J. Sleep Res. (2004) **13,** 373–383

Daytime variation in performance and tiredness/sleepiness ratings in patients with insomnia, narcolepsy, sleep apnea and normal controls

C. SCHNEIDER[1], S. FULDA[2] and H. SCHULZ[1]
[1]Clinic of Neurology, HELIOS Klinikum Erfurt, Erfurt, Germany and [2]Max Planck Institute of Psychiatry, Munich, Germany

Accepted in revised form 12 August 2004; received 9 March 2004

SUMMARY Daytime tiredness or sleepiness and deficits in cognitive performance are common complaints in sleep disordered patients. Till now there are few studies comparing patients from different diagnostic groups of sleep disorders in the same experimental protocol. We studied the time course of cognitive functions and subjective alertness in a parallel group design with four groups of patients [narcolepsy, untreated or treated obstructive sleep apnea (OSA), or psychophysiological insomnia] and a control group of subjects without sleep complaints. Each group consisted of 10 subjects, matched for age and gender. After a night with polysomnography, subjects were studied for 10 h from 08:00 hours to 18:00 hours at 20 min intervals under standardized environmental conditions. Four psychological tests were applied, (1) a critical flicker fusion (CFF) test to measure optical fusion threshold (alertness); (2) a paper-and-pencil visual line tracking test (selective attention); (3) a visual analog scale (VAS) for tiredness/sleepiness; and (4) the Tiredness Symptoms Scale (TSS), a 14 items check list. Each test session lasted for 8 min, followed by a 12 min pause. The level and time course of cognitive performance and self-rating data were analysed with hierarchical linear mixed effects models. Cognitive tests showed decrements in alertness and selective attention in untreated patients with insomnia, narcolepsy, and sleep apnea. Narcoleptic patients and untreated OSA had a lower CFF threshold than controls, and for narcoleptic patients the time course differed from that of all other groups. In the visual tracking test the performance of all groups of patients was worse compared with normal controls. Self-rated tiredness/sleepiness was significantly more pronounced in the three groups of untreated patients than in control subjects.

KEYWORDS alertness, cognitive performance, critical flicker fusion, insomnia, narcolepsy, obstructive sleep apnea, selective attention, sleepiness, tiredness

INTRODUCTION

Daytime sleepiness, tiredness and cognitive impairment are major consequences of various sleep disorders (Broughton and Broughton, 1994; Martikainen *et al*., 1992; Roth *et al*., 1995). Psychological studies have shown a rather consistent pattern of cognitive impairment for patients with narcolepsy and sleep related breathing disorders (SRBD) (Décary *et al*., 2000; Engleman and Joffe, 1999; Engleman *et al*., 2000; Naëgelé *et al*., 1995; Rieger *et al*., 2003; Valley and Broughton, 1981) while the existence of cognitive impairment is far less established for patients with insomnia (Edinger *et al*., 1997; Hauri, 1997; Riedel and Lichstein, 2000; Szelenberger and Niemcewicz, 2000). In a recent narrative meta-analysis on cognitive dysfunction in sleep disorders (Fulda and Schulz, 2001), it was shown that tests for selective attention and driving simulation display the most consistent impairment in

Correspondence: Dr Hartmut Schulz, Clinic of Neurology, Section Neuropsychology and Sleep, HELIOS Klinikum Erfurt, Nordhäuser Str. 74, 99089 Erfurt, Germany. Tel.: +49 361 781–6469; fax: +49 361 781-2132; e-mail: hschulz@erfurt.helios-kliniken.de

narcoleptic patients and in patients with SRBD. In the latter group of patients, improvement of performance by continuous positive airway pressure (CPAP) treatment was consistently found for some, but not all impaired cognitive functions (Bédard *et al*., 1991; Naëgelé *et al*., 1995). Results on cognitive dysfunction in patients with insomnia are less conclusive. Across different psychological functions, reduced performance was found in only one-fifth (22.9%) of all comparisons for patients with insomnia, while this number was well one-third for narcolepsy (34.6%) and SRBD (36.9%) (Fulda and Schulz, 2001).

The great majority of psychological studies applied a 'transversal' design, i.e., different functions were tested either once or with only a few repetitions, while longitudinal studies with multi-observations are rare in this area (Danker-Hopfe *et al*., 2001; Kraemer *et al*., 2000) with the exception of circadian studies (Folkard *et al*., 1993; Monk *et al*., 1997). For this reason, present knowledge on the time course of cognitive performance in patients with sleep disturbances is limited, although such investigations may be essential to estimate the relevance of any impairment for professional and other long-term activities, such as car driving.

The objective of the present study was to investigate daytime variations of performance in psychological tests in parallel with subjective estimates of sleepiness or tiredness in patients with insomnia, narcolepsy and sleep apnea to see, whether they differ in the level or time course of cognitive performance over a time-span of 10 h. A 10 h observational period was chosen to match temporal demands of a normal working day.

We applied two psychological tests measuring alertness and selective attention. Alertness was measured by the critical flicker fusion (CFF) test (Curran *et al*., 1990), and selective attention by a paper-and-pencil visual tracking test. This latter test is the 'visualization' (V) subtest of the Repetitive Psychometric Tests series (RPM-V) (Moran and Mefferd, 1959). In earlier studies we found that patients with narcolepsy and obstructive sleep apnea (OSAS) had a lowered optical fusion threshold in the CFF test, and performance of OSAS patients improved under treatment with CPAP (Schneider *et al*., 1999; Schulz and Wilde-Frenz, 1995; Schulz *et al*., 1992; Schulz *et al*., 1997). The RPM-V was selected to measure selective attention. This test can be administered repeatedly, since 20 parallel forms are available (Moran and Mefferd, 1959). Subjective sleepiness or tiredness was assessed with (i) a 10 cm visual analog scale (VAS) and (ii) the Tiredness Symptom Scale (TSS), a 14-items symptom list to measure tiredness (Schulz *et al*., 1991).

METHODS

Subjects

The study included four groups of patients with either narcolepsy (NAR; ICSD-R 347), obstructive sleep apnea syndrome (OSAS; ICSD-R 780.53-0), untreated or treated, or psychophysiological insomnia (INS; ICSD-R 307.42-0) according to the International Classification of Sleep Disorders (ICSD, 1991) and a control group of non-complaining normal sleepers. All subjects had a medical check up and a neurological examination as part of the routine diagnostic process in our sleep disorders clinic. Diagnostic decisions were based on anamnestic data, sleep questionnaires, medical examination and polysomnography (PSG). A Multiple Sleep Latency Test (MSLT) was performed for diagnostic purposes in all patients with probable narcolepsy and in other patients if indicated. Patients who fulfilled the ICSD criteria for OSAS were either newly diagnosed and still untreated (OSAu), or they were under steady-state treatment with nasal CPAP (OSAt). Normal control subjects (CON) had to have an unremarkable medical check-up, they should not complain on disturbed sleep, and should not have signs or symptoms of cardiorespiratory disturbances, periodic limb movements in sleep, or any other PSG indication for a sleep disturbance. The extend of daytime sleepiness was measured during the screening phase by the Epworth Sleepiness Scale (ESS; Johns, 1991). All subjects completed the Beck Depression Inventory (BDI; Beck *et al*., 1961, German version by Hautzinger *et al*., 1994) to rule out clinically significant psychiatric morbidity. Each group comprised 10 subjects, matched for age, sex and verbal intelligence [Mehrfachwahl-Wortschatz-Test, MWT-B (*Multiple Word Selection Test*); Lehrl, 1995].

Procedure

Subjects arrived in the late afternoon before the study day and stayed in the sleep unit until the end of the last test session on the next day. After arrival subjects were instructed about the procedures and a training test session was performed to familiarize them with the psychological tests. Later in the evening they were hooked up for a standard 8 h PSG recording with four EEGs (F1-A1, C3-A2, C4-A1, O2-A2), two EOGs, four EMGs (M. mentalis, M. masseter, M. tibialis anterior left and right side), ECG, three channels for respiratory signals (oro-nasal flow, respiratory effort from chest and abdomen), and oxygen saturation. PSG recording was performed with a computerized system (MKE Medizintechnik, Willroth, Germany). All recordings were digitally stored on magneto-optical disc and visually analysed according to the standard criteria (Rechtschaffen and Kales, 1968). In the morning after bathroom visit and breakfast, electrode resistance was controlled and electrodes replaced if needed to allow continuous recording with a reduced number of electrophysiological variables (four EEGs, two EOGs and two EMGs) throughout the test session. The results of the electrophysiological recordings will be reported separately.

The test session lasted for 10 h from 08:00 hours to 18:00 hours with test runs at 20 min intervals. During each test run four tasks were performed, the CFF test (3 min), the RPM-V (3 min), the TSS and the VAS. Each test run lasted for

about 8 min and was followed by a 12 min rest period. Subjects stayed during the whole 10 h in an artificially lighted and climatised room together with the experimenter. Between the test sessions subjects could read or talk with the experimenter. Light snacks and water or fruit tea were available throughout the day, but no regular meals.

Critical flicker fusion test

The CFF test measures the optical fusion threshold. The test was performed with an electronic device (ZAK, Simbach/Inn, Germany) which uses a forced choice paradigm, a modification of the method of constant stimuli (Aufdembrinke, 1982). The subject sits in front of a tube in which four light emitting diodes (LEDs) are arranged circular at 90° intervals. One of the LEDs emits a flickering light, while the other three LEDs emit constant light. The subject has to detect the flickering light and to indicate the critical LED by a pointer which can be shifted to any of the four LEDs by means of a movable bar on top of the viewing tube. During each test session a sequence of 50 stimuli was presented, each for 3 s. The task of the subject was to identify the critical stimulus and to respond by shifting the pointer into the area next to the critical stimulus within 3 s. After 3 s a new stimulus was presented automatically. The 50 stimuli per session were distributed in a ± 5 Hz range around the assumed threshold, with a total of five stimuli for each 1 Hz bin. Stimuli were presented in random order. Because the optical fusion threshold differs between subjects, the threshold was determined in a pretest session. This training session took place the day before the study day. From the CFF data two test scores were computed, one for the fusion threshold, the other for the stability of the threshold. The fusion *threshold* was defined as the frequency where 50% of the presented stimuli were detected correctly. The *stability* of the threshold is highest if all stimuli below the threshold are correctly identified and none above the threshold, while the stability is lowest if the detection curve is flat and hits and misses are evenly distributed around the assumed threshold. The stability score varies between 100% (high stability) and 0% (low stability). The CFF test was selected as a presumed measure of cortical activation (Bobon *et al*., 1982). The test is pharmaco-sensitive (Allain *et al*., 1992) and impaired by sedation (Maddock *et al*., 1993).

Repeated psychometric measures, subtest visualization (RPM-V)

The RPM was developed by Moran and Mefferd (1959) to measure Thurstone's factors of general ability. Visualization (V) is one of these factors. The RPM-V test is a paper-and-pencil test which consists of eight sets of 10 tangled lines. The starting points of the lines are numbered on the left side. The task of the subject is to track one line after the other by eye and to place the number of the line into the appropriate cell on the right. The test score is the number of cells correctly identified within 3 min.

Visual Analog Scale for tiredness/sleepiness

The Visual Analog Scale for tiredness/sleepiness (VAS-T/S) is a 10 cm VAS used to self-assess the current degree of tiredness or sleepiness. The scale was verbally anchored by 'extremely tired' (in German: 'extrem müde') on the left, and 'wide awake' (in German: 'munter, hellwach') on the right side. The subject had to indicate his or her current degree of tiredness/sleepiness by placing a tick mark on the 10 cm line. We will use throughout this paper the double terminology tiredness/sleepiness for the following reason. In the German language, the terms 'tired' and 'sleepy' both can describe a state of sleepiness, while the reverse use, i.e. to describe tiredness by the term sleepy is extremely unusual. As a consequence of this one-sided overlap in terminology, patients with excessive daytime sleepiness describe themselves either as tired or sleepy, while insomniac patients show a strong preference to describe their state rather as tired than sleepy. For this reason, the verbal anchors of the VAS fit to any patient, sleepy or tired.

Tiredness Symptoms Scale

The TSS is a 14-items checklist with physical and mental symptoms, associated with tiredness (Schulz *et al*., 1991). The TSS consists of an introductory remark (At this moment I notice) followed by 14 items: (1) heavy head, (2) sore eyes, (3) watering eyes, (4) heavy eyelids, (5) heavy legs, (6) general weakness, (7) feeling cold, (8) sensitivity to noise, (9) yawning, (10) loss of interest, (11) poor concentration, (12) irritability, (13) little desire to speak with others, (14) urge to move around. Every item has to be answered with 'yes' or 'no'. The total TSS score varies between 0 (no item confirmed) and 14 (all items confirmed). Since the subject is asked to describe his or her actual state, the TSS can be given repeatedly at short intervals. It has been shown that the TSS total score increases monotonously under conditions of short-term (one night) sleep deprivation in normal subjects (Bes *et al*., 1992).

Data analysis and statistical model

We applied linear mixed models that provide a comprehensive statistical framework to model time-dependent changes. They can incorporate different variance-covariance and error structures, and regard inter-individual differences in time-dependent changes. With this approach, measurements are assumed to be nested within an individual, with measurements being the level-1 and subjects the level-2 units. In our case there were 1.495 observations nested within 50 individuals (a total of five measurements from three subjects were missing).

Each of the five response variables (RPM-V, CFF threshold, CFF threshold stability, VAS tiredness/sleepiness, and TSS) was modeled separately. Model building procedures followed the recommendations by Bliese and Ployhart (2002) and Verbeke and Molenberghs (1997). In a preliminary analysis we regressed each response variable on group status alone ignoring any time related fluctuations to test whether groups

differed in overall mean level of performance (overall *F*-test followed by pairwise comparisons).

The extended statistical modeling was performed in four steps. We started with a preliminary mean structure that included polynomial time trends, group status (diagnosis) and the covariates age, gender, depression score and verbal intelligence. At this stage over-parameterization is deliberately favored to get consistent estimators of the covariance structure (Verbeke and Molenberghs, 1997). In a second step we selected the random effects for the polynomial time trends, thus allowing for inter-individual differences in time-dependent trajectories. The likelihood-ratio test was used to determine the significance of the random effects. In a third step the residual covariance structure was determined, regarding three different error structures. Beside a first-order autoregressive structure, time-dependent and group dependent heteroskedasticity was incorporated into the model and retained if the likelihood-ratio test indicated improved model fit. In a last step parameter reduction was intended. It was investigated whether the random effects were still needed in the model, and those fixed effects that did not significantly contribute to the overall model (*t*-test) were stepwise excluded. Hypothesis testing concerning the differences between the five experimental groups were tested with an overall *F*-test, followed by planned contrasts. All models were estimated with the restricted maximum likelihood (RML) which adjusts for the uncertainty of fixed effects and provides unbiased estimates of the variance components which makes the procedure attractive in a small sample situation. Data analysis was undertaken with *R* (Ihaka and Gentleman, 1996) and the *nlme* library in *R* (Pinheiro and Bates, 2000).

Several parameters, derived from the model building process, can be meaningfully interpreted. As described above, starting with group status as the only explanatory variable allows testing for differences in mean level of performance. The intercept of the model relates to performance at the beginning of the experimental period (08:00 hours) and the polynomial time parameters depict the pattern of performance over time.

Results of the model fitting procedure are given in Table 3, which shows the parameters and parameter estimations of the final model. To illustrate the relationship with the raw data, Fig. 1 shows for one performance test (RPM-V) the individual responses, the individual predicted responses by the final model and the mean predicted and observed responses. The predicted values are the Best Linear Unbiased Predictors (BLUPs) derived from the final model.

The study conforms to the Declaration of Helsinki and was approved by the Ethics Committee of the Chamber of Physicians of Thuringia.

RESULTS

Patients

All 50 subjects (26 males and 24 females) who agreed to participate and fulfilled inclusion criteria completed the study. The m/f ratio was 5/5 in all groups except NAR patients where it was 6/4. The overall mean age (±SD) was 52.7 ± 10.0 years, and there were only minor age differences between the groups. Mean BMI was higher in three of the patient groups (NAR: 30.9 ± 5.9, OSAu: 32.4 ± 4.1, OSAt: 31.5 ± 4.1) than in CON subjects (26.6 ± 1.8), while it was lower in the group of INS patients (24.3 ± 1.8). The mean ESS score was above the cut off score 10 for sleepiness in two groups (NAR: 19.1 ± 4.1 and OSAu: 14.8 ± 4.4). The proportion of subjects with an ESS score >10 was 1/10 for CON, 10/10 for NAR, 9/10 for OSAu, 1/10 for OSAt and 0/10 for INS, i.e., sleepiness was dominant in the two patient groups with NAR and OSAu. BDI scores were highest in the groups OSAu (11.0 ± 5.3) and INS (10.1 ± 5.6), but still well below the cut-off score of 16 for clinically relevant depression. Only three patients from the OSAu group and two from the INS group had a BDI score >16. Even in these five cases BDI scores reached a maximum score of 20 and none of the patients fulfilled criteria for clinically relevant depression. Finally, all groups had a mean verbal IQ above 100, with the lowest mean IQ in the OSAu group (106.9 ± 11.5) and the highest in the CON group (119.5 ± 11.8). Mean values and SDs for these variables are shown in Table 1. All subjects and patients were free of drugs acting on the central nervous system with the exception of anticataplectic medication in NAR patients.

Visual tracking test (RPM-V)

Regarding the overall mean level, CON subjects performed significantly better than all other groups (Table 2). NAR subjects performed significantly poorer than all other groups except OSAu patients. INS, OSAt and OSAu patients did not differ from each other. The group difference was already apparent at the first test session with a higher intercept for the CON subjects (Table 3). The intercept ($F_{4,43} = 3.10$, $P = 0.01$), the linear ($F_{4,1433} = 16.27$, $P < 0.001$), and the quadratic time component differed between groups ($F_{4,1245} = 11.66$, $P < 0.001$). With the exception of NAR patients, all other groups showed an initial increase in performance (positive linear component), which was less pronounced in OSAt patient (see parameter estimates Table 3). NAR patients did also differ from the four other groups in the quadratic time component.

RPM-V was modeled by a fourth order trend (*t*-test, all $P < 0.001$) with random effects for the intercept, the linear and the quadratic trend (likelihood-ratio tests, all $P < 0.001$) but not for the cubic and the fourth order trend. There was significant heteroskedasticity between groups (likelihood-ratio test, $P < 0.001$) and over time (likelihood-ratio test, $P < 0.001$). NAR patients exhibited the greatest variability, followed by untreated OSAu patients, controls, and treated OSAt patients with INS subjects showing the least variability. Additionally, responses became less variable over time (delta −0.006). The model fit further improved with a first order autoregressive error structure (Phi = 0.11, likelihood-ratio test, $P < 0.001$).

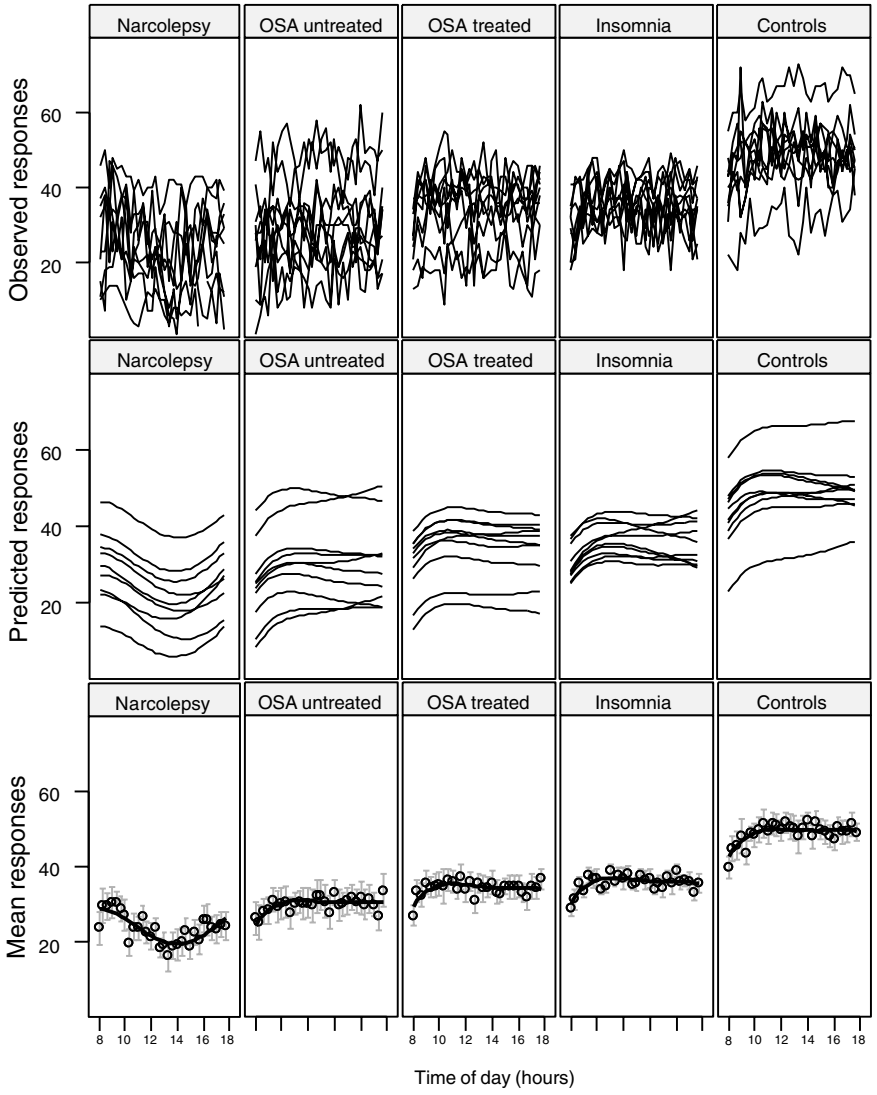

Figure 1. Observed and modeled responses for the RPM-visualization task. Observed (upper panel) and predicted (BLUPs, see text, middle panel) responses for each subject. In the lower panels are the averaged predicted (solid line) and observed (open circles, errors bars represent ±SEM) values for each group.

Age and verbal intelligence were related to RPM-V performance (Table 3). While increasing age was associated with lower performance, a higher score on the verbal intelligence test was related to a higher RPM-V performance. Fig. 1 shows the raw data, the predicted individual profiles and the mean predicted and observed responses. The absence of the linear trend at the beginning, as well as the pronounced midday decline of performance, is apparent in NAR patients but not in the other groups.

Table 1 Characteristics [mean (SD)] of the five study groups

Groups	*Age (years)*	*BMI*	*ESS*	*Verbal IQ*	*BDI*
CON	51.5 (9.5)	26.64 (1.76)	5.9 (3.5)	119.5 (11.8)	4.9 (3.6)
NAR	53.1 (11.6)	30.92 (5.87)	19.1 (4.1)	113.3 (13.9)	8.6 (4.6)
OSA_u	54.8 (8.0)	32.43 (4.05)	14.8 (4.4)	106.9 (11.5)	11.0 (5.3)
OSA_t	53.3 (11.8)	31.50 (4.09)	6.9 (3.6)	112.2 (12.2)	4.2 (3.0)
INS	50.7 (10.1)	24.30 (1.78)	4.5 (3.5)	117.5 (9.8)	10.1 (5.6)
Total	52.7 (10.0)	29.16 (4.86)	10.2 (6.8)	113.9 (12.2)	7.8 (5.2)

CON, Control group; NAR, narcolepsy; OSA_u, Obstructive Sleep Apnea Syndrome, untreated; OSA_t, Obstructive Sleep Apnea Syndrome, treated; INS, insomnia; SD, standard deviation; BMI, Body Mass Index; ESS, Epworth Sleepiness Scale; IQ, Intelligence Quotient; BDI, Beck Depression Inventory.

CFF threshold

The mean fusion threshold was significantly higher in group CON than in OSAu or NAR patient groups. No other differences in mean level threshold emerged between groups (Table 2).

The time course of the CFF threshold was best described by a cubic trend (*t*-test, all $P < 0.001$) with random effects for the intercept, the linear and the quadratic component (likelihood-ratio test, all $P < 0.001$). There was significant heteroskedasticity between groups (likelihood-ratio test, $P < 0.001$) with NAR patients having the greatest variability in performance followed by OSAu patients, CON, OSAt patients, and INS again exhibiting the least

Table 2 Overall mean level of performance

	CON (1)	*INS (2)*	*OSAt (3)*	*OSAu (4)*	*NAR (5)*	*Pairwise comparisons ($P_{F\ 1,\ 45} < 0.05$)*
RPM-V	48.7 ± 8.6	35.6 ± 4.5	34.1 ± 8.3	29.9 ± 10.8	23.4 ± 9.4	1 versus 2, 3, 4, 5; 2, 3 versus 1, 5; 5 versus 1, 2, 3
CFF threshold	27.9 ± 2.8	26.3 ± 1.5	25.6 ± 3.9	25.2 ± 1.7	24.1 ± 3.9	1 versus 4, 5
CFF stability	76.4 ± 7.0	74.8 ± 7.7	77.3 ± 6.7	72.2 ± 5.8	67.6 ± 6.1	5 versus 1, 2, 3
VAS	8.4 ± 1.3	4.7 ± 2.0	8.0 ± 1.9	5.5 ± 2.0	4.2 ± 2.3	1, 3 versus 2, 4, 5
TSS	1.0 ± 1.4	3.8 ± 2.3	1.4 ± 1.6	3.2 ± 2.0	4.6 ± 2.7	1, 3 versus 2, 4, 5

CON, control; INS, insomnia; OSAt, OSA treated; OSAu, OSA untreated; NAR, narcolepsy.
Values are mean ± SD. Pairwise comparisons were undertaken only if the overall *F*-test indicated significant differences between groups.

Table 3 Fitted models for the five response variables

RPM-V			*CFF threshold*			*CFF stability quotient*			*VAS tiredness/sleepiness*			*TSS*		
Effect		*Estimation (SE)*	*Effect*		*Estimation (SE)*	*Effect*		*Estimation (SE)*	*Effect*		*Estimation (SE)*	*Effect*		*Estimation (SE)*
Intercept			*Intercept*	β_1	26.67	*Intercept*	β_1	78.86	*Intercept*	β_1	5.70	*Intercept*	β_1	2.26
Age	β_1	−0.30			0.32			1.07			0.33			0.28
		0.11	Age	β_2	−0.18	*Time*			Depresssion	β_2	−0.24	Depression	β_2	0.23
Intelligence	β_2	0.26			0.03	C, IN	β_2	−0.195			0.05			0.05
		0.09	*Time*					0.046	*Time*			*Time*		
C	β_3	41.35	Depression	β_3	−0.014	O*t*	β_3	−0.367	Intelligence	β_3	−0.013	IN	β_3	0.079
		2.55			0.004			0.069			0.005			0.027
IN, O*t*, O*u*, N	β_4	28.52	N	β_4	−0.177	O*u*, N	β_4	−1.352	C, O*t*, O*u*	β_4	0.390	Ou	β_4	−0.500
		1.31			0.019			0.182			0.056			0.118
Time			*Time*^{2}			*Time*^{2}			IN	β_5	0.330	N	β_5	−0.225
C, IN, O*u*	β_5	1.797	Depression	β_5	0.0010	O*u*, N	β_5	0.0420			0.059			0.012
		0.173			0.0003			0.0062	*Time*^{2}			*Time*^{2}		
Ot	β_6	1.744	C, IN, O*t*, O*u*	β_6	−0.0053				Intelligence	β_6	0.0024	O*u*	β_6	0.0831
		0.179			0.0005						0.0008			0.0160
Time^2			*Time*^{3}						N	β_7	−0.0300	N	β_7	0.0791
C, IN, O*t*, O*u*	β_7	−0.1667	Depression	β_7	−0.00003						0.0062			0.0266
		0.0240			0.00001				*Time*^{3}			*Time*^{3}		
N	β_8	−0.1151	time^3	β_8	0.00014				Intelligence	β_8	−0.00013	O*u*, N	β_8	−0.00423
		0.0195			0.00002						0.00004			0.00084
Time^3									time^3	β_9	0.00164	*Time*^{4}		
time^3	β_9	0.00624									0.00042	O*u*, N	β_9	0.000069
		0.00135							*Time*^{4}					0.000014
Time^4									Intelligence	β_{10}	0.000002			
time^4	β_{10}	−0.000080									0.000001			
		0.000025							time^4	β_{11}	−0.000023			
											0.000008			
Random effects: Intercept, Time, Time^2			Random effects: Intercept, Time, Time^2			Random effects: Intercept, Time, Time^2			Random effects: Intercept, Time, Time^2			Random effects: Intercept, Time, Time^2		

C, Control; IN, Insomnia; O*t*, OSA treated; O*u*, OSA untreated; N, narcolepsy.

variability. Model fit also improved with a first order autoregressive error structure (Phi = 0.15, likelihood-ratio test, $P < 0.001$).

Both age and the depression score were related to the CFF threshold. Age was negatively related to the overall level of the CFF threshold and the depression score was related to all three time parameters (Table 3). Groups did not differ in regard to the intercept ($F_{4,43} = 1.23$, $P = 0.312$) or the cubic trend ($F_{4,1420} = 1.08$, $P = 0.360$) but in the linear ($F_{4,1439} = 4.84$, $P < 0.001$) and the quadratic component ($F_{4,1427} = 3.05$, $P = 0.02$). NAR patients showed a significant and negative linear trend but no quadratic time trend. The other four groups, which did not differ from each other, exhibited a negative quadratic time trend but no overall linear component (Table 2). The positive cubic component, however, was uniform for all groups.

CFF threshold stability

NAR patients had a less stable optical fusion threshold than the other groups (Table 2, Fig. 2b). The stability of the CFF threshold was best modeled by a quadratic trend (*t*-tests, all $P < 0.001$) with random effects for the intercept, the linear and the quadratic trend (likelihood-ratio tests, all $P < 0.01$). There was significant heteroskedasticity between groups (likelihood-ratio test, $P < 0.001$). Again, NAR patients showed the greatest variability, followed by OSAt patients, INS and OSAu patients, while CON subjects exhibited the least variability. The model fit further improved with a first order autoregressive error structure (Phi = 0.06, likelihood-ratio test, $P < 0.01$).

None of the covariates was related to the stability of the CFF threshold. Groups did not differ in the intercept ($F_{4,45} = 1.88$, $P = 0.13$) but in the linear ($F_{4,1439} = 8.27$, $P < 0.001$) and the quadratic component ($F_{4,1435} = 6.86$,

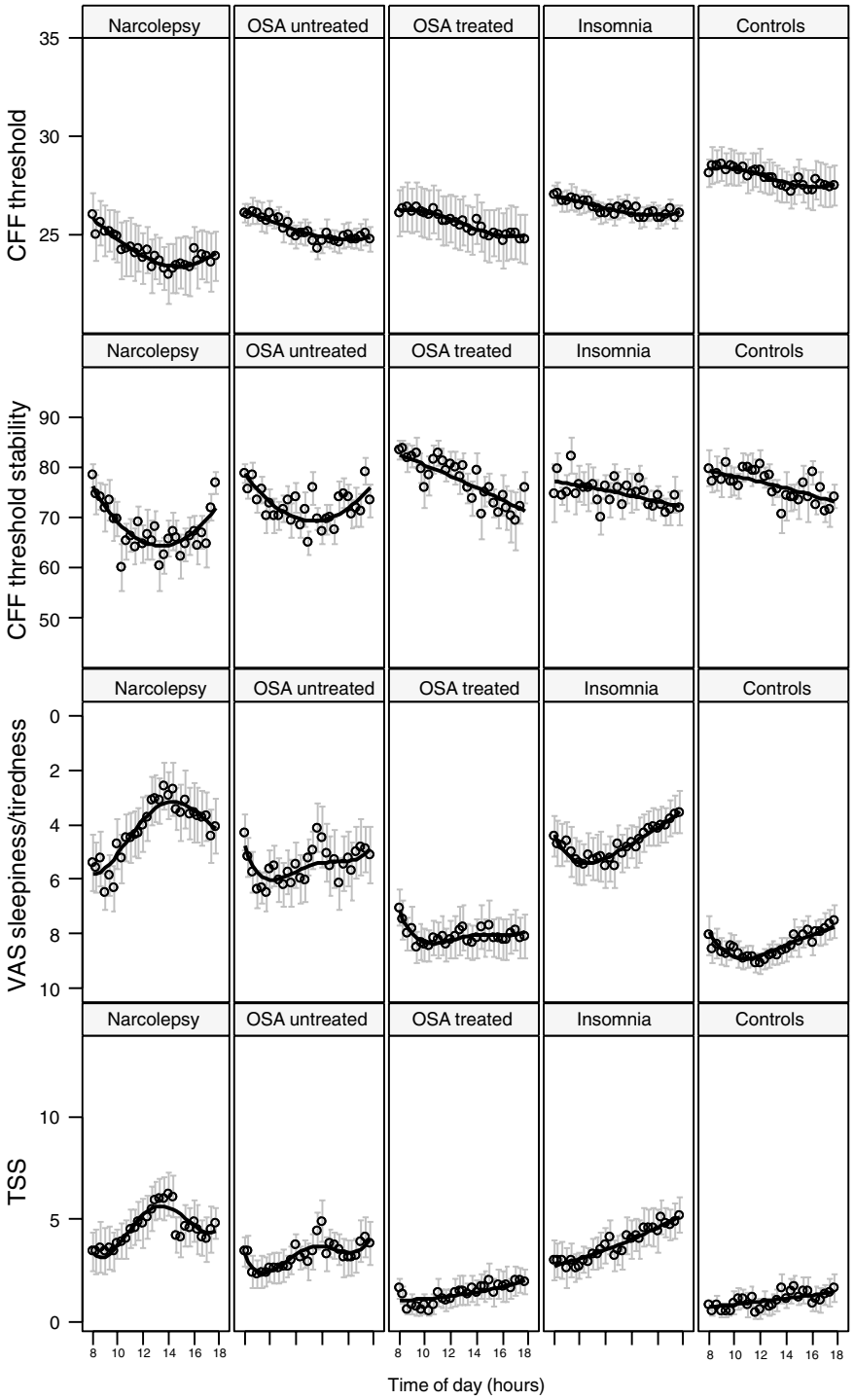

Figure 2. Averaged predicted (solid line) and observed (open circles, error bars represent ±SEM) values for each group for the CFF task (upper panel), the safety quotient of the CFF task (upper middle panel), VAS sleepiness/tiredness (lower middle panel), and TSS scale (lower panel).

$P < 0.001$). NAR patients and OSAu patients, which did not differ from each other, showed the most pronounced linear decline. This trend was less pronounced in OSAt patients and least pronounced in CON and INS patients. Whereas CON, INS and OSAt patients displayed only a linear decrease over time, NAR and OSAu subjects showed an additional quadratic trend with an increase in stability at the end of the testing period (Fig. 2b). However, from the five response variables the model for the CFF threshold stability captured the least variance of the data.

Visual Analog Scale for tiredness/sleepiness

Overall self-rated tiredness/sleepiness was lowest in CON and OSAt subjects, while INS, OSAu and NAR patients had significantly higher values and did not differ from each other (Table 2). Subjective sleepiness was best modeled by a fourth order trend (t-tests, all $P < 0.01$) with random effects for the intercept, the linear and the quadratic component (likelihood-ratio tests, all $P < 0.01$). There was significant heteroskedasticity between groups (likelihood-ratio test, $P < 0.001$). OSAu patients exhibited the largest variability, followed by NAR, OSAt and INS patients, and CON subjects. In addition, a first order autoregressive structure was included (Phi = 0.51, likelihood-ratio test, $P < 0.001$).

Sleepiness/tiredness depended also on the depression score with higher self-rated depression being associated with higher tiredness/sleepiness (Table 3). Verbal intelligence, on the other hand, was related to all four time trends. Interestingly, groups did not differ in the intercept ($F_{4,44} = 2.30$, $P = 0.07$) when the depression score was included in the model. Excluding this variable showed a pronounced difference between INS patients and all other groups, with INS patients rating themselves more tired than the other groups at the first test session. Group differences were apparent in the linear ($F_{4,1432} = 5.10$, $P < 0.001$) and the quadratic component ($F_{4,1426} = 4.03$, $P = 0.003$). All groups, with the exception of NAR patients, showed an initial positive linear trend (decreasing tiredness/sleepiness) which was most pronounced in CON, treated, and untreated OSA patients and less pronounced but still apparent in INS patients (Table 3, Fig. 2c). All groups showed a negative quadratic trend which was numerically smaller in NAR patients, but as with the RPM-V test, it was more obvious in NAR subjects as a result of the missing linear trend.

Tiredness Symptoms Scale

The mean number of perceived symptoms of tiredness, measured by the 14-items TSS, was very low for CON and OSAt subjects, while it was markedly higher in the other three patient groups (Table 2).

Response on the TSS was best modeled by a fourth order time trend (t-tests, all $P < 0.01$) with random effects for the intercept, the linear and the quadratic trend (likelihood-ratio tests, all $P < 0.001$). There was significant heteroskedasticity between groups (likelihood-ratio test, $P < 0.001$; OSAt > CON > NAR > OSAu > INS, in decreasing order). In addition, a first order autoregressive structure was included (Phi = 0.39, likelihood-ratio test, $P < 0.001$).

Again, the depression score was positively related to the TSS score, with higher depression scores matching more tiredness symptoms (Table 3). Groups differed in all four time components (linear: $F_{4,1436} = 6.57$, $P < 0.001$; quadratic: $F_{4,1429} = 8.77$, $P < 0.001$; cubic: $F_{4,1421} = 6.85$, $P < 0.001$; fourth order: $F_{4,1425} = 2.44$, $P = 0.04$). CON subjects and OSAt patients showed no significant change with time. In contrast, NAR patients and OSAu patients exhibited an initial negative linear trend (decreasing tiredness), while INS showed a consistent positive linear trend. Quadratic, cubic and fourth order trends became apparent only in NAR and OSAu patients (Fig. 2d). The TSS thus proved to be sensitive in discriminating the time course of self-rated sleepiness in the different groups.

Comparing groups

NAR patients displayed the most pronounced and consistent impairment in selective attention, alertness and subjective rating of tiredness/sleepiness. The dominant pattern was a curvilinear trend with a strong quadratic component. In all three measures of cognitive functioning performance decreased between 08:00 hours and about 14:00 hours, while thereafter it increased again (RPM-V, CFF threshold stability) or leveled off (CFF threshold). In parallel to the time course of performance measures, subjective ratings (VAS-T/S, TSS) showed an inverse time course with increasing tiredness/sleepiness from the morning to the early afternoon, followed by a decrease in the late afternoon hours. A comparable quadratic trend was seen in OSAu patients for CFF threshold stability, but not for the other measures. For OSAt patients the quadratic trend was transformed into a linear, decreasing trend in this test (cf. Fig. 2b). OSAt patients performed somewhat better than OSAu patients in all three cognitive functions and, in parallel, VAS-T/S and TSS ratings were lower in the former than in the latter group. Although these mean differences were systematic, there were large interindividual differences in the observational data and in the individual predicted responses.

INS patients performed on a significantly lower level than CON subjects in the RPM-V test for selective attention while they did not differ from the other patient groups. In addition, subjective ratings of tiredness/sleepiness for INS patients were as high as for NAR and OSAu patients, while OSAu and CON scored substantially lower (Fig. 2c, d).

DISCUSSION

The diagnostic groups differed in performance on two cognitive functions (alertness and selective attention) and in self-rated vigilance (tiredness/sleepiness). For most comparisons within group variability was highest in narcoleptic patients and lowest in the group of insomniac patients.

Intercept

At the beginning of the test session groups differed only in the selective attention task (RPM-V), where control subjects performed better than all patient groups. This result has to be replicated before one can conclude that this task, which measures concentration or selective attention, is especially sensitive to distinguish between normal sleepers and those with different forms of sleep disturbances, including insomnia. Performance in the RPM-V test depended also on age and intelligence. However, significant group differences emerged even after controlling for differences in age and intelligence.

The fact that the intercept was not significantly different for the CFF test or subjective measures (VAS-T/S, TSS) strongly suggests that group differences are not so pervasive that they can be reliably assessed by a single test run of a performance test or a rating scale. This may explain to some extent why results from earlier studies with cognitive performance measures in sleep-disturbed patients, especially patients with insomnia, were inconsistent or contradictory (see Hauri, 1997).

Overall level of performance

Narcoleptic patients differed from control subjects in the mean level of performance, regardless of the type of task. The same was true for untreated OSA patients, with the exception of the CFF stability quotient, a parameter which describes the quality of performance. In contrast to this, treated OSA patients and controls did not differ from each other in the mean level of performance, except for the RPM-V. This result is in line with the assumption of a cognitive enhancing effect of CPAP treatment (Naëgelé *et al*., 1998; Sateia, 2003).

Insomniac patients differed from controls on the RPM-V task but not the CFF threshold, a measure of alertness. The present findings add some new aspects to the scarce information on cognitive performance of insomniac patients in the literature. Earlier studies found no deficit of insomniac patients in tasks measuring auditory vigilance (Hauri, 1997; Sugerman *et al*., 1985), visual vigilance (Bonnet and Arand, 1995) or divided attention (Hauri, 1997) but decreased performance in simple and choice reaction time tasks (Hauri, 1997; Pedrosi *et al*., 1995). The present CFF findings would support the assumption that vigilance and alertness are not impaired in insomniac patients, while insomniacs performed on a lower level than normal sleeping control subjects in a speed test of selective attention (RPM-V) which requires sustained concentration for some minutes.

Although performance in the CFF test was lower in all patient groups compared with controls, these differences were not significant for the limited sample sizes of 10 patients per group. We were also unable to replicate the finding of a significant difference in CFF performance between OSAu and OSAt patients, which our group had reported earlier (Schulz *et al*., 1997). Reasons for this may be the small sample size and differences in statistical analysis. The present results confirm the known fact that CFF test performance is age dependent with a decrease of the optical fusion threshold with increasing age (Schulz *et al*., 1992).

Differences in mean level were most apparent for self-rated sleepiness/tiredness. While control subjects and treated OSA patients rated themselves as quite alert, insomniac, narcoleptic and untreated OSA patients were more sleepy or tired, and these latter three groups were indistinguishable one from the other.

Time course of performance

The time-course of performance yielded substantial differences between groups with some convergence across tasks. Most notably, NAR patients displayed a U-shaped pattern in all three cognitive measures and an inverted U-shaped curve for subjective sleepiness. Performance was highest early and late in the 10 h-session and lowest performance in the early afternoon. This pattern corresponds with Broughton's concept of a biphasic, circasemidian rhythm of sleep propensity (Broughton, 1994) with a major peak during subjective night and a secondary, less pronounced peak halfway between two consecutive sleep phases. Similar temporal distributions of vigilance states have also been observed in subjects isolated from external time cues and deprived of self-determined activities (Campbell, 1984) and in subjects confined to an ultrashort sleep-waking cycle (Lavie, 1986).

A comparable U-shaped performance pattern was seen in OSAu patients for the CFF threshold stability and to a lesser degree for the self-rated sleepiness/tiredness. This suggests that the appearance of a significant midday trough in performance in a repeated measurement design is limited to the sleepiest subjects.

The relationship between the time course of performance measures and tiredness/sleepiness ratings will be presented separately, since such an analysis has to take into account time lags and other potential effects which need adequate statistical analysis.

Methodological issues

Because we used a multi-observation repeated measurement design, psychological testing was restricted to two tasks only, presumably measuring selective attention (RPM-V) and alertness (optical fusion threshold, CFF). Both tests were short (3 min each) and given 30 times at equidistant intervals for a total observational time of 10 h. The choice of a 10-h observational period was arbitrary and should correspond to the temporal demands of a normal working day. We did not intend to assess circadian variations in performance, which would have required measurements for 24 h and, ideally, for multiple circadian cycles (Folkard *et al*., 1993; Monk *et al*., 1997). Problems associated with repeated measurements of psychological functions are training and adaptation effects. While such effects are presumably small for the CFF test, larger training effects can be expected for performance on the RPM-V, a speed test. We tried to minimize such effects (i) by

familiarization with the task the day before the test session and (ii) by using multiple parallel forms of the test (Moran *et al*., 1964). The results suggest that there was no training effect in the CFF test but a systematic one in the RPM-V test. This training effect was seen in all groups. Its duration was shortest in NAR patients, who showed an early decline in performance on this test. In the other four groups the performance gain because of training lasted 1–2 h, thereafter group means reached a steady state for the rest of the test session.

We decided to use a polynomial model with linear parameters to depict the time course of the performance measures. A non-linear model might have been a valid alternative as it is usually both, parsimonious and valid beyond the observed range of data. Thus, the present results do not allow generalizing beyond the time frame from 08:00 hours to 18:00 hours. The data modeling approach provided a suitable tool to explore the complex, time-structured observations. The specific design of the study, however, must be taken into consideration when interpreting the results. Foremost, performance was measured every 20 min with a 12-min break in-between testing sessions. This does not correspond strictly to the continuous performance curves represented in Figs 1 and 2. The 12-min breaks might well have allowed to restore performance to a higher level than would have been the case with uninterrupted performance. As many real-life tasks, e.g. driving, usually take more than a few minutes, it seems reasonable to assume that the time course of frequently measured performance depicted here might be a too optimistic estimate when compared with high demand real-life performance.

Other aspects which may limit the generalizability of the results are sample size and sample selection. We have studied five groups with 10 participants each. Although patients were approached consecutively, given the inclusion criteria, this selection process is semi-randomly at best. Together with the fact, that each diagnostic group consisted of 10 participants only, generalizations of the results should be treated with appropriate caution. The issue of the low number of participants is of special importance when testing differences between groups.

In summary, hierarchical linear mixed effects models offer a powerful tool to explore complex, time-structured data. The specific advantages of this flexible modeling approach could be even further exploited with the accumulation of empirical evidence and theoretical developments (Broughton, 1998; Van Dongen and Dinges, 2002). Results as the present ones may serve to refine models of repeated measurement performance.

ACKNOWLEDGEMENTS

The study was supported by a grant from the Deutsche Forschungsgemeinschaft, DFG (SCHU 640/4-1).

REFERENCES

Allain, H., Lieury, A., Brunet-Bourgin, F., Mirabaud, C., Trebon, P., Le Coz, F. and Gandon, J. M. Antidepressants and cognition: comparative effects of moclobemide, viloxazine and maprotiline. *Psychopharmacology (Berlin)*, 1992, 106(Suppl.): 56–61.

Aufdembrinke, B. The measurement of CFF: some methodological considerations. *Pharmacopsychiatry*, 1982, 15(Suppl. 1): 5–8.

Beck, A.T., Ward, C. H., Mendelson, M., Mock, J. and Erbaugh, J. An inventory for measuring depression. *Arch. Gen. Psychiatry*, 1961, 4: 561–571.

Bédard, M. A., Montplaisir, J., Richer, F., Rouleau, I. and Malo, J. Obstructive sleep apnea syndrome: pathogenesis of neuropsychological deficits. *J. Clin. Exp. Neuropsychol*., 1991, 13: 950–964.

Bes, F., Wilde-Frenz, J. and Schulz, H. Application of the Tiredness Symptoms Scale during sleep deprivation. In: Coenen, A. M. L. (Ed.) *Sleep-Wake Research in The Netherlands, Vol. 3*, Dutch Society for Sleep–Wake Research, Leiden, The Netherlands, pp. 19–20. 1992.

Bliese, P. D. and Ployhart, R. E. Growth modeling using random coefficient models: model building, testing and illustrations. *Organ. Res. Meth*., 2002, 5: 362–387.

Bobon, D. P., Lecoq, A., von Frenckell, R., Mormont, I., Lavergne, G. and Lottin, T. La fréquence critique de fusion visuelle en psychopathologie et en psychopharmacologie. Essai de synthèse de la littérature. *Acta Psychiatr. Belg*., 1982, 82: 7–112.

Bonnet, M. H. and Arand, D.L. 24-hour metabolic rate in insomniacs and matched normal sleepers. *Sleep*, 1995, 18: 581–588.

Broughton, R. Important underemphasized aspects of sleep onset. In: R. D. Ogilvie and J. R. Harsh (Eds) *Sleep Onset*. American Psychol. Assoc., Washington, DC, 1994: 19–35.

Broughton, R. J. SCN controlled circadian arousal and the afternoon "nap zone". *Sleep Res. Online*, 1998, 1: 166–178.

Broughton, W. A. and Broughton, R. J. Psychosocial impact of narcolepsy. *Sleep*, 1994, 17(Suppl.): S45–S49.

Campbell, S. S. Duration and placement of sleep in a ''disentrained'' environment. *Psychophysiology*, 1984, 21: 106–113.

Curran, S., Hindmarch, I., Wattis, J. P. and Shillingford, C. Critical Flicker Fusion in normal elderly subjects; a cross-sectional community study. *Curr. Psychol. Res. Rev*., 1990, 9: 25–34.

Danker-Hopfe, H., Kraemer, S., Dorn, H., Schmidt, A., Ehlert, I. and Herrmann, W. M. Time-of-day variations in different measurements of sleepiness (MSLT, pupillography, and SSS) and their interrelations. *Psychophysiology*, 2001, 38: 828–835.

Décary, A., Rouleau, I. and Montplaisir, J. Cognitive deficits associated with sleep apnea syndrome: A proposed neuropsychological test battery. *Sleep*, 2000, 23: 369–381.

Edinger, J. D., Fins, A. I., Sullivan, J. R. Jr, Marsh, G. R., Dailey, D. S., Hope, T. V., Young, M., Shaw, E., Carlson, D. and Vasilas, D. Do our methods lead to insomniacs' madness?: Daytime testing after laboratory and home-based polysomnographic studies. *Sleep*, 1997, 20: 1127–1134.

Engleman, H. and Joffe, D. Neuropsychological function in obstructive sleep apnoea. *Sleep Med. Rev*., 1999, 3: 59–78.

Engleman, H. M., Kingshott, R. N., Martin, S. E. and Douglas, N. J. Cognitive function in sleep apnea/hypopnea syndrome (SAHS). *Sleep*, 2000, 23(Suppl. 4): 102–108.

Folkard, S., Totterdell, P., Minors, D. and Waterhouse, J. Dissecting circadian performance rhythms: implications for shift work. *Ergonomics*, 1993, 36: 283–288.

Fulda, S. and Schulz, H. Cognitive dysfunction in sleep disorders. *Sleep Med. Rev*., 2001, 5: 423–445.

Hauri, P. Cognitive deficits in insomnia patients. *Acta neurol. belg*., 1997, 97: 113–117.

Hautzinger, M., Bailer, M., Worall, H. and Keller, F. *Beck-Depressions-Inventar (BDI)*. Testhandbuch. Verlag Hans Huber, Bern, 1994.

ICSD. *International Classification of Sleep Disorders. Diagnostic and Coding Manual.* Diagnostic Classification Steering Committee, Thorpy MJ, chairman. American Sleep Disorders Association, Rochester, MN, 1991.

Ihaka, R. and Gentleman, R. R. A language for data analysis and graphics. *J. Comput. Graph. Stat.*, 1996, 5: 299–314.

Johns, M.W. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep*, 1991, 14: 540–545.

Kraemer, S., Danker-Hopfe, H., Dorn, H., Schmidt, A., Ehlert, I. and Herrmann, W. M. Time-of-day variations of indicators of attention: performance, physiological parameters, and self-assessment of sleepiness. *Biol. Psychiatry*, 2000, 48: 1069–1080.

Lavie, P. Ultrashort sleep-waking schedule: III. Gates and "forbidden zones" for sleep. *Electroencephal. Clin. Neurophysiol.*, 1986, 63: 414–425.

Lehrl, S. *Mehrfachwahl-Wortschatz-Intelligenztest, MWT-B.* Spitta, Nürnberg, 1995.

Maddock, R. J., Casson, E. J., Lott, L. A., Carter, C. S. and Johnson, C. A. Benzodiazepine effects on flicker sensitivity: role of stimulus frequency and size. *Prog. Neuropsychopharmacol. Biol. Psychiatry*, 1993, 17: 955–970.

Martikainen, K., Urponen, H., Partinen, M., Hasan, J. and Vuori, J. Daytime sleepiness: a risk factor in community life. *Acta Neurol. Scand.*, 1992, 86: 337–341.

Monk, T., Buysse, D. J., Reynolds, C. F. III, Berga, S. L., Jarrett, D. B., Begley, A. E. and Kupfer, D. J. Circadian rhythms in human performance and mood under constant conditions. *J. Sleep Res.*, 1997, 6: 9–18.

Moran, L. J. and Mefferd, R. B. Repetitive psychometric measures. *Psychol. Rep.*, 1959, 5: 269–275.

Moran, L. J., Kimble, J. P. and Mefferd, R. B. Repetitive psychometric measures: equating alternative forms. *Psychol. Rep.*, 1964, 14: 335–338.

Naëgelé, B., Thouvard, V., Pépin, J. L., Lévy, P., Bonnet, C., Perret, J. E., Pellat, J. and Feuerstein, C. Deficits of cognitive executive functions in patients with sleep apnea syndrome. *Sleep*, 1995, 18: 43–52.

Naëgelé, B., Pepin, J. L., Lévy, P., Bonnet, C., Pellat, J. and Feuerstein, C. Cognitive executive dysfunction in patients with obstructive sleep apnea syndrome (OSAS) after CPAP treatment. *Sleep*, 1998, 21: 392–397.

Pedrosi, B., Roehrs, T. A., Rosenthal, L., Portier, J., Roth, T. Daytime function and benzodiazepine effects in insomniacs compared to normals. *Sleep Research*, 1995, 24: 48.

Pinheiro, J. C. and Bates, D. M. *Mixed-Effects Models in S and S-PLUS*. Springer, New York 2000.

Rechtschaffen, A. and Kales, A. A. *Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages in Human Subjects*. University of California, Brain Information Service, Los Angeles, 1968.

Riedel, B. W. and Lichstein, K. L. Insomnia and daytime functioning, Sleep Med. *Review*, 2000, 4: 277–298.

Rieger, M., Meyer, G., Gauggel, S. Attention deficits in patients with narcolepsy. *Sleep*, 2003, 26: 36–43.

Roth, T., Roehrs, T., Rosenthal, L. Hypersomnolence and neurocognitive performance in sleep apnea. *Curr. Opin. Pulm. Med.*, 1995, 6: 488–490.

Sateia, M. J. Impairment and quality of life in obstructive sleep apnea. *Clin. Chest Med.*, 2003, 24: 249–259.

Schneider, C., Walther, B. W. and Schulz, H. The effect of Modafinil on long-term vigilance performance and self-rated tiredness in patients with excessive sleepiness. *Sleep Res. Online*, 1999, 2(Suppl. 1): 166.

Schulz, H. and Wilde-Frenz, J. The disturbance of cognitive processes in narcolepsy. *J. Sleep Res.*, 1995, 4: 10–14.

Schulz, H., Volk, S. and Yassouridis, A. Measuring tiredness by symptoms. *Sleep Res.*, 1991, 20A: 515.

Schulz, H., Wilde-Frenz, J., Volk, S. and Geisler, P. Narcolepsy and the pathological aspects of multiple napping. In Stampi, C. (ed.) *Why We Nap. Evolution, Chronobiology and Function of Polyphasic and Ultrashort Sleep*. Birkhäuser, Boston, 1992: 258–270.

Schulz, H., Wilde-Frenz, J. and Grabietz-Kurfürst, U. Cognitive deficits in patients with daytime sleepiness. *Acta Neurol. Belg.*, 1997, 97: 108–112.

Sugerman, J. L., Stern, J. A., Walsh, J. K. Daytime alertness in subjective and objective insomnia: some preliminary findings. *Biol. Psychiatry*, 1985, 20: 741–750.

Szelenberger, W. and Niemcewicz, S. Severity of insomnia correlates with cognitive impairment. *Acta Neurobiol. (Warsz.)*, 2000, 60: 373.

Valley, V. and Broughton, R. Daytime performance deficits and physiological vigilance in untreated patients with narcolepsy-cataplexy compared to controls. *Rev. EEG Neurophysiol.*, 1981, 1: 133–139.

Van Dongen, H. P. A. and Dinges, D. F. Chronic partial sleep deprivation data point to a novel process regulating waking behavioral alertness. *J. Sleep Res.*, 2002, 11(Suppl. 1): 232.

Verbeke, G., Molenberghs, G. *Linear Mixed Models in Practice: An SAS-Oriented Approach*. Springer, New York, 1997.