephalogr 1997;28:64-7.
32. Whitney CW, Gottlieb DJ, Redline S, et al. Reliability of scoring respiratory disturbance indices and sleep staging. Sleep 1998;21:749-57.
33. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74
34. De Gennaro L, Ferrara M, Bertini M. Topographical distribution of spindles: variations between and within NREM sleep cycles. Sleep Res Online. 2000;3:155-60.
35. Happe S, Anderer P, Gruber G, Klosch G, Saletu B, Zeitlhofer J. Scalp topography of the spontaneous K-complex and of delta-waves in human sleep. Brain Topogr 2002;15:43-9
36. McCormick L, Nielsen T, Nicolas A, Ptito M, Montplaisir J. Topographical distribution of spindles and K-complexes in normal subjects. Sleep 1997;20:939-41
37. Werth E, Achermann P, Borbely AA. Fronto-occipital EEG power gradients in human sleep. J Sleep Res 1997;6:102-12
38. Crowley K, Trinder J, Kim Y, Carrington M, Colrain IM. The effects of normal aging on sleep spindle and K-complex production. Clin Neurophysiol 2002;113:1615-22
39. Jacobs L, Feldman M, Bender MB. Eye movements during sleep. I. The pattern in the normal human. Arch Neurol 1971;25:151-59
40. Feinberg I, Braun M, Koresko RL. Vertical eye-movement during REM sleep: effects of age and electrode placement. Psychophysiology 1969;5:556-61
41. Hakkinen V, Hirvonen K, Hasan J, et al. The effect of small differences in electrode position on EOG signals: application to vigilance studies. Electroencephalogr Clin Neurophysiol 1993;86:294-300
42. Hauri PJ, Harris CD, Silber MH. Physiologic assessment of sleep. In: Daube JR, ed. Clinical neurophysiology. New York: Oxford University Press; 2002. p. 493-512
43. Davis H, Davis P, Loomis A, Harvey E, Hobart G. Electrical reactions of human brain to auditory stimulation during sleep. J Neurophysiol 1939;2:500-14.
44. Kuhlo W, Lehmna D. Das Einschlaferleben und seine neurophysiologischen korrelate. Arch Psychiatr Nervenkr 1964;205:687-716.
45. Liberson WT, Liberson CW. EEG records, reaction times, eye movements, respiration, and mental content during drowsiness. Rec Adv Biol Psychiatry 1965;8:295-302.
46. Maulsby RL, Kellaway P, Graham M, eds. The normative electroencephalographic data reference library. Final Report, Contract NAS 9-1200, National Aeronautics and Space Administration. 1968.
47. Tanaka H, Hayashi M, Hori T. Statistical features of hypnagogic EEG measured by a new scoring system. Sleep 1996;19:731-8.
48. Santamaria R, Chiappa KH. The EEG of drowsiness in normal adults. J Clin Neurophysiol 1987;4:327-82.
49. Ogilvie RD, Wilkinson RT. Behavioral versus EEG-based monitoring of all night sleep/wake patterns. Sleep 1988;11:139-55.
50. Trinder J, Whitworth F, Kay A, Wilkin P. Respiratory instability during sleep onset. J Appl Physiol 1992;73:2462-9.
51. Fogel RB, White DP, Pierce RJ, et al. Control of upper airway muscle activity in younger versus older men during sleep onset. J Physiol 2003;553:533-44.
52. Mezzanotte WS, Tangel DJ, White DP. Influence of sleep onset on upper-airway muscle activity in apnea patients versus normal controls. Am J Respir Crit Care Med 1996;153:1880-7.
53. Worsnop C, Kay A, Pierce R, Kim Y, Trinder J. Activity of respiratory pump and upper airway muscles during sleep onset. J Appl Physiol 1998;85:908-20.
54. Kay A, Trinder J, Kim Y. Progressive changes in airway resistance during sleep. J Appl Physiol 1996;81:282-96.
55. Burgess HJ, Kleiman J, Trinder J. Cardiac activity during sleep onset. Psychophysiology 1999;36:298-306.
56. Spielman AJ, Zhang G, Yang C, et al. Intracerebral hemodynamics probed by near infrared spectroscopy in the transition between wakefulness and sleep. Brain Res 2000;866:313-25.
57. Kjaer TW, Law I, Wiltschiotz G, Paulson OB, Madsen PL. Regional cerebral blood flow during light sleep - a H15 O-PET study. J Sleep Res 2002;11:201-7.
58. de Lugt D, Loewy DH, Campbell KB. The effect of sleep onset on event related potentials with rapid rates of stimulus presentation. Electroencephalogr Clin Neurophysiol 1996;98:484-92.
59. Cote KA, De Lucht DR, Campbell KB. Changes in the scalp topography of event-related potentials and behavioral responses during the sleep onset period. Psychophysiology 2002;39:29-37.
60. De Gennaro L, Ferrara M, Ferlazzo F, Bertini M. Slow eye movements and EEG power spectra during wake-sleep transition. Clin Neurophysiol 2000;111:2107-15.
61. Porte HS. Slow horizontal eye movement at human sleep onset. J Sleep Res 2004;13:239-49.
62. Niiyama Y, Sekine A, Fushimi M, Hishikawa Y. Marked suppression of cortical auditory evoked response shortly before the onset of REM sleep. Neuroreport 1997;8:3303-8.
63. Roth M, Shaw J, Green J. The form voltage distribution and physiological significance of the K-complex. Electroencephalogr Clin Neurophysiol Suppl 1956;8:385-402.
64. Bremer G, Smith JR, Karacan I. Automatic detection of the K-complex in sleep electroencephalograms. IEEE Trans Biomed Eng 1970;17:314-23.

65. Bastien CH, Crowley KE, Colrain IM. Evoked potential components unique to non-REM sleep: relationship to evoked K-complexes and vertex sharp waves. Int J Psychophysiol 2002;46:257-74.
66. Halasz P. Hierarchy of micro-arousals and the microstructure of sleep. Neurophysiologie Clinique 1998;28:461-75.
67. Raynal D, et al. K-alpha events in hypersomniacs and normals. Sleep Res 1974;3:144.
68. Halasz P, Ujszaszi J. Chewing automatism in sleep connected with micro-arousals: an indicator of propensity to confusional awakening? Sleep 1988;86:235-9.
69. Berger, H. Über das Elektroencephalogram des Menschen. Sechste Mitteilung. Arch Psychiatr Nervenkr 1933;99:555-74
70. Zygierewicz J, Blinowska KJ, Durka PJ, Szelenberger W, Niemcewicz S, Androsiuk W. High resolution study of sleep spindles. Clin Neurophysiol 1999;110:2136-47.
71. Werth E, Achermann P, Dijk DJ, Borbely AA. Spindle frequency activity in the sleep EEG: individual differences and topographic distribution. Electroencephalogr Clin Neurophysiol 1997;103:535-42
72. Monasterio C, Silvia Vidal JD, Montse Ferrer M, et al. Effectiveness of continuous positive airway pressure in mild sleep apnea—hypopnea syndrome. Am J Respir Crit Care Med 2001;164:939-43.
73. Zeitlhofer J, Gruber G, Anderer P, Asenbaum S, Schimicek P, Saletu B. Topographic distribution of sleep spindles in young healthy subjects. J Sleep Res 1997;6:149-55.
74. Gondeck AR, Smith JR. Dynamics of human sleep sigma spindles. Electroencephalogr Clin Neurophysiol 1974;37:293-7.
75. De Gennaro L, Ferrara M, Bertini M. The spontaneous K-complex during stage 2 sleep: is it the 'forerunner' of delta waves? Neurosci Lett 2000;291:41-3.
76. Steriade M, Amzica F. Slow sleep oscillations, rhythmic K-complexes, and their paroxysmal developments. J Sleep Res 1998;7, suppl 1:30-5.
77. Van Cauter E, Plat L, Copinschi G. Interrelations between sleep and the somatotrophic axis. Sleep 1998;21:553-66.
78. Holl RW, Hartmann ML, Veldhuis JD, Taylor WM, Thorner MO. Thirty-second sampling of plasma growth hormone in man: correlation with sleep stages. J Clin Endocrinol Metab 1991;72:854-61.
79. Gronfier C, Luthringer R, Follenius M, et al. A quantitative evaluation of the relationships between growth hormone secretion and delta wave electroencephalographic activity during normal sleep and after enrichment in delta waves. Sleep 1996;19:817-24.
80. Van Cauter E, Plat L, Scharf M, et al. Simultaneous stimulation of slow-wave sleep and growth hormone secretion by gamma-hydroxybutyrate in normal young men. J Clin Invest 1997;100:745-53.
81. Bunnell DE, Agnew JA, Horvath SM, Jopson L, Wills M. Passive body heating and sleep: influence of proximity to sleep. Sleep 1988;11:210-19.
82. Horne JA, Staff LH. Exercise and sleep: body-heating effects. Sleep 1983;6:36-46.
83. McGinty D, Szymusiak R, Thomson D. Preoptic/anterior hypothalmic warming increases EEG delta frequency activity within non-rapid eye movement sleep. Brain Res 1994;667:273-77.
84. Aeschbach D, Cajochen C, Landolt H, Borbely AA. Homeostatic sleep regulation in habitual short sleepers and long sleepers. Am J Physiol 1996;270:R41-53.
85. Bonnet MH (2005) Acute sleep deprivation. In: Kryger MH, Roth T, Dement WC, eds. Principles and practice of sleep medicine. Philadelphia: Elsevier Saunders; 2005. p. 51-66
86. Smith JR, Karacan A, Yang M. Ontogeny of delta activity during human sleep. Electroencephalogr Clin Neurophysiol 1977;43:229-37.
87. Finneli LA, Landolt H-P, Buck A, et al, Functional neuroanatomy of human sleep states after zolpidem and placebo: A $H_2{}^{15}O$-PET study. J. Sleep Res 2000;9:161-73.
88. Stein MB, Millar TW, Larsen D, Kryger, MH. Irregular breathing during sleep in patients with panic disorder. Am J Psychiatry 1995;152:1168-73.
89. Tan X, Campbell IG, Feinberg I. Inter-night reliability and benchmark values for computer analyses of non-rapid eye movement (NREM) and REM EGG in normal young adult and elderly subjects. Clin Neurophysiol 2001;112:1540-52.
90. Webb WB, Dreblow LM. A modified method for scoring slow wave sleep of older subjects. Sleep 1982;5:195-9.
91. Brendel DH, Reynolds CF, Jennings JR. Sleep stage physiology, mood, and vigilance responses to total sleep deprivation in healthy 80-year-olds and 20-year-olds. Psychophysiology 1990;27:667-85.
92. Carskadon MA, Dement WC. Sleep loss in elderly volunteers. Sleep 1985;8:207-21.
93. Reynolds CF 3rd, Kupfer DJ, Hoch CC, Stack JA, Houck PR, Berman SR. Sleep deprivation in healthy elderly men and women: effects on mood and on sleep during recovery. Sleep 1986;9:492-501.
94. Reynolds CF 3rd, Kupfer DJ, Hoch CC, et al. Sleep deprivation as a probe in the elderly. Arch Gen Psychiatry 1987;44:982-90.
95. Moe KE, Larsen LH, Vitiello MV, Prinz PN. Estrogen replacement therapy moderates the sleep disruption associated with nocturnal blood sampling. Sleep 2001;24:886-94.
96. Vitiello MV, Larsen LH, Moe KE, Borson S, Schwartz RS, Prinz PN. Objective sleep quality of healthy older man and women is differentially disrupted by nighttime periodic blood sampling via indwelling catheter. Sleep 1996;19:304-11.
97. Dijk DJ, Duffy JF, Riel E, Shanahan TL, Czeisler CA. Ageing and the circadian and homeostatic regulation of human sleep during forced desynchrony of rest, melatonin and temperature rhythms. J Physiol 1999;516:611-27.
98. Dijk DJ, Duffy JF, Czeisler CA. Age-related increase in awakenings: impaired consolidation of non-REM sleep at all circadian phases. Sleep 2001;24:565-77.
99. Duntley SP, Kim AH, Silbergeld DL, Miller JW. Characterization of the mu rhythm during rapid eye movement sleep. Clin Neurophysiol 2001;112:528-31.
100. Wiegand L, Zwillich CW, Wiegand D, White DP. Changes in upper airway activation and ventilation during phasic REM sleep in normal men. J Appl Physiol 1991;71:488-97.
101. Bahill AT, Adler D, Stark L. Most naturally occurring human saccades have magnitudes of 15 degrees or less. Invest Ophthalmol 1975;14:468-9.
102. Herman JH, Barker DR, Roffwarg HP. Similarity of eye movement characteristics in REM sleep and the awake states. Psychophysiology 1983;20:537-43.
103. Fukuda T, Wakawura M, Ishikawa S. Comparative study of eye movements in the alert state and rapid eye movement sleep. Neuro-Ophthalmology 1981;1:253-60.
104. Jeannerod M, Mouret J. Recherches sur les mechanisms des mouvements des yeux observes au cours de la veille et du sommeil. Pathologie-Biologie 1963;11:1053-60.
105. Yasoshima A, Hayashi H, Iijima S, et al. Potential distribution of vertex sharp waves and saw-toothed wave on the scalp. Electroencephalogr Clin Neurophysiol 1984;58:73-6.
106. Sato S, McCutchen C, Graham B, Freeman A, Von Albertini-Carletta I, Alling DW. Relationship between muscle tone changes, sawtooth waves and rapid eye movements during sleep. Electroencephalogr Clin Neurophysiol 1997;103:627-32.
107. Kato M, Kajimura N, Okuma T, et. al. Association between delta waves during sleep and negative symptoms in schizophrenia. Pharmaco-eeg studies by using structurally different hypnotics. Neuropsychobiology 1999;39:165-72.
108. Atienza M, Cantero JL, Stickgold R, Hobson JA. Eyelid movements measured by Nightcap predict slow eye movements during quiet wakefulness in humans. J. Sleep Res 2004;13:25-9.

Table 3— Evidence Table

Author/ Ref Number	Evidence level	Nature of Study	Subjects (number /M-F /age- mean (range) / Normal or patients)	Methods/ Montage	Relevant Conclusions
Crowley/38	2	Cross-sectional study	34/19M, 15F/ young group 21.4+/-2.4 (18-25), old group74.6+/-6.8 (M0, 76.7+/-5.8 (F)/normal subjects	C_3-A_2, O_1-A_2/transverse EOG, submental EMG, spindles 11-16 Hz, K complexes defined according to R & K, density of spindles and K complexes compared, duration, amplitude, and frequency of spindles compared	Spindle density falls with age, also K complex density but less so. Spindle amplitude and duration also fall with age
Danker-Hopfe/26	1	Clinical Series	98/73M,25F/ Males: 52.3+-12.1, females: 49.5+-11.9/Patients	27 patients with psychiatric diseases, 15 with Parkinson disease, 5 with PLMS, 51 with OSA. Scored by 2 different centers, blinded to each other. Standard R&K rules used. No prior cross training. Experience of scorers not stated/ Minimum 6 EEG, Fp1, C_3, O_1 to M_2 and Fp2, C_4, O_2 to M_1, slightly modified EOG, submental EMG	Even with disordered sleep, overall inter-rater reliability falls in the substantial range. Main problems are confusion between stage 1 and 2 followed by stage 2 and SWS, and stage 1 and wake. REM was reliably scored. Age affects reliability of stage 2 and SWS scoring.
DeGennaro/75	5	Clinical series	10/10M/10 normal male students, mean age 23.4 years (SEM = 0.87, age range 20-30 years)	EEG signals filtered at time constant 0.3 sec, HFF 30 Hz. 7 EEG channels: C_3-A_2, C_4-A_1, FPz-A_1, F_z-A_1, C_z-A_1, P_z-A_1, O_z-A_1 Bipolar horizontal EOG (1 cm from medial and lateral canthi of the dominant eye) and vertical EOG electrodes placed 3 cm above and below the right eye pupil	K complexes appear to be a forerunner of delta sleep
DeGennaro/34	5	Clinical series	10/ 10M, 0F/ 23.4 (NA)/ Normal subjects - paid volunteers	C_3/A_2/C_4/A_1, FP_2, F_Z, C_Z, P_Z, O_Z / A_1 Horizontal & vertical EOG; bipolar submental EMG	1. C_3, C_4 best place to visualize spindles 2. Confirms visual observation of reciprocal oscillations between sigma and delta within NREM
Duntley/99	5	Clinical series	12 (scalp recordings) 5 (subdural grids)/ NA/ NA/ Patients with epilepsy, none with central foci	Visual analysis of EEG for 7-10 Hz central rhythms identical in morphology to waking mu/ Full- 24 EEG derivations plus EOG (RE-LE, undefined) and chin EMG	7-10 Hz mu like rhythms recorded centrally appears to be a feature of REM sleep seen in about 50% of patients. Frequency of occurrence in normal subjects unknown
Feinberg/40	4	Clinical series	9/3M, 3F/young group: 25.2(18.7-35.9), old group 75(64.7-86.8/normal subjects	LOC-M1/M2, ROC-M1/M2, LOC-midline forehead, ROC-midline forehead. EM during REM analyzed. REM defined atypically (pre-R&K study)	Fewer EM in REM recorded with mastoid compared to forehead reference. REM% may be reduced using mastoid references.
Finelli/87	5	Clinical series	8/ 8M, 0F/ 23y ± 0.46 (21-25)/ Normal subjects	Full–digital system sampling rate 128 Hz 27 EEG electrodes placed, extended version of International 10-20 System. Also, C_3/A_2 and C_3/A_1 & A_2 recorded. Average ref (all 27 derivations) used in power spectra	Reported individual differences, highly stable, frequency-specific patterns of EEG power distribution in NREM sleep. Points to importance of attention to individual traits rather than restricting analysis to common features

Author/ Ref Number	Evidence level	Nature of Study	Subjects (number /M-F /age- mean (range) / Normal or patients)	Methods/ Montage	Relevant Conclusions
Hakkinen/41	4	Clinical series	7/ 7M, 0F/ NA (24-39)/ Normal subjects	20 electrodes applied above left and below right eyes ranging in distance from 1-3 cm from eye margin in vertical plane and between mid-eyeball and 1 cm lateral to outer canthus in horizontal plane. Eye electrodes referenced to M1. Subjects asked to track through 30 angular degrees in all directions. Amplitudes of signals measured visually	Voluntary tracking EM in wakefulness best recorded with G1 electrodes 1 cm above and below eyes at horizontal level of outer canthi if referenced to M1. No data on EM potentials in sleep
Happe/35	5	Clinical series	10/ 3M, 7F/ 25.8 (20-35)/ Normal subjects	Full 10-20 electrode system, referenced to averaged M1/ M2 K complex defined: >75 µV/0.5 sec Delta waves: <2 Hz/>75 µV Spindle waves: 11-16 Hz /0.5 sec duration min FFT used for power maps	spontaneous K complexes max midline/frontal, delta in stage 4 max frontal
Jacobs/39	4	Clinical series	8/4M,4F/14-50 (NA)/ Normal subjects, 98 REM periods recorded	DC recording of eye movements/ EEG, EOG: 2 channels, lateral to each outer canthus, above and below R eye	REM sleep eye movements were more oblique than horizontal or vertical
Kim/27	5	Clinical series	2 or 3/all male/ "Young," low 20s/ normal 10 scorers	R & K and then slow wave rescored using a modified definition to separate 2 adjacent slow waves Epoch-epoch correlation (20-sec epoch)	The agreement rates for stages 2 and R high; those for stages 1, 3, and 4 were low. When a new rule was applied for scoring single slow waves (to distinguish 2 slow waves from 1 bimodal wave), the agreement improved for stage 3 (73%) but agreement for stage 4 remained low (48%)
Martin/28	3	Observational	5/5M,0F/NA (17-21) /Normal subjects, 3 human scorers	C3-M2; O1-M2; 2 channels of horizontal EOG electrodes attached to outer canthus], submental EMG R&K criteria. 3 human scorers versus a computer algorithm Epoch-epoch comparisons	Least agreement for wake, stage 1, stage 3, even when an O1 electrode was available
McCormick/36	5	Clinical series	8/2M, 6F/22.8 14-28/Normal subjects, 3 nights	Full 10-20 electrodes to A_1-A_2 linked reference, EOG, submental EMG. 4 15 min epochs of stage 2 sampled from each record. K complexes def with 75 microvolt min amplitude	Spindles are most detectable at C_3 (5.2/min). K complexes are most detectable at Fp1 and Fp3 (2.3/min.) Fewer than half the K complexes detected at Fp1 were detected at C_3
Monroe/29	4	Clinical series	1/1M,0F/23 (NA)/ Normal subject 27 scorers	F_4-A_2, P_4-A_1, R & K EOG, no EMG, 30-sec epochs, no criteria for stages presented to scorers (study pre-R&K)	Poor reliability for stages 3 & 4 but greater reliability when combined as delta sleep
Niiyama/62	5	Clinical series	12/12M, 0F/ 23.5 (21-25)/ Normal subjects	C_3, C_4, C_Z / A_1 & A_2, EOG, submental EMG Cortical AEP recorded during different stages of sleep	Physiologic changes of REM sleep may occur after last spindle or K complex but before muscle atonia develops
Norman/30	2	Clinical series	62/44M,18F/13-71/38 SDB, 10 CPAP titrations, 14 normals	C_3-A_2, O_1-A_2, EOG, submental EMG. 5 experienced technicians scored each record. 30-sec epoch-epoch comparisons	Most disagreement with stage 1 (esp. compared to stage 2), other stages reasonable reliability (some SWS-stage 2 disagreement)

Author/ Ref Number	Evidence level	Nature of Study	Subjects (number /M-F /age- mean (range) / Normal or patients)	Methods/ Montage	Relevant Conclusions
Ogilvie/49	3	Clinical trial	11/5M, 6F/ 22 (33-45)/ Normal subjects	Tones below arousal threshold administered. Probability (P) of responding to the tone calculated for each sleep stage /R & K montage	Behavioral and electro-physiologic measures of sleep onset differ. W-stage 1 junction is a more valid point of sleep onset than age 1-2 junction
Sangal/31	2	Clinical series	30/ 24M, 6F/ 53.9±10.4 (NA)/ Patients with OSA	Montage unclear ("central EEG", "left and right EOG", submental EEG) 2 scorers per PSG	Conventional R & K epoch based scoring has high interscorer reliability
Santamaria/48	5	Clinical series	50/ NA/ Adult/ Normal subjects	Qualitative visual observation/ Full-16 EEG channels, EOG outer canthus to outer canthus, from above middle of 1 eye to below middle of other, ECG	Many EOG and EEG features of drowsiness present before disappearance of alpha. Identification of drowsiness in patients without alpha especially dependent on EOG.
Smith/86	4	Clinical series	25/23M, 2 F/5 groups: ages 3-5, 13, 25-34, 43-53, 67-79/ normal subjects	F_1-F_7, C_3-A_2, signals digitized and analyzed, delta defined as 0.5-3 Hz, amplitude at least 5 µV.	Delta activity does not decline in abundance with age, but amplitude drops by 40 years
Van Sweden/19	4	Clinical series	10; 1F, 9M; men age 42.8 (range 25-610; patients	C_4-A_1 versus Fpz-Cz and Pz-Oz. Epoch-by-epoch comparison of scoring by the different montages for each of 2 scorers	Bipolar derivations are an acceptable alternative to R & K derivations.
Webb/90	5	Clinical series	40/ NA/ 55.3 (50-60)/ Normal subjects	Limited F_1/F_7, P_1, T_5, C_3/A_2, 1 channel EOG, no EMG Used one minute scoring epoch, stage 3: >13 seconds, <30 sec, 0.5-3 Hz Stage 4: ≥30 sec, 0.5–3 Hz, no amplitude criterion	1-Stage 3 & 4 can be scored reliably by using frequency and eliminating amplitude criterion 2-If amplitude criterion dropped, SWS% increases in this group 3-Age sensitive measure: amplitude, age resistant measure: frequency
Werth/37	5	Clinical series	20/20M,0F/23.2 (20-26)/Normal subjects	Power analysis performed at different frequency bands over different head regions during NREM and REM sleep/ F_3-C_3, C_3-P_3, P_3-O_1, also all electrodes to M2	Slow wave activity (presumably including K complexes) maximal frontally
Werth/71	5	Clinical series	6/ 6M, 0F/ 27.1±3.3 (23-32)/ Normal subjects	Staging in 20-sec epochs compared between R & K (C_3-A_2, submental EMG and R & K EOG) and R & K EOG with submental EMG but not C_3-A_2. Epoch –to-epoch comparisons not done/ Full	Stage 2 is decreased (cannot identify spindles as easily) and stage 4 sleep decreased when scoring with EOG and not C_3-A_2. This technique does not allow adequate accuracy in staging.
Whitney/32	4	Cross-sectional study	20 random selected records for intrascorer epoch-by-epoch reliability and 30 selected records for interscorer epoch-by-epoch reliability from Sleep Heart Health Cohorts	C_3-A_1 and C_4-A_2 EEG, right and left EOG, submental EMG	Good inter- (and intra-rater reliability as measured by percent positive agreement for wakefulness, stage 2 and REM.

Author/ Ref Number	Evidence level	Nature of Study	Subjects (number /M-F /age- mean (range) / Normal or patients)	Methods/ Montage	Relevant Conclusions
Wiegand/100	5	Clinical series	6/6M,0F/27+/-6 (23-40)/ Normal subjects	Ventilation determined in NREM, tonic REM and phasic REM sleep/ ventilation measured via pneumotachometer (nasal only, mouth taped shut), genioglossus and alae nasi EMG monitored	Reduction of tidal volume and increase in respiratory frequency characterize phasic but not tonic REM sleep

*The Visual Task Force members participating in consensus decisions to derive visual scoring rules included: Michael H. Silber (chair), Sonia Ancoli-Israel, Michael H. Bonnet, Sudhansu Chokroverty, Madeleine M. Grigg-Damberger, Max Hirshkowitz, Conrad Iber, Sheldon Kapen, Sharon Keenan, Meir Kryger, Thomas Penzel, and Mark Pressman.