Sleep and Sleep Stages Regulation

Validation of the S and C Components of the Three-Process Model of Alertness Regulation

*Torbjörn Åkerstedt and †Simon Folkard

*IPM and Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden;
and †Department of Psychology, University College of Swansea, Swansea, United Kingdom*

Summary: This paper summarizes work to validate and develop further the homeostatic and circadian component of a quantitative (computerized) three-process model for predicting alertness/sleepiness in daily living. The model uses sleep data as input and contains circadian and homeostatic components (amount of prior wake and amount of prior sleep), which are summed to yield predicted alertness on a scale between 1 and 16. The present validation was carried out using regression analysis, with sleepiness-related electroencephalographic parameters (alpha power density) from field and laboratory studies as criteria. The results showed that variations in alpha-power density in truck drivers, train drivers and laboratory subjects could be predicted with considerable accuracy ($r^2 > 0.70$) from the model, as could subjective alertness. Levels ≤ 7 on the 16-point scale were defined as critically low alertness. The paper also describes a simplified, graphic, paper version of the computation model, visualized as a two-dimensional "alertness nomogram". It is suggested that the studied components of the model may serve as tools for evaluating work/rest schedules in terms of sleep-related safety risks. **Key Words:** Sleepiness—EEG—Alertness—Model—Circadian—Homeostasis.

Sleepiness is a serious problem for individuals who have to displace sleep and/or work hours away from the conventional time. Severe subjective sleepiness and sleep-related electroencephalographic (EEG) changes characterize, for example, night-working truck drivers (1) and train drivers (2). In process operators, bouts of sleep actually occur during work (3). As may be expected, sleepiness at work also leads to accidents, including major events such as nuclear disasters (4).

In order to explain and predict variations in alertness related to irregular sleep/wake patterns, we have constructed a quantitative model that uses sleep/wake timing to predict alertness (5). The work was inspired by the two-process model of sleep regulation, which had shown that sleep length and slow-wave activity could be described by a combination of homeostatic (prior time awake or amount of prior sleep) and a circadian influence (6). Using subjective alertness data from a number of experiments of altered sleep/wake patterns, we found that alertness was also predictable from a

circadian and a homeostatic component in combination with a component for sleep inertia. The output of the model has been validated against subjective ratings, performance and electrooculogram (EOG) measures of sleepiness, and has shown considerable accuracy (5).

To use the model for practical purposes, however, further validation was needed. In particular, we needed to know whether the model would predict changes in objective indicators of alertness in individuals at work. Interest was focused on EEG alpha (8–12 Hz) power density, because this measure is highly related to subjective sleepiness (7) and performance failure (8). In addition, there was a need to establish criteria for interpretation of the predictions from the model, that is, to identify at what predicted levels EEG changes would occur and sleepiness be perceived. Finally, there was also need to establish the optimum acrophase setting (of the circadian component) for shift workers. For these purposes the present paper presents analyses of the correspondence between predictions from the model and the empirical results of a series of previously published laboratory and field studies. The sleep inertia component is not included here because of scarcity of field data with sufficient temporal resolution.

FIG. 1.

2

3
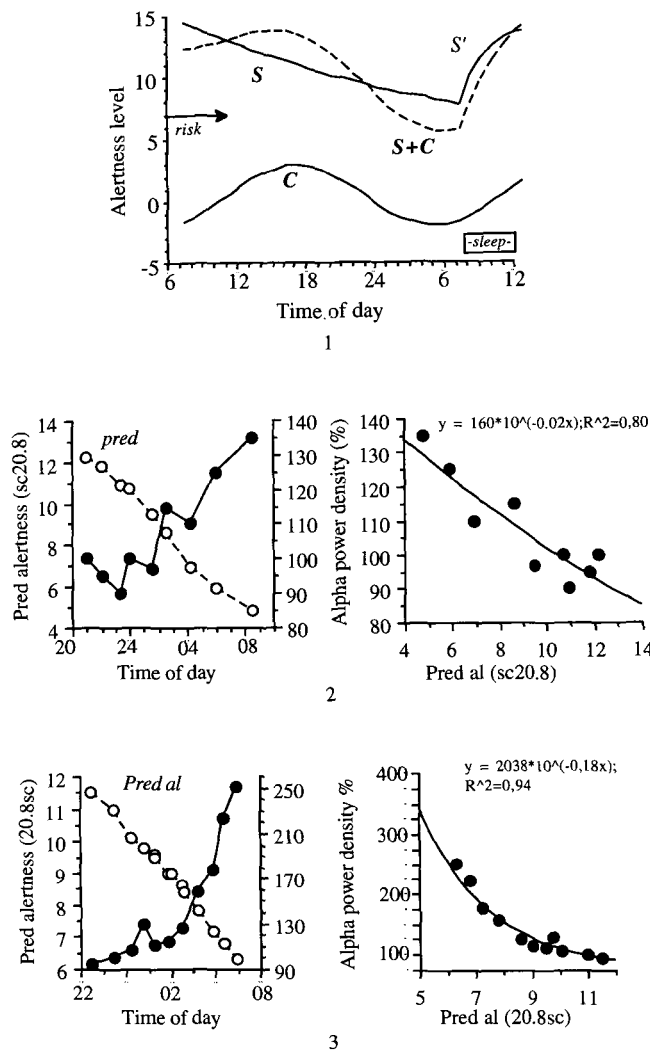
FIG. 1. Parameters of the three-process model of alertness regulation. S = homeostatic component during waking [S = (Sa − L)e ∧ − 0.0353t + L; where Sa = value of S (14) at awakening, L = lower asymptote (2.4) and t = time since awakening]. S′ = homeostatic component during sleep [S′ = U − (U − Sr)e ∧ − 0.381t; where Sr = value of S at retiring (7.96), U = upper asymptote (14.3) and t = time since falling asleep]. C = circadian component [C = M cos(t − p)π/12; where M = amplitude (2.5), p = acrophase (in decimal hours) and t = time of day (in decimal hours)]. S + C = the alertness prediction (excluding W). 7 = level of risk. The values in parentheses are the present default values of the model.

FIG. 2. Left: Mean alpha power density (filled circles) and predicted alertness plotted against time of day for truck drivers working at night. Right: Mean alpha power density plotted against predicted alertness for the same group. 100% = value during first 30 minutes.

FIG. 3. Left: Mean alpha power density (filled circles) and predicted alertness plotted against time of day for train drivers with night work. Right: Alpha power density plotted against predicted alertness for the same group. 100% = mean value for daytime drive.

METHODS

As mentioned above, the model was derived using subjective alertness data from a number of experiments of altered sleep/wake patterns. Modelling these

data we found that alertness was, indeed, predictable from three parameters: S, C and W (for details, see Fig. 1). **Process C** represents sleepiness due to circadian influences and has a sinusoidal form with an afternoon peak (Fig. 1). **Process S** is an exponential function of the time since awakening; it is high on awakening, falls rapidly initially and gradually approaches a lower asymptote. At sleep onset Process S is reversed and called **S′**; recovery occurs in an exponential fashion that initially increases very rapidly but subsequently levels off toward an upper asymptote. Total recovery is usually accomplished in 8 hours. The final component (not in Fig. 1) is the wakeup, **Process W** or sleep inertia. This component is not part of the present model and will not be further discussed. The input to the model is the times of rising and going to bed for the period investigated.

The predicted alertness is expressed as the arithmetic sum of the two (W presently excluded) functions above (S + C in Fig. 1). The scale of the model ranges normally from 1 to 16, but in practice 3 corresponds to extreme sleepiness, 14 to high alertness and 7 to a sleepiness threshold (EOG slow eye movements) (5). In Fig. 1 S + C shows predicted alertness when wakefulness is extended by 8 hours (to 24 hours), as is frequently the case, for example, with a first night shift. This particular prediction assumes that awakening occurs at 0700 hours in the morning after an 8-hour sleep period, and thereafter no sleep occurs (due to the night shift) until 0700 hours the following day. The combined effect of S + C (long time awake and the circadian downswing) yields a fall of alertness during the night, with a trough in the early morning. After sleep is started, the sleep recovery of factor S, together with the circadian upswing, causes a rapid increase in (latent) alertness during sleep.

The basic approach involves using the sleep data from the empirical studies as input to the model and comparing them with the model prediction through linear regression. The input used was the time of rising and prior bedtime before the period to be predicted, as well as the timing of any subsequent sleep taken during the period to be predicted. The parameters used for the model were the standard ones presented in Fig. 1, with the exception that the acrophase of C was varied from 1548 hours to 2148 hours to maximize the variance predicted in the dependent variable. Only the maximum solution is presented below.

Four studies were selected—two field studies and two laboratory studies. The EEG studies involved ambulatory measurement from standard derivations (see below). All involved fast Fourier transform analysis with a time window of 0.25 seconds, yielding intervals of 7.5 seconds and a sampling rate of 68 Hz. The resulting spectra were integrated across the delta (0.5–

3.9 Hz), theta (4–7.9 Hz), alpha (8–11.9 Hz) and beta (12–14.9 Hz) bands. Artifacts were removed after visual inspection. The results were expressed as a multiple or percentage of baseline levels (as in the original studies).

The studies of subjective sleepiness used the Karolinska Sleepiness Scale. The scale (7) is verbally anchored and has been validated against EEG and EOG parameters. It contains the steps 1–9, where 1 = very alert, 3 = alert, 5 = neither alert nor sleepy, 7 = sleepy but not fighting sleep, and 9 = very sleepy/fighting sleep/effort to keep awake. The intermediate steps do not have verbal anchors.

RESULTS

The first analysis used data from 15 truck drivers during a 500-km night drive between Malmö and Stockholm in southern Sweden (1). The drivers were male volunteers recruited from local trucking companies and had a mean age of 45 years. In connection with a night drive they would rise in the morning around 1030 hours after 8 hours of sleep, go about their private business during most of the day, and report to work at the depot in Malmö or Helsingborg around 1700 hours for an "early group" and 1900 hours for a "late group". After arriving they had electrodes applied for EEG (CzOz) and EOG (horizontal and oblique) recording. Shortly afterwards, the driver took charge of a ready-loaded truck and started his drive toward Stockholm on highway E4. The early group started at 1820 hours and the late group started at 2030 hours. Most of the route was either motorway or wide-lane highway. During the drive 1.5 breaks were taken (total time = 65 minutes) and the drivers arrived in Stockholm at 0407 hours and 0724 hours, respectively, making a total of 10–12.5 hours of work.

All EEG values were expressed as percent of the value during the first 30 minutes of driving. Alpha and theta power density increased significantly with duration of the drive (or lateness of the hour), as did subjective sleepiness. Figure 2 illustrates the results for the alpha band, combined for the two groups. It also contains the prediction from the model based on input from the mean group values for prior sleep and the nap during driving. Maximum R^2 for predicting alpha power density was reached for a phase of 2048 hours. The predicted alertness is approximately the mirror image of the EEG data. The close ($R^2 = 0.80$), slightly curvilinear (exponential) relation between the two is illustrated in the right part of the Fig. 2. Maximum alpha power density (30–40% increase) was obtained at predicted levels around 5.

A similar analysis was carried out with 11 male train drivers (mean age 42 years) during a 350-km night drive between Stockholm and Nässjö (2). The drivers were volunteers employed by Swedish Railways. Before the night drive they rose around 0900 hours after 8 hours of sleep, reported to work an hour before the drive, had EEG equipment (O_2P_4 and oblique EOG) applied, and started driving a train set at 1821 hours or 2121 hours (early and late group means, respectively). No naps were taken. The drive lasted 4.5 hours and included 5–6 stops along the route.

The results showed that the mean power spectral density in the alpha and theta bands increased significantly during the night, as did the amount of slow, rolling eye movements (a sleep onset indicator) and subjective sleepiness. The group mean data on prior sleep were used as input to the alertness model and predicted alertness was obtained, again with a maximum R^2 for an acrophase of 2048 hours. Figure 3 combines the alpha power density results for the early and late group and shows a marked increase, together with the decrease of predicted alertness. The regression analysis showed the best fit between the two to be exponential ($R^2 = 0.94$). This means that alpha power density shows very little change as alertness goes from high (12–14 units) to medium (9–11 units), but accelerates as it falls further (below 7 units). Maximum levels (140% increase) were observed at predicted alertness around 6.

In order to validate predictions also in a well-controlled laboratory situation, we chose a laboratory study of 28 hours of continuous activity, under isolation from external time cues, with a controlled intake of food and drink every 2 hours (7). Eight males between 25 and 45 years of age participated. The subjects rose from sleep at 0700 hours after an 8-hour sleep period, reported to the sleep laboratory at 1900 hours, had electrodes applied (CzOz) and started the recordings at 2200 hours. In 2-hour intervals until 1100 hours the next morning, the subjects carried out a drowsiness test. This involved sitting in a sound-insulated room with the instruction to focus on a mark 1 m in front of the eyes for 5 minutes. The EEG from this recording was analyzed in the same way as in the previous studies.

The results showed a strong, two- to three-fold increase in alpha and theta power density (as well as in slow eye movements and rated sleepiness) up to 0500 or 0700 hours. Thereafter a small reduction was seen. The timing of the prior sleep was input to the model, and the resulting prediction (maximum acrophase at 1948 hours) is approximately the opposite to that of the alpha power density (Fig. 4). The figure also shows a clear exponential relation ($R^2 = 0.95$) between alpha power density and predicted alertness. The exponential rise of alpha power density starts around predicted values of 7 and rapidly increases below that.
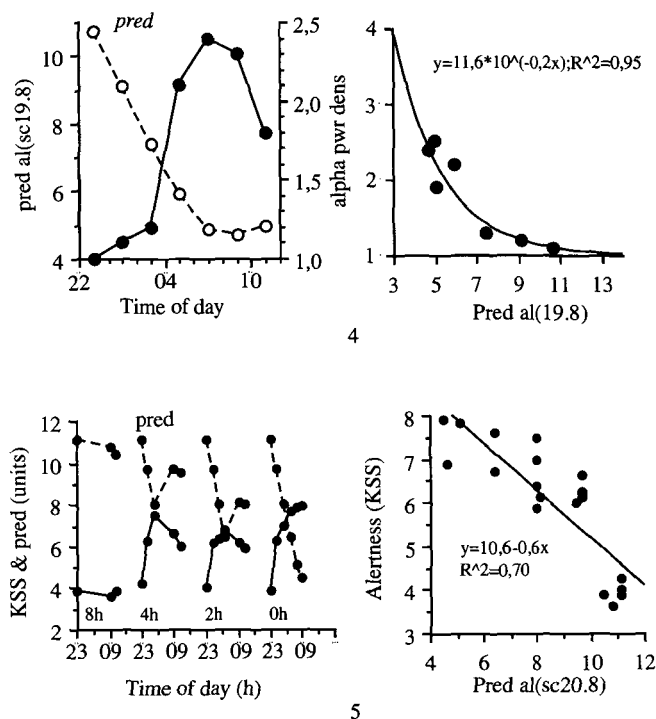
FIG. 4. Mean alpha power density (filled circles and continuous line) during 5-minute drowsiness test and predicted alertness plotted against time of day for subjects in a laboratory study of 28 hours of wakefulness. Right: Alpha power density plotted against predicted alertness for the same study. 1 = value during measurement at 2300 hours.

FIG. 5. Mean sleepiness (KSS) for four conditions (8, 4, 2, or 0 hours of sleep) of irregular sleep experiment, plotted with predicted alertness against time of day (left) and regressed on predicted alertness (right). Broken line = predicted alertness.

To obtain verbal descriptions of the levels of predicted alertness we used data from a laboratory study of irregular sleep with self-ratings of sleepiness performed every 2 hours using the Karolinska Sleepiness Scale (9). The eight male subjects slept in the laboratory before the experiment and were isolated from time cues (clocks, daylight) during the experiment. Four conditions occurred—all lasted for 28 hours from 0700 hours to 1100 hours the next day. They contained either no intervening sleep or 8, 4, or 2 hours of sleep (all terminated at 0700 hours). The sleep before each condition was a normal baseline sleep between 2300 hours and 0700 hours. Sleepiness ratings were made from 2300 hours until 1100 hours while the subjects were awake. Ratings immediately on awakening were discarded for the present purpose because of the sleep inertia effect. Figure 5 shows the raw data and predictions based on the sleep data input. Note that the number of measurement points decreases with the amount of intervening sleep. The variation across time was significant, except for the 8-hour condition. Ratings ranged from 3 or 4 (alert) in the late morning or evening

after a normal night of sleep to 8 in the late morning after a night of sleep loss.

For prediction of alertness the sleep data for each condition were entered into the model and the whole series of data was predicted (with a maximum acrophase at 2048 hours). This yielded an R^2 of 0.70 for the whole series, and the regression function became mean sleepiness (KSS) = 10.6 − 0.6x, where x = predicted sleepiness. (The prediction could be improved if each condition were predicted separately.) The regression function yields approximately the following pairs of data: high subjective sleepiness (a rating of 9, fighting sleep/an effort to keep awake) coincides with a predicted alertness level of 3–4. A KSS value of 7 (sleepy but not fighting sleep) corresponds to a predicted alertness of 7. Intermediate KSS (a rating of 5) corresponds to a predicted alertness of 10–11, whereas a KSS level of 2–3 (alert) corresponds to a predicted alertness of about 14.

DISCUSSION

Taken together, the validation data suggest that the alertness model is rather successful at predicting subjective and EEG-based sleepiness measures in situations of irregular sleep and waking, particularly when considering the number of factors outside the model that may influence alertness.

One weakness with the earlier version of the model was that the acrophase had not been sufficiently validated. To obtain information for setting suitable default values, we varied the acrophase between 1548 and 2148 hours. We found 2048 hours to produce consistently optimum prediction. This is 4 hours later than what was seen in the original derivation of the model (5). The reason for the discrepancy might be that the present studies mainly involved individuals on irregular sleep schedules, and it appears that such groups may be pronounced evening types, even before entering such systems (10). There is also the possibility of a tendency to free-run (which may yield a similar impression of a delay). In any case, we suggest that an acrophase between 2000 and 2100 hours be used as a default value when predicting shift-worker alertness. The original acrophase of 1648 hours should be retained for day workers pending further modelling.

Another weakness with the previous version of the model was the difficulty in interpreting the predictions in operational or similar terms. Some help in this respect may be gained from the analyses of the EEG data. Thus, medium to high predicted alertness (ratings of 8–14) seems to lack increased alpha activity, whereas the increase is sharp below 7—more than doubled in two of the studies. Such changes may be incompatible with adequate functioning in perception of signals

(8,11). Very similar results were observed for the EOG in the original validation studies (5). Thus, it seems that predicted values around 3–5 may be characterized by pronounced intrusions of sleep-related EEG and EOG changes even in the active individual; that is, there is a clear risk of behavioral malfunction.

The percentage increase of alpha activity was similar in the laboratory subjects and train drivers; predicted levels of 5–6 were associated with values 2–2.5 times baseline levels. The truck drivers showed much lower alpha levels, however. This could be due to the high demands on the truck drivers, penalizing even small attention lapses by increased risk. On the other hand such differences should be interpreted with caution because the recording techniques differed considerably between the studies.

With respect to the verbal meaning of predicted alertness, it seems that values around 3–5 were associated with very sleepy/fighting sleep/an effort to keep awake (KSS = 9). A predicted sleepiness of 7 corresponded to sleepy but not fighting sleep. Predicted values around 12 were associated with ratings of alert (KSS = 3).

Our tentative interpretation of the output of the model may be summarized in the following way (see also Fig. 6): Level 3 and below makes sleep intrusions highly likely even in the active individual, and this is perceived as difficulty staying awake. Around 7 the acute pressure of sleep is lowered and sleep is no longer perceived as difficult to resist. This may be regarded as a turning point. Above 10 sleep is unlikely and subjects tend to rate themselves as alert. Even if these interpretations have to be adjusted in the future, they may serve as guidelines for interpreting predictions in real-life situations.

It should be emphasized, however, that factors outside the model parameters will influence alertness greatly. One such factor is the ultradian rhythm, which causes a bimodal sleep pattern with an increased sleep tendency during the afternoon (12). This component will be added pending further work. Another influence is individual differences such as "diurnal type" (13,14), age and gender. Modifying the model for such influences will make it possible to predict individual data also. At present it should be emphasized that the model applies only to group mean data. Other influencing factors are monotony, stress, sleep conditions, drugs, health status, etc. These do not, however, belong in a general model of alertness and sleep regulation, but will have to be considered when interpreting output from the model.

The predictions from the model were obtained via an interactive computer program developed for Macintosh. However, to simplify and increase accessibility we have developed a "barefoot" version of it—an
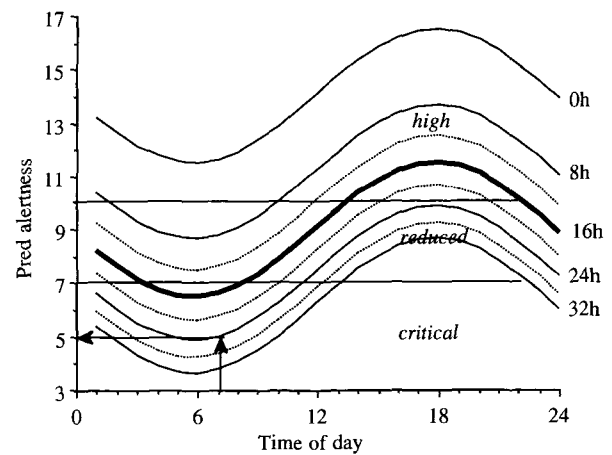

FIG. 6. Alertness nomogram. Each curve represents the circadian variation of alertness after a certain number of hours of prior waking. The arrows describe the use of the nomogram. Labels "critical" to "high" refer to verbal interpretation of predicted levels.

"alertness nomogram". It is envisaged that such an instrument may be used to train shift workers to optimize sleep/wake strategies or to guide administrators toward more physiological shift scheduling.

The nomogram was constructed by using the program (with C set to 1648 hours) to simulate sleepiness for each time of day and for 0–32 hours of prior time awake. The result is depicted in Fig. 6 and contains the predictions based on time of day and hours since the preceding major sleep period. Naps and shortened major sleep periods are not included yet. Each curve represents the circadian variation of alertness after a certain number of hours of prior waking. To use the nomogram, the subject first identifies the point in time (see arrows for example) for which an alertness estimate is desired. From this point a vertical line is drawn upwards to the curve that represents the number of hours of prior waking (for that point in time). From this intersection a horizontal line is drawn to the left and predicted alertness is read off on the y-axis. In the example, a prediction is sought for 0700 hours at the end of the first night shift. The point is thus preceded by approximately 24 hours of wakefulness (if no naps have occurred). The prediction lands on a value of 5, that is, very low alertness with clear risks of involuntary sleep onsets.

It is concluded that the homeostatic and circadian components of the three-process model of alertness are able to predict with considerable accuracy objective and subjective alertness changes due to altered sleep/wake patterns. Conceivably, these variables may constitute a tool for evaluating work/rest schedules in terms of sleep-related safety risks.

REFERENCES

1. Kecklund G, Åkerstedt T. Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics* 1993;36(9):1007–17.

2. Torsvall L, Åkerstedt T. Sleepiness on the job: continuously measured EEG changes in train drivers. *Electroencephalogr Clin Neurophysiol* 1987;66:501–11.

3. Torsvall L, Åkerstedt T, Gillander K, Knutsson A. Sleep on the night shift: 24-hour EEG monitoring of spontaneous sleep/wake behaviour. *Psychophysiology* 1989;26(3):352–8.

4. Mitler MM, Carskadon MA, Czeisler CA, Dement WC, Dinges DF, Graeber RC. Catastrophes, sleep and public policy. Consensus report. *Sleep* 1988;11:100–9.

5. Folkard S, Åkerstedt T. A three process model of the regulation of alertness and sleepiness. In: Ogilvie R, Broughton R, eds. *Sleep, arousal and performance: problems and promises.* Boston: Birkhäuser, 1991:11–26.

6. Borbély A. Sleep regulation: circadian rhythm and homeostasis. In: Ganten D, Pfaff D, eds. *Sleep. Clinical and experimental aspects.* Berlin: Springer Verlag, 1982:83–104.

7. Åkerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. *Int J Neuroscience* 1990;52:29–37.

8. Torsvall L, Åkerstedt T. Extreme sleepiness: quantification of EOG and spectral EEG parameters. *Int J Neurosci* 1988;38:435–41.

9. Åkerstedt T, Gillberg M. Sleep duration and the power spectral density of the EEG. *Electroencephalogr Clin Neurophysiol* 1986;64:119–22.

10. Knutsson A, Åkerstedt T. The healthy-worker effect: self-selection among Swedish shift workers. *Work Stress* 1992;6:163–7.

11. O'Hanlon JF, Beatty J. Concurrence of electroencephalographic and performance changes during a simulated radar watch and some implications for the arousal theory of vigilance. In: Mackie RR, ed. *Vigilance.* New York: Plenum Press, 1977:189–202.

12. Broughton RJ. Chronobiological aspects and models of sleep and napping. In: Dinges DF, Broughton RJ, eds. *Sleep and alertness: chronobiological, behavioral, and medical aspects of napping.* New York: Raven Press, 1989:71–98.

13. Horne JA, Östberg O. A self-assessment questionnaire to determine morningness–eveningness in human circadian rhythms. *Int J Chronobiol* 1976;4:77–110.

14. Torsvall L, Åkerstedt T. A diurnal type scale. *Scand J Work Environ Health* 1980;6:283–90.