Timing of human sleep: recovery process gated by a circadian pacemaker

DAAN, SERGE, DOMIEN G. M. BEERSMA, AND ALEXANDER A. BORBÉLY. Timing of human sleep: recovery process gated by a circadian pacemaker. Am. J. Physiol. 246 (Regulatory Integrative Comp. Physiol. 15): R161–R178, 1984.—A model for the timing of human sleep is presented. It is based on a sleep-regulating variable (S)—possibly, but not necessarily, associated with a neurochemical substance—which increases during wakefulness and decreases during sleep. Sleep onset is triggered when S approaches an upper threshold (H); awakening occurs when S reaches a lower threshold (L). The thresholds show a circadian rhythm controlled by a single circadian pacemaker. Time constants of the S process were derived from rates of change of electroencephalographic (EEG) power density during regular sleep and during recovery from sleep deprivation. The waveform of the circadian threshold fluctuations was derived from spontaneous wake-up times after partial sleep deprivation. The model allows computer simulations of the main phenomena of human sleep timing, such as 1) internal desynchronization in the absence of time cues, 2) sleep fragmentation during continuous bed rest, and 3) circadian phase dependence of sleep duration during isolation from time cues, recovery from sleep deprivation, and shift work. The model shows that the experimental data are consistent with the concept of a single circadian pacemaker in humans. It has implications for the understanding of sleep as a restorative process and its timing with respect to day and night.

Sleep deprivation; wakefulness; computer simulation; slow-wave sleep; circadian rhythm; non-rapid-eye-movement sleep

I. INTRODUCTION

No behavior so regular and pervasive characterizes our daily life and is yet so little understood as sleep. It is not known why we sleep, what triggers sleep, or what causes awakening. These questions have stirred much research in the past decade. In this Journal, a discussion of the theoretical principles of human sleep timing has recently been initiated (54, 55, 92, 95, 96). We present an alternative view.

There are two classic approaches to the mechanism of sleep and wakefulness. Circadian research has specifically addressed the spontaneous timing of sleep. Physiological sleep research has been especially concerned with the events within sleep. In each of these fields, a major principle is emerging, which is of significance for both the function and the regulation of sleep. There is a circadian principle, which may be formulated as "the longer we are active, the shorter we sleep." Consecutive alpha (activity time) and rho (rest time) values are commonly negatively correlated in animals as well as in humans (7). On the other hand, there is a homeostatic principle: "the longer we are active (and, perhaps, the more we are active), the deeper our sleep." Numerous reports show increased slow-wave sleep (SWS) after sleep deprivation (13, 63, 66, 81, 94) or decreased SWS after reduced waking time (39). A quantitative analysis of the electroencephalogram (EEG) signal has revealed that the power density of the EEG in the low-frequency range is increased during recovery sleep after sleep deprivation (20).
A combination of the two principles seems ideally suited to allow the organism to generate compensatory responses to variations in its need for sleep, while confining sleep to the appropriate part of its daily cycle. On the basis of these principles, we have recently developed the ingredients of a model for the timing of human sleep and wakefulness (18, 30). By computer simulation, we shall account for the main phenomena of human sleep timing (onset and duration) in a variety of experimental protocols. The analysis is based, in particular, on the extensive work by Wever (91).

Various parameters of the model were derived from existing empirical data. Subsequently these parameters were varied to simulate behavior in conditions of temporal isolation, forced bed rest, and the period after sleep deprivation. Here we present a full account of the model, indicate designs for its experimental testing, and discuss its implications for the understanding of sleep function.

A. Previous Models

Various models of activity-rest cycles exist, each designed to explain specific phenomena. Early approaches used physical and mathematical oscillators to elucidate the self maintenance of circadian rhythms under constant conditions and the principles of entrainment by (external) zeitgebers. One seminal theory assumed that inactivity occurs whenever a circadian oscillation falls below a constant threshold (6, 88). This tradition culminated in an elaborate model by Enright of circadian pacemakers composed of a multitude of neuronal elements (36). The single-oscillator approach demonstrated the main features of circadian self-sustainment and entrainment. It further convincingly revealed how the remarkable circadian precision may result from an ensemble of relatively sloppy and even damped neuronal oscillators (37).

Enright’s basic assumption was that the activity of neurons in the proposed pacemaker is reflected in the activity of the organism. However, the proposition that the oscillator itself directly generates sleep-wake alternations is incompatible with the fact that sleep-wake cycles have a range of frequencies much wider than is typical of circadian rhythms. Sleep-wake cycles exhibit 1) ultradian rhythmicity, both in a variety of adult mammals and in early human ontogeny; 2) midday naps persisting into adulthood; and 3) internal desynchronization in conditions of isolation, with frequencies ranging from the ultradian to the circadian range (periods of 12–65 h). These phenomena suggest that the occurrence of sleep is not directly determined by the phase of a circadian oscillator or pacemaker, although both ultradian and infradian sleep-wake cycles always appear subject to circadian modulation.

In the 1970s, two phenomena, splitting and internal desynchronization, led to the postulation of dual-oscillator models. In splitting, the circadian activity rhythm of small mammals shows two separate frequencies until synchrony with the two components in stable antiphase is attained (47, 73). Theoretical analysis has shown that oscillators have to be precise and nearly identical to produce such behavior, which suggested that two parts of the same pacemaker, probably located in the suprachiasmatic nuclei (SCN), are involved (31). Recent experimental evidence in hamsters, in which unilateral SCN lesions obliterated the splitting phenomenon without destroying the circadian rhythm (70), is consistent with this view. Similar evidence is available for insects (52, 93).

The other phenomenon that led to a two-oscillator model is the internal desynchronization in humans living isolated from time cues (9). On the basis of such experiments, Wever (90) proposed a system of two coupled circadian oscillators: a strong one controlling body temperature, and a weaker one controlling activity and rest. Kronauer et al. (55) formalized Wever’s model as a computer algorithm with two van der Pol oscillators. These authors assumed that sleep occurs during two-thirds of that part of the cycle in which one of the two oscillating variables is below average. This model accounts qualitatively for major aspects of spontaneous human sleep-wake behavior (55). However, the two-oscillator approach contains the counterintuitive assumption that the sleep-wake rhythm is generated by a circadian oscillator with an unusually broad and partly non-circadian frequency range. Physiological correlates of the parameters in the van der Pol equations remain unspecified and have so far not led to experimental predictions or tests. Adding further oscillators (51) increases the degree of freedom but does not seem to extend our physiological insight.

A different hypothesis was advanced by one of us to account for the structure of sleep after regular and extended waking periods (16, 18). It is assumed that sleep propensity, determined by a regulating variable (S), builds up during wakefulness and declines during sleep. Sleep or wakefulness occurs when S is above or below a threshold, respectively. The threshold itself is subject to a circadian rhythm. This model can explain various experimental results. However, it does not account for internal desynchronization, because the single threshold oscillation imposes its own frequency on the sleep-wake rhythm. We have therefore introduced the additional assumptions that during wakefulness the threshold operates at a higher level than during sleep and that this level depends on the circumstances prevailing during waking (30).

B. General Principles of the Model

The general structure of the present model is outlined in Fig. 1. We postulate a single circadian pacemaker, entrained by an external zeitgeber. The pacemaker is presumably located in the SCN of the hypothalamus and normally functions as a single unit, possibly on the basis of mechanisms suggested by Enright (36). It can be entrained by the light-dark cycle via the retinohypothalamic tract. Through its efferents it may generate numerous physiological circadian oscillations or synchronize these, if there is a reason to assume that they have self-sustaining capacity in the absence of the SCN, e.g., body temperature in primates (42, 75). Among these are oscillations in the two thresholds (L, H) for S. S increases monotonically during wakefulness until it reaches H, at
which point sleep is initiated. $S$ declines monotonically during sleep until it reaches $L$, when sleep is terminated. The system acts like a thermostat that switches off at a higher threshold than it switches on. Since we will use $S$, $H$, and $L$ as dimensionless variables (i.e., as fractions of the minimum-maximum range of $S$), the sign of the $S$ change is equivocal. We adopt the convention that $S$ decreases during sleep and increases during wakefulness. $S$ may be thought of as a chemical sleep factor, whereas $H$ and $L$ may reflect the sensitivity of hypothetical brain receptors for $S$. Although these assumptions are not necessary for the present purpose, they facilitate their conceptualization and may eventually lead to specific neurochemical hypotheses.

In the “somnostat” of the present model, the frequency of sleep-wake alternations depends on the interval between the two threshold levels and on the rate of buildup and breakdown of $S$. We assume that external conditions affect the threshold levels. Sleep deprivation experiments are simulated by suspending the upper threshold ($H$), thereby allowing $S$ to increase further. In contrast, bed rest, warmth, darkness, or the absence of social stimulation lowers $H$ so that sleep is precipitated. Culturally determined habits such as naps cause a transitory depression of the upper threshold. The wake threshold ($L$), operative during sleep, seems less influenced by the environment, although the ringing of an alarm clock may be effectively equivalent to a sudden rise in $L$.

The sleep-wake behavior may feed back to the circadian pacemaker (Fig. 1; see also sect. VID). Self-selection of the light-dark cycle should exert an effect on the entrainment of the circadian pacemaker. There is, however, no solid evidence for such a feedback. Finally, the sleep-wake cycle may directly affect many physiological oscillations or may exert a masking influence. This should be taken into account when comparing experimental data and theoretical predictions from the model. A discussion of possible entraining and masking influences is beyond the scope of the present paper. Since too little is known about phase responses to light in humans, theoretical explorations do not appear worthwhile at present.

In the following sections, we will derive estimates of
parameter values for average healthy human adults from experimental data and explore the consequences of varying some values by simulation. We are aware that the present model may also account for sleep patterns in animals and in human infants. To date, however, evidence sufficient to make our efforts fruitful has been accumulated only for adult humans.

II. SLEEP RENEWAL PROCESS

To make quantitative predictions we need to estimate the rates of increase and decrease of S. Following earlier suggestions (18, 20), we have used for this purpose empirical data on the integrated power density of the human sleep EEG. Although the physiological significance of this parameter remains to be specified, it has the required properties of decreasing progressively in the course of sleep and of increasing as a consequence of extended wakefulness.

A. Derivation of the Time Constants

The decrease in EEG power density in the 0.75- to 25-Hz domain during sleep is essentially exponential. From the data of Borbely et al. (20), an instantaneous breakdown rate (r) of 0.238/h can be calculated. In the simulations we used 30 min periods as a convenient time unit. We calculated $S_i$, the value of $S$ at time $i$ during sleep, by multiplying the preceding value $S_{i-1}$ by a factor $e^{-0.238/2} = 0.888$. S asymptotically approaches zero during sleep. We further assume that during wakefulness, $S$ approaches an upper asymptote (U). The rate of this buildup can be derived by fitting an exponential curve of the form $S_i = U - e^{-r/2}(U - S_{i-1})$ using three data points, as indicated in Fig. 2. These points are the relative EEG power densities at the end of a normal sleep night ($t = 0$; Fig. 2, h 7), at the end of a normal waking day of 17 h ($t = 34\ h$ of 30 min; Fig. 2, h 24), and after 24 h of sleep deprivation (41 h of waking, $t = 82$; Fig. 2, h 48). The instantaneous rate ($r'$) amounts to 0.055/h. For convenience, we have set the upper asymptote U equal to 1, thus making S dimensionless. In the simulation procedure, S is calculated iteratively per half hour from $S_i = 1 - 0.973(1 - S_{i-1})$.

The rates $r$ and $r'$ are based on data from 8 healthy young subjects (20). The maximum EEG power density occurs during SWS, which is well known to predominate in the first part of the sleep period (38). Since most studies of sleep structure have been based on sleep stages, there are few other data adequate to verify these parameters. In one study describing the overnight decay in relative power density in the 0.5- to 2-Hz band ($\delta$-sleep) an exponential decay function was derived (48). The instantaneous rate of decrease was $-0.27 \pm 0.11$ (SD) h$^{-1}$, which is close to the rate of $-0.238 \ h^{-1}$ derived from Borbely’s data. A study of Bos et al. (24) further revealed that relative power densities (fractions of the initial power density) are independent of electrode placement. The exponential decay in power density may thus represent a fundamental quantitative property of the sleep EEG.

In deriving these time constants, we have tacitly assumed that the breakdown and buildup processes are monotonic. Future research may force us to the position that $S$ decreases only during non-rapid-eye-movement (NREM) sleep and remains constant or, in view of the EEG similarities with wakefulness, even increases during rapid-eye movement (REM) sleep. If this were the case, the decay rate would be higher during NREM sleep, whereas the overall average decay rate would still remain close to $-0.238 \ h^{-1}$.

B. Effect of Threshold Oscillations

We have assumed that the thresholds at which $S$ induces and terminates sleep are subject to circadian oscillations and have introduced these initially (30) as sine waves to obtain a qualitative evaluation. In Fig. 3, examples of simulations obtained with different circadian amplitudes (A) and mean levels of the thresholds ($\bar{L}$, $\bar{R}$) are shown. Obviously the average period of the sleep-wake alternation ($T_{SW}$) is dependent primarily on the interval between the thresholds, whereas the amplitude of the sine waves affects the extent to which the sleep-wake rhythm is synchronized. Raising $\bar{R}$ or reducing $A$ both result in “internal desynchronization” between the circadian rhythm and the sleep-wake cycle. Lowering $\bar{R}$ causes a higher frequency of the sleep-wake cycle, or “sleep fragmentation.” Thus, in a qualitative sense, variation in sleep-wake frequency with a persisting circadian modulation is a characteristic property of the model.

![FIG. 2. Derivation of rates of decrease and increase of $S$ from relative EEG power densities. Left panel: regression lines of log power densities during base-line sleep (lower curve) and after 24 h sleep deprivation. (From Ref. 20.) Right panel: same curves plotted on linear scale. $S_0$, terminal value on base-line nights; $S_{34}$, initial value on base-line nights, 17 h after $S_0$. $S_{82}$, initial value on nights following sleep deprivation. Exponential curve was fitted through these points. (Modified from Ref. 30.)](http://ajpregu.physiology.org/doi/abs/10.1210/jc.29-2.246)
Since sleep deprivation involves a forced extension of wakefulness, it can be simulated by a temporary suspension of H. This results in a prolonged exponential increase of S. The duration of the recovery sleep depends on the circadian phase of sleep onset. Initial simulations on the basis of a sine wave circadian oscillation have shown (18, 30) that recovery sleep is gradually shortened after deprivation periods lasting up to about 16 h beyond normal sleep onset (Fig. 4, lower panel). After still longer sleep deprivation periods (>17 h), the duration of sleep suddenly exceeds the base-line level and then gradually decreases. These simulations show a remarkable correspondence with sleep deprivation experiments carried out in six human subjects by Åkerstedt and Gillberg (3). Figure 4 (upper panel) illustrates the sudden increase in spontaneous sleep duration of individuals around the predicted time.

Sleep deprivation experiments are well suited to illustrate the two basic principles on which the model is based: 1) the circadian principle, with longer wakefulness followed by shorter sleep, at least within a normal range of variation, and 2) the homeostatic principle, with higher S (or more EEG power) following longer wakefulness. Since the model also accommodates qualitatively the main features of isolation experiments, it seemed worthwhile to obtain a quantitative estimate of the circadian threshold variations.

III. CIRCADIAN SLEEP PROPENSITY

Quantitative estimates of the circadian threshold variations can be derived by measuring either the tendency to go to sleep during wakefulness or the tendency to wake up during sleep. The first approach could exploit circadian changes in self-ratings of fatigue during prolonged sleep deprivation (2). This would raise the problem of how fatigue ratings translate into actual decisions to go to bed. We used the alternative approach and derived the form of the circadian oscillation from the S level at the time of awakening. This level can be calculated directly from the duration of sleep in subjects waking up at different times of the day.
A. Derivation of Wake-Up Threshold Oscillation

If 7:00 A.M. and 11:00 P.M. represent the average times of spontaneous sleep termination and onset, respectively, as was the case in the study of Akerstedt and Gillberg (3), then two points on the circadian threshold curves, S₀ and S₃₂ can be calculated from

\[
S₀ = S₀^* \times 0.888^{16} \text{ (at 7:00 A.M.)}
\]

\[
S₃₂ = 1 - (1-S₀^*) \times 0.973^{32} \text{ (at 11:00 P.M.)}
\]

The resulting values are S₀ = 0.093 and S₃₂ = 0.622. S values at the end of x h of sleep deprivation (beyond the regular sleep time) are then calculated from

\[
S(x) = 1 - (1-S₃₂^*) \times 0.973^{2x}
\]

and those for spontaneous awakening times after y h of sleep as

\[
S(x, y) = S(x) \times 0.888^y
\]

By using the average sleep durations (y) following sleep deprivation (over x h), as presented by Åkerstedt and Gillberg (3), seven S(x, y) points were obtained (Fig. 5). A curve was fitted by eye through these points. This curve can be approximated closely by a skewed sine wave. In each phase position θ, we calculated θ', where θ' is determined by sin θ' = α ÷ (φ - φ'). A value of α = 2.07 gave the best fit between the resulting skewed sine function and the data.²

The threshold curve thus derived has a maximum at about 2:00 P.M. and has a minimum around 5:00 A.M., approximating the curve of self-rated fatigue in sleep deprivation experiments of Fröberg et al. (41) and Åkerstedt and Fröberg (2) (Fig. 6). It is therefore not unreasonable to assume that the curves for H (the tendency to go to bed during wakefulness) and L have similar waveforms. Assuming that the value of S₃₂ = 0.622 lies on the upper curve, H can be obtained by displacing curve L upward by 0.50 S units. The whole circadian process can then be described by four parameters, which for entrained conditions are estimated as 1) A = 0.12; 2) \(τ = 24 \text{ h}; 3) L = 0.17; and 4) H = 0.67.

B. Precision

In everyday life, the accuracy of the human sleep-wake system may be largely under control of conscious decisions in a framework of personal and cultural habits, work schedules, and other social factors. The intrinsic precision of the system can be evaluated from its performance in the absence of external time cues. Relevant data are available from the extensive analyses by Wever (91). To account for stochastic variation in the model, variance may be introduced either in the buildup and decay of S, in the thresholds, or both. Variance in the thresholds would satisfy the notion that the response to fatigue is not a fully deterministic process, since unpredictable influences may precipitate or postpone the decision to go to bed. Therefore, although the outcome is not fundamentally different, we have chosen to introduce variance in the thresholds rather than in S.

Variance was introduced by adding to the threshold level a variable of the form \( y; = \zeta(y_i - \bar{y}) + x \) for each half-hour interval, where x is drawn from a Gaussian distribution with mean 0 and standard deviation p. For p, a value of 0.022 was empirically established in order to approximate the variability in sleep times as observed in conditions of temporal isolation (91).

A FORTRAN simulation program calculated H, L, and S for each 30-min period and examined whether S exceeded H or fell below L. Simulations were usually run for 60 circadian cycles. In the simulations presented below, we modified \( \tau \), A, and \( \Pi \) for temporal isolation (see sect. IVA) and H for sleep-inductive conditions (see sect. IVB) and shift work schedules (see sect. IVD) and set \( L = H = 1 \) for sleep deprivation (see sect. IVC).

IV. SIMULATIONS OF EMPIRICAL RESULTS

A. Spontaneous Sleep Timing in Isolation from Time Cues

The values of the parameters in the model (except the noise parameter p) were defined on the basis of data

²Unfortunately the resulting curve cannot be described by a simple analytic function, although its computational algorithm is simple enough. However, since the series expansion of the skewed sine wave converges quickly, it can also be described by \( f(\phi) = 0.97 \sin \phi + 0.22 \sin 2 \phi + 0.07 \sin 3 \phi + 0.03 \sin 4 \phi + 0.01 \sin 5 \phi \). The advantage of the skewed sine wave is that its periodicity is easily understood and that it shows only a slight (but significant) deviation from the waveform we originally chose for convenience.
obtained from healthy adults, under either natural or shifted sleep-wake schedules. These subjects were exposed to various reinforcers in their social environments which help the sleep-wake process to remain in synchrony with other circadian rhythms as well as with the external world. Conditions of isolation from time cues may be expected to have a number of consequences.

1) The period of the circadian rhythm no longer equals that of the earth’s rotation cycle. Since in Wever’s data (91) the average period of the body temperature rhythm is $25.00 \pm 0.50$ (SD) h with remarkably small interindividual variance, we have, in most cases, set $\tau = 25.0$ to simulate temporal isolation.

2) In a social environment subjects will try to synchronize to the generally accepted bed times, which leads to relatively small variance in the phase of the sleep onset. In conditions of temporal isolation we therefore expect an increased interindividual variance. This is consistent with the proposition that the absence of social stimuli (boredom or improved concentration) may lead to either a lowered or a raised $H$. We shall analyze the consequences of varying this threshold.

3) A third consequence of isolation from time cues is probably a slight reduction of the amplitude of the circadian threshold oscillation. Although daily rhythms in biology generally persist in the absence of zeitgebers, their amplitude usually decreases. This effect is dramatic in plant circadian rhythms (25) but is also present in widely different animal rhythms (30). In humans, Weitzman et al. (86) and Wever (91) presented evidence that the circadian amplitudes of body temperature and cortisol circadian rhythms are reduced under free-running conditions. Moreover it is obvious from physical oscillator theory that resonance with a zeitgeber increases the amplitude of a self-sustained oscillator. In some simulations we have therefore gradually reduced $A$ from the default value of 0.12 to 0.08 and in others we have chosen arbitrary constant values between 0.08 and 0.12.

1. Distribution of sleep wake frequencies. The effect of varying $H$ in temporal isolation was analyzed by making a series of simulations in which the value of $H$ was increased from 0.36 to 1.06 in steps of 0.0005. Average periods of the ensuing sleep-wake cycle ($\tau_s$) were calculated and plotted in Fig. 7 (lower panel) and were compared with the values obtained from humans in isolation (Fig. 7, upper panel). Simulation and empirical data correspond to the extent that both distributions show a high incidence of 25-h (circadian) and 50-h (circadian) periods of the sleep-wake cycle and a low incidence of intermediate values. However, two major discrepancies between the two distributions emerge.

1) In the simulations there are relatively more circadian cases than in reality. The incidence of extreme $\tau_s$ values is higher than in the data. We assume that this is due to our choice of a uniform distribution of levels of $H$ instead of a Gaussian distribution around a mean level, which would have reduced the extreme values. However, we have no way of adequately selecting a mean and a standard deviation of the distribution of levels independently from Wever’s data.

2) Wever’s results (Fig. 7, upper panel) show a total absence of sleep-wake periods ($\tau_s$) of 21–23 and 27–30 h, an observation that has been interpreted as supporting the two-oscillator model (91, p. 216 and 217). In the simulations, no reduction is seen in these frequency ranges. We surmise that this discrepancy is due to differences in the analyses of the data. Average periods of 22 and 28 h in our simulations are the result of sleep-wake rhythms incidentally displaying complete phase jumps over a circadian cycle. Between such jumps the sleep-wake rhythm runs in synchrony with the circadian cycle (Figs. 8 and 10). In computing the frequency distribution of Fig. 7 (upper panel), Wever did not take average periods over the entire experiment but selected sections in which no obvious phase jumps occurred. We suggest that this procedure excludes a priori the incidence of periods close to the circadian period and that their absence is an artifact of the data analysis rather than an intrinsic feature of the system. This is supported by Zulley’s reanalysis of some of the same data (97).
2. Internal synchronization. In true internal synchronization, where the period of the free-running sleep-wake rhythm equals that of other physiological circadian rhythms, some peculiar correlations of sleep timing with the circadian period have been described. In analyzing data from experiments where the circadian period was reduced by exposing subjects to an electromagnetic 10-Hz field, Wever (91, Fig. 104) observed the following simultaneous trends: 1) a slight increase in the wake-sleep ratio (α:φ); 2) an increase in the precision of sleep timing, expressed as the cycle period divided by the mean of the standard deviations of the end (σr) and onset (σo) of wakefulness times \[ 2 \cdot \frac{r}{(σ_r + σ_o)} \]; and 3) a change in the relative precisions of the sleep onset and awakening time such that sleep onset became less precise than awakening (σr/σo increased). Since these changes are not intuitively obvious consequences of any specific model, it was interesting to determine whether the same patterns would emerge in simulations. In 24 runs, \( T \) was set to correspond to the observed circadian periods (91, Table 8), whereas A was reduced to 0.08 to allow for conditions of isolation. The resulting average values of \( \alpha:φ \), the precision, and the relative standard deviations of sleep onset and termination are compared with Wever's results in Fig. 9. There is, on the whole, a reasonable

Fig. 7. Frequency distribution of period (\( T_c \)) of sleep-wake cycle. Upper panel: empirical data from Wever's study (91), derived from sections of individual records. Lower panel: average values over 60 cycles in each of 1,400 simulations with R varying stepwise from 0.36 to 1.06 in increments of 0.0005.

Fig. 8. Example of simulation resulting in sleep-wake cycle with average period (\( T_c \)) of 22 h. R = 0.575; A = 0.08; \( T = 25 \) h.

Fig. 9. Dependence of characteristics of free-running sleep-wake cycle on period of circadian oscillation. Lines connect measured values in individual free running in 2 different situations leading to long and short periods (91, Fig. 104). Group averages (dashed lines) are compared with averages obtained in 2 x 12 simulations, with \( T \) set equal to observed \( T \) values (dots). Left panel: ratio of wakefulness and sleep time. Middle panel: ratio of \( r \) and mean (σ) of standard deviations of wake onset (σo) and wake end (σe). Right panel: ratio of standard deviations σe and σo.
correspondence between the simulations and the empirical data. To some extent this result is trivial, because the model parameters were selected to produce $\alpha \rho$ of about 2.0, a value that generally applies to both entrained and isolated conditions, and because in the selection of the noise parameter ($p = 0.022$) we have attempted to match Wever's data on average precision. The following consequences, however, are not trivial: 1) all three measures shown in Fig. 9 are different for long and short periods, both in the simulated and the empirical data; and 2) the standard deviation of the wake-up times ($\sigma_w$) is consistently smaller than that of sleep onset times ($\sigma_s$), again both in simulated and actual data. These features were not built into the design of the model but are consequences of the angles at which $S$ intersects with the threshold curves. The ratio of the standard deviations in the timing of the two events was slightly larger in the actual data than in the simulations. This suggests to us that the variance of the upper (sleep) threshold may be larger than that of the lower (awakening) threshold. This corresponds with the intuitive notion that external conditions are subject to more day-to-day variation during wakefulness than during sleep.

The data on $\alpha \rho$ and on precision fully describe the sleep-wake rhythm in internal synchronization. The average phase of the sleep-wake rhythm is such that sleep onset typically precedes the minimum of the threshold oscillation by about 6 h (Fig. 5). This is very similar to the phase angle difference between sleep onset and the body temperature minimum. It is interesting to examine whether the circadian rhythms of body temperature and of the hypothetical sleep thresholds correspond also in other respects.

3. Internal desynchronization. In Fig. 3 we have seen that a reduction in the amplitude of the threshold oscillations leads to the occasional skipping of a sleep episode, whereupon sleep occurs at an earlier phase and at a higher level of $S$ in the next cycle. Over a long simulation run this results in a longer average sleep-wake cycle ($\tau_s$). In Fig. 10 we have plotted one such simulation in standard actogram format to compare it with one of Wever's (91, Fig. 3b) well-known records. In this case, $A$ was gradually reduced from 0.12 to 0.08 over 20 cycles and then held constant at 0.08. Spontaneous skipping of sleep episodes led to a longer average sleep-wake cycle ($\tau_s = 28.1$ h) in the second half of the run. The simulation and the real data correspond in considerable detail. In the three temporally synchronized sections (Fig. 10, lower right panel) $\tau_s$ is longer (26.1 h) than the preset circadian period ($\tau = 25.0$ h), an effect that is already present to a minor extent in the beginning of the record. These differences are due to a gradual change in the phase relationship over the course of the run. Corresponding changes were observed in the empirical data (Fig. 10, left panel). We emphasize that although no feedback effects from the sleep-wake cycle on the circadian threshold curve have been incorporated into the model, differences in period between the two sections of the simulation were observed. Therefore such differences as consistently observed in human experiments (91, p. 52) can be explained without invoking a feedback effect, although the existence of such an effect has not been disproved. It should also be emphasized that the correspondence of simulation and observation in Fig. 10 is dependent on the precise choice of the parameter value $\Lambda = 0.08$ and $R = 0.85$. Lower values of $\Lambda$ and higher values of $R$ lead to more frequent "jumps." Higher $A$ values prevent internal desynchronization. As we have no independent estimates of parameter values for individuals, the model does not explain why the subject in Fig. 10 behaved in this particular way. It is the detailed consistency between simulation and observation that suggests the usefulness of the approach.

Inspection of Fig. 10 further shows that the three spontaneous phase jumps with the skipping of a sleep episode are preceded by a late awakening relative to the maximum of $I$, and is followed by a long sleep episode. In a series of simulations of conditions of internal desynchronization, this leads to a peculiar relationship between the phase of sleep onset and sleep duration (Fig. 10).
A. Distribution of Spontaneous Sleep Onset Times with Respect to Circadian Minimum of Rectal Temperature

In the phase (ψ)-duration (D) plane, two areas of concentration are present. One centers around ψ = -24, D = 8 h, the other around ψ = -8, D = 14 h. In the case of internal synchronization, the latter concentration is absent. A similar distribution (Fig. 11A) has been described by two laboratories studying sleep-wake timing relative to the body temperature rhythm in conditions of isolation (28, 98). The two concentrations of points lie closer together in the data than in the simulations. Since the former are based on measured body temperature, the well-known hypothermic effect of sleep may have led to an earlier occurrence of the temperature minimum in the case of early sleep onset.

B. Sleep Enhancement and Fragmentation by Continuous Bed Rest

Even in the presence of daily time cues, external conditions can exert a considerable influence on the timing of sleep and wakefulness. Large modifications of the sleep-wake cycle have been obtained by subjecting people to continuous bed rest and preventing social and intellectual activities. Such a treatment leads to a fragmentation of sleep (27, 65). We have attempted to simulate this striking alteration of the sleep-wake cycle.

We assume that continuous bed rest and social deprivation lower H. This is consistent with the proposition that external conditions may affect the system via the thresholds (Fig. 1) and with the experience that boredom and bed rest are sleep-inducing conditions. Lowering H causes an increase in the average frequency of the sleep-wake cycle (Fig. 12). Simultaneously total sleep duration is prolonged. This is because the rising and falling parts of the S curve are now situated at a lower average value, which entails a more rapid rise and a slower decay. Since buildup and decay on average must balance, the wake-sleep ratio decreases. The model predicts quantitatively how, in the average subject, the percentage of sleep will vary with the average period of the sleep-wake cycle (Fig. 13). To obtain comparable empirical data we recalculated the original records (S. Campbell, unpublished research) from human subjects during bed rest. The experiments had been performed on nine healthy male adults who were confined to bed in a dark, clockless room for up to 60 h.

The results show an average sleep-wake cycle of 7.31 h (normal 24 h) and an average fraction of the time spent asleep of 50.8% (normal 33%). The simulations (with H = 0.245) show for a frequency of 7.38 h a sleep fraction of 51%. This result does not discouragingly deviate from

FIG. 11. A: distribution of spontaneous sleep onset times with respect to circadian minimum of rectal temperature in 10 subjs showing internal desynchronization. Upper panel: histogram of frequency of sleep onset times. Lower panel: dependence of sleep duration on sleep onset time. (Data from Ref. 98.) B: distribution of sleep onset times with respect to circadian minimum of threshold oscillation in 7 simulations showing internal desynchronization. Conventions as in A.

FIG. 12. Proposed effect of continuous bed rest on upper threshold and its consequences for sleep wake pattern.
the average values observed. However, there was a large variation between individuals, in both the sleep-wake period and the percentage of sleep. There was no interindividual correlation between these variables. Thus, whereas on the basis of parameter values derived from averages the model is able to predict the average results, it fails so far to account for interindividual variance. It is possible that the interindividual variation of \( H \), the parameter manipulated in the present simulations, is smaller than the variation of other parameters. We await individual assessment of parameter values for further testing of the model.

C. Sleep Deprivation

Sleep deprivation has been repeatedly used as a tool for investigating sleep regulation. As we have shown for the model with sine wave threshold curves (sect. IIC), simulated and actual data for sleep duration following sleep deprivation show a fair correspondence. Since some quantitative aspects of the model were based on sleep duration data following different sleep deprivation schedules, they cannot be used for independent prediction of sleep duration. However, two other predictions can be derived. The first concerns the changes in EEG power density after various deprivation schedules. The model predicts that sleep deprivation of increasing duration will be followed by increasing initial power densities. In contrast, power density before sleep termination should be low after short (0–4 h) and long (16–24 h) sleep deprivation, but higher after intermediate (8–12 h) deprivation. Although Åkerstedt and Gillberg (3) did not present EEG power density data, they recorded the percentage of time spent in the sleep stage with maximum power (20). The percentage of SWS provides a rough approximation of power density. The predicted curves of \( S \) and the measured curves of SWS show a surprisingly close correspondence, including an increasing trend in initial SWS and a periodicity in terminal power and SWS in the seven deprivation schedules (Fig. 14).

A further prediction concerns the second recovery night after sleep deprivation. In Fig. 15 we have indicated the predicted durations of first and second sleep periods following sleep deprivation. Although the duration of the first recovery sleep corresponds with the results of Åkerstedt and Gillberg (3), this result is trivial, since their data were used to quantify the model. The simulated second sleep shows peculiar variations that are not intuitively obvious. Both a short (0–8 h) and a long (16–24 h) sleep deprivation are virtually compensated in the first sleep period, such that the second sleep is of normal length (ca. 8 h). However, sleep deprivation lasting 8–16 h has repercussions on the second sleep period. Sleep deprivation of 8–12 h is followed by two short sleep episodes, sleep deprivation of 13–16 h by two long sleep episodes. Since, to our knowledge, the spontaneous timing of the second sleep night following sleep deprivation has not been experimentally measured, this prediction awaits validation.

D. Sleep in Shift Workers

Shift work schedules often involve repeated sleep deprivation. They can be simulated by letting \( L \) and \( H \) equal 1.0 during an 8-h work period and during two 1.5-h periods before and after work to allow for meals and transportation (i.e., 11 h of forced wakefulness). We have...
also taken into consideration that most shift workers continue to live in a social environment that favors sleeping at night [elegantly documented, for instance, in a field study on train drivers (4) (Fig. 1)]. We accounted for this cultural pattern by lowering the mean level of the upper threshold between 11:00 P.M. and 7:00 A.M. to $H = 0.245$, the value derived from the bed-rest data (see sect. IVB). Between 7:00 A.M. and 11:00 P.M. the standard value of $H = 0.67$ was maintained. These values do not apply during the forced wakefulness.

Figure 16 presents the average sleep durations predicted for a series of 24 work shift onset times. These are compared with questionnaire data from shift workers in extremely varied employment conditions (53). There is a reasonable correspondence between simulated and empirical data. In particular, the model predicts the minimal sleep duration known for night shifts. Discrepancies between model and observations may be partly due to the fact that the empirical data originate from numerous different rotating schedules, whereas the simulations assume a steady state.

V. EXPERIMENTAL DESIGNS FOR TESTING THE MODEL

A model does not derive its usefulness from the functional correspondence with the system it is designed to describe. It may, however, improve our insight when it generates predictions of unknown behavior. These predictions may point the way to new experimental designs, which, in turn, may serve to validate or reject the theory.

Appropriate tests should concern the three main ingredients of the model: 1) the relaxation oscillation (process S), supposedly reflected in the EEG power density; 2) the existence of a threshold determining sleep termination (L); and 3) the existence of a threshold determining sleep onset (H).

The first point can be tested by almost any sleep-wake schedule, since the model generates precise predictions on the value of S and its changes, and hence on the EEG power density during NREM sleep. A consistent increase in EEG power density over consecutive NREM sleep episodes, or a decrease over a wake period, would require a fundamental revision of the model. Data on quantitative variations between individuals may allow us to extend predictions to the level of the individual.

The existence of a wake-up threshold (L) can be tested by determining the time of spontaneous awakening in imposed sleep schedules. In particular, it should be possible to obtain an EEG power value during early sleep and predict NREM-EEG power at the time of awakening. Moreover it is necessary to establish whether the threshold is solely a function of circadian phase and, if so, whether it obeys the rules for other circadian rhythms, e.g., phase shifting in response to light pulses as observed in animal experiments (32).

Testing the existence of a sleep onset threshold (H) is more problematic. 1) The time course of S during wakefulness has not yet been based on a measured variable but has to be extrapolated from the terminal EEG power in the preceding sleep period. A nap study might provide more direct information. 2) Internal and external influences are likely to affect H to a much greater extent during wakefulness than they affect L during sleep. Constant conditions as realized in Campbell’s experiments (Ref. 27 and unpublished data) are mandatory for further test. Such conditions presumably lower the upper thresh-

![FIG. 16. Sleep durations of shift workers in simulation of work shifts starting at different hours of the day. Open circles, observed averages (+ 1 SD) in shift work studies. Data are pooled for morning shifts (starting 0500-1000), midday shifts (1000-1400), evening shifts (1400-1800), and night shifts (2200-0500). (From Ref. 53.)](http://ajpregu.physiology.org/)

![FIG. 15. Durations of 1st and 2nd spontaneous sleep times after sleep deprivation of different duration beyond normal waking time predicted by model simulations. Vertical lines, ranges of variation for 1st sleep durations as measured by Åkerstedt and Gillberg (3).](http://ajpregu.physiology.org/)
old to a level at which it is attained by S up to four times a day, thereby enhancing the opportunity for testing sleep onset time as a function of prior waking and circadian phase. If the assumptions of the model are valid, it should be possible to predict the time of sleep onset and the initial EEG power density on the basis of the EEG power curves of the preceding sleep period.

VI. IMPLICATIONS AND PERSPECTIVE

We have shown that a model containing only a single circadian pacemaker that gates a noncircadian homeostatic sleep-wake process is sufficient to explain in quantitative detail the major phenomena of human sleep timing. This is the case both for experimental conditions where either sleep is manipulated (sleep deprivation and forced bed rest) or zeitgeber input to the circadian pacemaker excluded and for nonexperimental schedules as imposed by shift work. So far the parameters of the model have been specified exclusively by empirically derived average values. Therefore interindividual variations in sleep timing are not yet predictable. We feel confident, however, that by measuring parameters on an individual level and by generating predictions for individual behavior in experimental schedules, the model can be put to a critical test.

It is appropriate to discuss here aspects of human sleep timing that led others (55, 91) to postulate two circadian oscillators and to consider some relationships of the present proposition to sleep physiology and function.

A. Arguments for a Separate Circadian Sleep-Wake Pacemaker

In accordance with the model of Wever (91) and Kronauer et al. (55), the present model contains two rhythmic entities. However, in our view, the sleep-wake rhythm itself is not generated by a circadian pacemaker. Process S might be considered an oscillator, but it is certainly not a pacemaker (as in Ref. 55). In view of its large frequency range and its sensitivity to external conditions, it does not fulfill the requirements for circadian oscillation.

The conclusion that a circadian sleep-wake oscillator exists in humans was based on the fine detail of day-to-day variation of sleep timing in temporal isolation. Wever (91, p. 71) considered the negative serial correlation between activity and subsequent rest as evidence for an oscillatory origin of the sleep-wake rhythm. Indeed, due to the properties of process S, only positive serial correlations would be expected in the present model if the thresholds were kept at a constant level. However, in the presence of circadian threshold changes, the serial correlations between wake and sleep episodes become negative (sect. IIIB and Ref. 30). Hence the sign of this correlation cannot be used to identify the type of process underlying the sleep-wake cycle.

Kronauer et al. (55) have placed particular emphasis on the phenomenon of “relative entrainment” or “phase trapping” as indicating a separate sleep-wake pacemaker. A self-sustained oscillator, entrained by a second one with a period close to the limits of entrainment, may periodically change its phase relationship to the latter (33, 89). Relative entrainment would not be produced by the S process entrained by a circadian pacemaker. If phase trapping were a general phenomenon, it could be readily explained by feedback effects of the subjectively experienced light-dark cycle on the circadian pacemaker (Fig. 1). However, from the extensive material on human subjects, only a single record that was interpreted as phase trapping has been published (55, Fig. 1A). The record shows 3 days in which sleep onset was advanced relative to the temperature cycle. This observation seems too incidental to be accepted as solid evidence for the existence of a separate self-sustained oscillatory process underlying sleep and wakefulness. We therefore disagree with the contention that the published human data represent convincing evidence for the existence of multiple circadian pacemakers (e.g., 62, p. 157). Indeed the sudden fourfold increase in sleep-wake frequency during continuous bed rest (27) seems incompatible with the notion of a circadian sleep-wake pacemaker. In contrast, this result is predicted in quantitative detail by a homeostatic S process (Fig. 13).

B. Entrainment and Internal Coupling

The implication of the model for the interactions of external zeitgebers, the circadian and S processes, invite further comments. Kronauer et al. (55) have suggested that light-dark schedules entrain the circadian system by acting on the sleep-wake oscillator. They argued that entrainment of the temperature oscillator would arise from its coupling to the sleep-wake oscillator. However, there is evidence in humans that light-dark cycles are able to entrain the rhythm of body temperature, whereas the sleep-wake cycle free runs (10). In view of the evidence from other mammalian species, it is likely that light-dark schedules directly entrain the circadian pacemaker. On the other hand, light and dark have direct effects on the sleep-wake cycle itself. Wever (91, Fig. 9b) has shown that light-dark schedules involving total darkness may induce sleep-wake cycles with period far outside the circadian range. After discontinuing such schedules, the sleep-wake rhythm rapidly returned to a circadian frequency. In the present model, these influences of light are interpreted as direct effects via the threshold. Total darkness, even if occurring out of phase with the usual circadian sleep time, may precipitate a subject’s decision to go to bed. We presume that this in turn lowers H and leads to sleep onset. Since S is not a true self-sustained oscillator, this effect of light and dark would not be genuine circadian entrainment. This view is further supported by the observation that the sleep-wake cycle adopts a wide range of frequencies under such treatment.

Internal coupling of the circadian temperature oscillator to the sleep-wake cycle is a basic feature in the two oscillator models of Wever (91) and Kronauer (55). Sleep has direct effects on other variables (e.g., body temperature) under control of a circadian oscillator. Such effects are usually designated masking, although the term is clearly defined only for changes induced by external conditions (8). Masking effects do not constitute evidence for internal coupling.
The slight differences in the periods of body temperature rhythms, derived from the maximum values in synchronized and desynchronized sections of actograms (average difference 0.70 ± 0.38 h) were interpreted by Wever (91, p. 52) as being due to feedback from the sleep-wake cycle. However, in our model, similar differences emerge without internal coupling (Fig. 10). Aside from this argument, it should be noted that the assessment of the circadian period on the basis of the body temperature maximum alone is not without problems.

Although coupling of the circadian pacemaker to the sleep-wake cycle has yet to be demonstrated, it is likely that some feedback influenced by the sleep-wake behavior occurs via the perception of light and dark. In the relatively low light intensity of most experimental rooms, such effects are probably sufficiently small to be initially disregarded in a model.

The absence of strong feedback effects of the sleep-wake cycle on the circadian pacemaker seems reasonable from a functional point of view. The sleep-wake cycle may be considered as a "slave oscillator," which according to Pittendrigh (71) does not necessarily have the clocklike properties of self-sustained circadian oscillators. A strong feedback or coupling effect from the slave on the pacemaker would impair its precision in tracking environmental time (72). It seems to us that an organism's timing would be optimally served by a precise central timer, exerting flexible control over its slave processes, but being only slightly responsive to their action.

C. Sleep Structure

None of the current theories on the timing of sleep and wakefulness (30, 36, 55, 91) specifically considers the sleep stages and their typical temporal pattern across the night. The percentage of stages 3 and 4 (SWS) is known to be highest in the first NREM-REM sleep cycle and to decline over subsequent cycles (for references, see 18). This decreasing trend is a reflection of the progressive decline of the EEG power density in NREM sleep, a measure on which process S of the model was principally based. As suggested elsewhere (20), the conventional stages of NREM sleep (74) represent arbitrary subdivisions of an essentially continuous process. In contrast to SWS, the percentage of REM sleep shows an increasing trend over the sleep period. REM sleep episodes are characterized by a low EEG power density. They typically alternate with NREM sleep episodes at intervals averaging 108 min (Beersma, Daan, and van den Hoofdakker, unpublished research). A comprehensive model of sleep regulation should eventually account not only for the timing of sleep, but also for the opposite trends of NREM and REM sleep episodes as well as for their cyclic alternation.

The problem of the NREM-REM sleep cycle is beyond the scope of this paper and has been considered elsewhere (18). In the present context, it is relevant to examine the possibility of a circadian influence on REM sleep. It has been generally assumed that the REM sleep pattern is determined to a large extent by circadian factors (81, 82, 87), and hence the incidence of REM sleep has been used as a phase indicator of the circadian oscillator (28). Alternatively the REM sleep pattern may be solely determined by an intrinsic sleep process, such as the decay of S, which is entirely defined by the initial conditions, independent of circadian phase (the "Kharma" hypothesis; McCarley, personal communication). By elaborating the reciprocal interaction model for the generation of NREM-REM cycles (59), both McCarley (personal communication) and Beersma et al. (unpublished research) have found that both the typical and the pathological REM sleep patterns can be accounted for by the Kharma hypothesis.

On the other hand, it is difficult to explain the decreasing trend of REM sleep propensity during daytime hours, as demonstrated by nap studies (35, 58, 81), without assuming a circadian influence. In a previous, qualitative version of the model, one of us has proposed that both a sleep-dependent (S) and a circadian (C) process determines the occurrence of REM sleep (18). Therefore, when extending the model to incorporate REM sleep, we will have to examine whether the present version of the model is sufficient to account for the available data.

D. Sleep Deprivation in Affective Illness

There are numerous indications that sleep is impaired in mental disorders. Sleep has been studied in detail in endogenous depressives (43, 44). In view of the sleep disturbances associated with depression, it may seem paradoxical that sleep deprivation can exert a therapeutic effect (34, 49, 78). This suggests a causal link between sleep and depression. Various hypotheses have been advanced to account for this relationship. They have generally centered on the concept of internal desynchronization, postulating that biological rhythms vary in their mutual phase relationships. An explicit proposition was the phase-advance hypothesis, which has been developed in detail by Wehr and co-workers (83-85). The hypothesis states that in patients with endogenous depression, a number of circadian oscillations are phase advanced relative to the sleep-wake cycle, which itself is determined by the social environment. Sleep deprivation therapy, as well as advancing bedtime for several hours, would prevent sleep from occurring in a "wrong" (i.e., depressogenic) circadian phase. However, there is as yet no compelling evidence for an abnormal phase relationship of rhythms in depressive patients. In fact, the amplitude of circadian rhythms may be more affected than the phase (12).

Another explicit hypothesis has recently been advanced by Borbély and Wirz-Justice (23). To account for the pathognomic sleep changes as well as for the antidepressant effect of sleep deprivation, these authors have proposed that the S process of the present model is deficient in depressive patients. This hypothesis is open to experimental verification. Beersma et al. (unpublished research) have shown that differences in the REM sleep pattern between depressive patients and healthy subjects can be accounted for by assuming a low level of S at sleep onset.

In light of these considerations it is of interest to use the model to calculate the effect of various clinically used
sleep deprivation schedules on the level of process S. Figure 17 indicates the time course of S during total sleep deprivation, for three partial sleep deprivation schedules, and for base-line conditions. It is evident that the raised level of S following total sleep deprivation is most closely approximated by schedule 2 in which the subject goes to bed earlier and restricts his sleep period to the first part of the night. Limiting sleep to the second part of the night (schedules 3 and 4) gives rise to the levels of S that are only slightly higher than under base-line conditions. According to the hypothesis that the antidepressant effect of sleep deprivation is related to the increased level of S (23), the most effective therapy would consist of total sleep deprivation, whereas the least effective treatment would entail the restriction of sleep to the second part of the night. The few reports that have been published on this problem do not appear to contradict such a prediction (45, 69, 76, 77), although there are serious methodological problems (34). It is interesting to note that a raised level of S would also occur after the phase advance of bedtime by 6 h without sleep deprivation. Such a treatment schedule has been reported to induce a remission in depressives, an effect that has been attributed to the correction of a presumably abnormal phase relationship of circadian rhythms (84, 85). In conjunction with the considerations illustrated in Fig. 17, the present model may therefore provide an alternative explanation for the antidepressant effect of phase-advancing bedtime.

E. Sleep Function and Timing

We have tailored our model particularly to fit human sleep, which has been analyzed in more quantitative detail than animal sleep. However, there is no reason to assume that the main characteristics of the model are restricted to sleep in humans. Its preliminary versions were in fact established on the basis of animal data (14-17). In the rat, as in humans, the SWS fraction of NREM sleep gradually declines during the circadian inactivity period, whereas the REM sleep fraction of total sleep exhibits the opposite trend (21). Moreover, also in the rat, the decrease of the EEG power density in NREM sleep can be approximated by an exponential function (Borbély et al., unpublished work). Prolonged forced wakefulness gives rise to a massive increase in SWS recovery sleep. The important role of a circadian factor in sleep regulation in the rat was demonstrated by experiments in which a conflict was created between circadian and wake time dependent components of sleep propensity (21).

The application of the model to species other than humans may require only quantitative changes in parameter values. For instance, ultradian rhythms of activity and rest, modulated by circadian variation, are characteristic of many small mammals (5, 29) and may be generated by the model by changing only a single parameter. Lowering of H leads to an ultradian sleep-wake pattern (Fig. 18) quite reminiscent of the records obtained in rodents (in particular microtines). For quantitative simulation, further data on the timing and EEG of rodent sleep will be needed.

Thus we have reason to assume that the main features of the model correspond to basic principles of mammalian sleep regulation. The exponential time course of process S, in particular, seems ideally suited to mediate compensatory sleep responses in the intensity domain rather than in the time domain, allowing the animal to
remain in phase with its environment.

The concept of an intensity dimension in sleep (21, 38) has important implications for the existing theories of sleep function. These theories, summarized by Webb (80), comprise two controversial approaches. The classic view (e.g., 1, 46) is one of sleep as a recovery process of body and brain. More recently, ecotothetical theories have been advanced (60, 61, 79) according to which sleep serves to conserve energy and to reduce risk at times of day when wakefulness would be non-profitable or dangerous. These theories are defended mainly by the argument that the large variation in sleep time among species and among individuals is incompatible with a restorative function. Such a proposition is based on the premise of a constant rate of sleep. Behaviorally sleep is indeed a constant phenomenon, but physiologically and as a recovery process it may well have an intensity dimension.

If sleep is indeed primarily a repair process, the strategic decision an animal faces is when to carry out this recovery process. It may well have an intensity dimension. The large variation in sleep time among species and among individuals is incompatible with a restorative function. Such a proposition is based on the premise of a constant rate of sleep. Behaviorally sleep is indeed a constant phenomenon, but physiologically and as a recovery process it may well have an intensity dimension.

REFERENCES


Commentary on the article of Daan et al.

A CENTRAL ISSUE IN MODELING the human circadian timing system is that of essential model complexity; i.e., how many oscillators are necessary to account for observed rhythm phenomena. The debate is presently confounded by semantic confusion, which needs to be addressed before proceeding to comments on the model of Daan et al. and comparisons with other models.

Terminology

Daan et al. consider that the oscillator controlling the sleep-wake cycle in humans does not fulfill the requirements for being either a pacemaker or circadian “in view of its large frequency range and its sensitivity to external conditions” (sect. VI.A). They thus maintain that their model contains only a single circadian pacemaker (Abstract and sect. 6). Nevertheless the sleep-regulating process S is a “relaxation oscillation . . . supposedly reflected in the EEG power density” (sect. 5). The distinction implied is that a pacemaker exerts period and phase control over subsidiary oscillators and rhythmic processes; i.e., the suprachiasmatic nuclei of the hypothalamus (SCN) are postulated to be a pacemaker that controls the period of the S oscillation by regulating the oscillations in the upper and lower thresholds of the proposed somnostat. If we allow this definition, the model of Daan et al. involves one circadian pacemaker, but two oscillators in the regulation of the sleep-wake cycle, and introduces a third oscillator generating the body temperature rhythm in the absence of the SCN (sect. I.A).

Two types of experimental observation have traditionally been interpreted as indicating that the human circadian timing system contains (at least) two pacemakers (6, 8, 9). The first is the occurrence of spontaneous internal desynchronization between the sleep-wake and body temperature rhythms of subjects in temporal isolation. The second is that in primates some circadian rhythms are dependent on the integrity of the SCN, whereas others (body temperature and cerebrospinal fluid cortisol concentration) persist after total destruction of the SCN (1, 3, 7). Daan et al. relegate the mechanism(s) generating these persisting rhythms to the status of oscillators normally under the control of the SCN pacemaker. In contrast, we interpret these results as indicating the existence of another pacemaker outside the SCN.

Differences in Model Structure

In the model of Daan et al., the SCN are postulated to act as a pacemaker for both the body temperature rhythm and the oscillatory thresholds in the somnostat regulating sleep and wake. The periodicity of the thresholds is