

Sleep loss-related decrements in planning performance in healthy elderly depend on task difficulty

KATHARINA BLATTER, KLAUS OPWIS, MIRJAM MÜNCH, ANNA WIRZ-JUSTICE and CHRISTIAN CAJOCHEN

Centre for Chronobiology, University Psychiatric Clinics, Basel, Switzerland

Accepted in revised form 8 August 2005; received 30 July 2005

SUMMARY Prefrontal cortex (PFC)-related functions are particularly sensitive to sleep loss. However, their repeated examination is intricate because of methodological constraints such as practice effects and loss of novelty. We investigated to what extent the circadian timing system and the sleep homeostat influence PFC-related performance in differently difficult versions of a single task. Parallel versions of a planning task combined with a control group investigation were used to control for practice effects. Thirteen healthy volunteers (five women and eight men, range 57–74 years) completed a 40-h sleep deprivation (SD) and a 40-h multiple nap protocol (NAP) under constant routine conditions. Each participant performed 11 easy and 11 difficult task versions under either SD or NAP conditions. The cognitive and motor components of performance could be distinguished and analysed separately. Only by thoroughly controlling for superimposed secondary factors such as practice or sequence effects, could a significant influence of circadian timing and sleep pressure be clearly detected in planning performance in the more difficult, but not easier maze tasks. These results indicate that sleep loss-related decrements in planning performance depend on difficulty level, and that apparently insensitive tasks can turn out to be sensitive to sleep loss and circadian variation.

KEYWORDS circadian rhythm, cognitive performance, planning, prefrontal cortex, sleep deprivation.

INTRODUCTION

Cognitive functioning is strongly influenced by two interlinked systems, that of circadian timing and the sleep regulatory homeostat (for reviews see Cajochen *et al.*, 2004; Rogers *et al.*, 2003). Whereas the former regulates wake promoting mechanisms, the latter enhances the sleep drive with increasing duration of time awake. Both systems interact to determine the daily variations of sleepiness or alertness and consequently affect neurobehavioural functions.

To what extent neurobehavioural performance is influenced by these two systems depends to a certain degree on the characteristics of the task employed (e.g. assessed cognitive domain, function-related brain regions, task duration, method

of administration, measured variable and task difficulty) (Bonnet, 2000). Task difficulty has recently been investigated in a study comparing the effect of age and time of day on performance in two tasks of different complexity: the authors found a time of day effect on performance only in the less complex task (Bonnefond *et al.*, 2003). As the two tasks were not only differentially complex but also differed in their test constructs, it is not clear whether the authors tested task complexity *per se*. In an imaging study using functional magnetic resonance imaging, the effect of sleep deprivation (SD) on logical reasoning within different difficulty levels of a single test was examined (Drummond *et al.*, 2004). The cerebral compensatory response to SD was enhanced with increasing task complexity. Nevertheless the subjects' behavioural performance did not reflect this process. Likewise, no interference of task complexity with circadian rhythmicity was found in a study using controlled laboratory conditions (van Eekelen and Kerkhof, 2003).

Correspondence: Christian Cajochen, PhD, Centre for Chronobiology, University Psychiatric Clinics, Wilhelm Kleinstr. 27, CH-4025 Basel, Switzerland. Tel.: +41 61 325 53 18; fax: +41 61 325 55 77; e-mail: christian.cajochen@upkbs.ch

Since the initial description of the contribution of task specificity to the different circadian patterns of performance (Folkard *et al.*, 1983), a rather broad range of task-specific cognitive domains have been investigated. However, many of the tasks employed are related to attention and working memory and thus functionally associated with the prefrontal cortex (PFC) (Folkard *et al.*, 1993; Freivalds *et al.*, 1983; Gillooly *et al.*, 1990; Monk and Leng, 1986; Monk *et al.*, 1997). The main waking function of the human PFC is to enable generation and execution of novel goal-directed behaviour. Especially important is the aspect of novelty because with increased practice and routine the functional involvement of the PFC diminishes and shifts towards lower sensory and motor cortical regions (Fuster, 1997). PFC tests should therefore be novel, stimulating and interesting, as extensively demonstrated by a few authors (Harrison and Horne, 1998; Harrison *et al.*, 2000; Horne, 1993; Jones and Harrison, 2001). As repeated testing is essential in most study designs for circadian and sleep research, precisely this requirement of novelty constitutes a conceptual dilemma: with every repetition of a PFC-related test, its construct validity has to be questioned. One way to settle this problem is to test a sufficient number of subjects once only. This is unrealistic in the gold-standard chronobiology protocols such as constant routine and forced desynchrony protocols. A compromise is the use of parallel task versions for repeated testing. Still, repeated administration of a cognitive task leads to a practice effect, which can last up to 3 weeks and influences subsequent test performance (Jewett *et al.*, 2001). In order to estimate this additional effect, it is necessary to apply the same test procedure in a control group. However, even with varying stimulus material and with control of practice effects, the once novel behaviour becomes more routine. Hence PFC involvement diminishes. This approach is rather a methodical approximation than a truly alternative solution to test many people once only.

The aim of the present investigation was to test to what extent circadian timing and the homeostatic sleep system influence PFC-related performance in differentially difficult versions of one single task. We adopted the above described alternative with parallel task versions and a control group in order to meet the methodical constraints of frequent testing, and employed a paper and pencil-planning task in the manner of Porteus mazes (Porteus, 1965). Planning is a markedly PFC-dedicated behaviour, where automatism do not lead to optimal performance (Burgess, 1997). Maze test performance is significantly correlated with proficiency in a driving task (Sivak *et al.*, 1981). Therefore, its investigation is of high everyday relevance in the context of circadian variation and sleep-induced performance decrements. Furthermore, it is a practical task for repeated testing, as parallel versions can be created theoretically *ad libitum*. In addition, this test can be accomplished within a reasonable time frame, which is a consideration particularly relevant for protocols in which many tests are carried out.

The following questions were addressed:

- 1 Does the circadian timing system modulate planning performance?

- 2 To what extent does the homeostatic sleep system influence planning performance?
- 3 Are the observed effects similar for the two complexity levels of a single task?

METHODS

Test generation

Maze tracing task

The maze tracing task was conceptualized according to Porteus Mazes (Porteus, 1965) and to the Wechsler Intelligence Scale for Children (Wechsler, 1991). By means of 'Maze Maker' (<http://hereandabove.com/maze/mazeorig.form.html>, August 2005), 56 mazes were generated (size: 16 × 16 cm, 20 × 20 rows of 29 pixels width, wall width: 3 pixels; length factor: 8). In a test run with 28 subjects (undergraduate students and technicians from our unit; 50% women and 50% men, age range 21–61 years) the time to complete a maze successfully was measured. The degree of maze difficulty was determined by the averaged performance time for each maze. Thereafter a range of average performance ± standard deviation for all mazes was defined. Mazes with an averaged performance time outside this set range were discarded, resulting in 14 mazes being adopted.

Different difficulty levels

In order to have a sufficient number of mazes for both study blocks, we used these 14 originally selected mazes for the study block 1, and reused them in block 2 in a modified manner by rotating them 90° in counterclockwise direction and reversing the maze entrance and exit points (Fig. 1 upper and middle panel). This modification led to a significant increase in the number of arms branching-off near the entrance of the mazes (see below, and Fig. 1 lower panel). Based on the fact that with increased number of branching-offs the probability to take a blind alley increases, a maze is more difficult the more complex (with secondary and tertiary branching-offs) the arms are, that lead away from the correct trace, particularly when they are closely located to the entrance point of the maze where the exit point of the maze cannot yet be foreseen (as was generally the case for block 2 mazes). We statistically compared the degree of difficulty of block 1 and block 2 mazes by counting the number of primary arms leading off the correct trace and determining its median (Fig. 1 lower left panel). The median served as boundary value for the two different tracing directions. The number of arms below the median comprised the starting area of the original tracing direction of block 1, while the number of arms above the median comprised the starting area of block 2 (Fig. 1 lower right panel). Then the total number of branching-off points in the starting area of the mazes was counted for each block separately and compared with a *t*-test (block 1: 5.8 ± 3.7 versus block 2: 9.0 ± 4.2), yielding significantly more difficult mazes for block 2 ($P = 0.045$, two-sided).

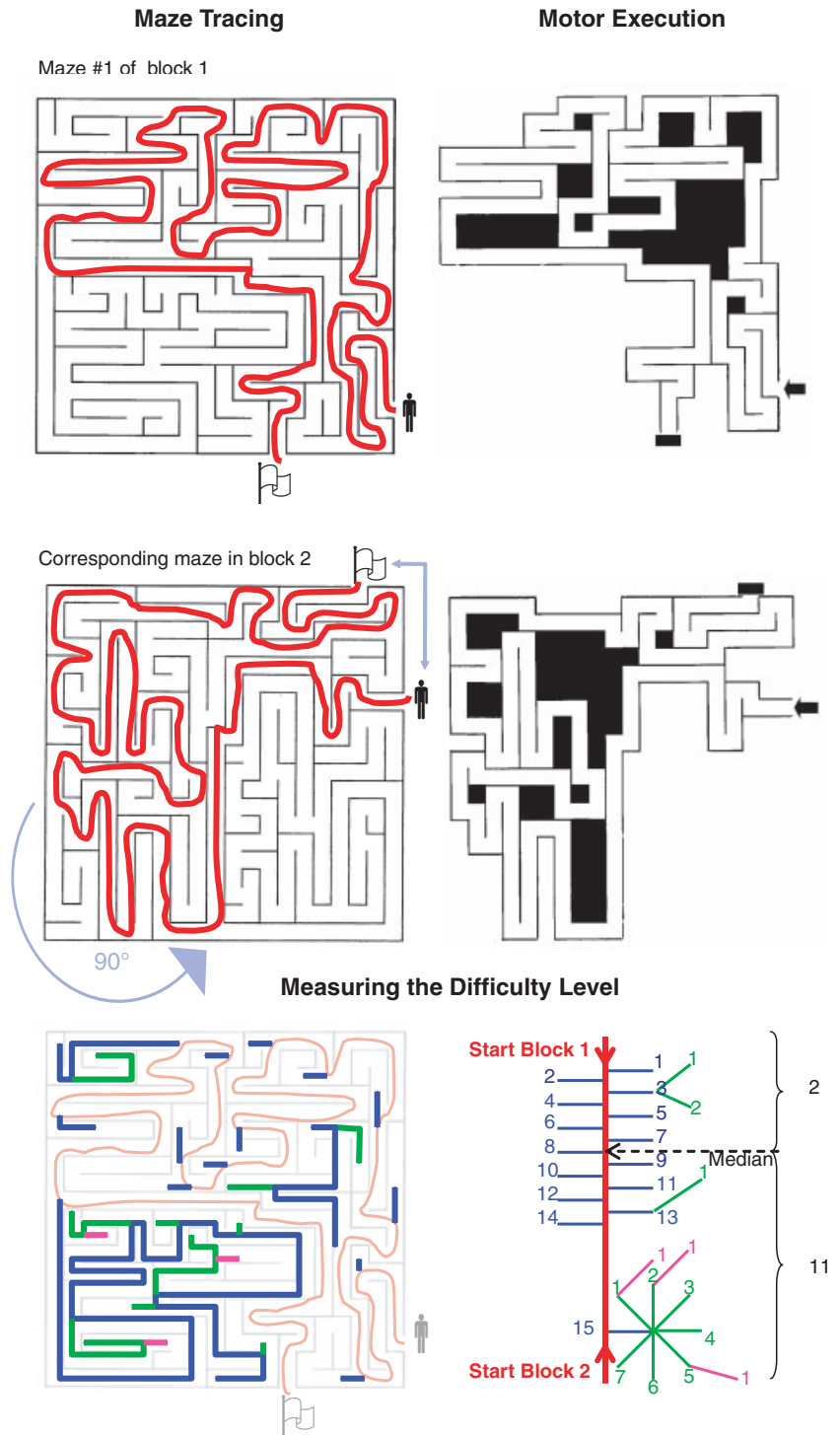


Figure 1. Maze Tracing Task. Upper panel: Example of an easy maze of block 1 (left) and the appropriate trace for the motor performance measurement (right). Middle panel: Maze from upper panel rotated 90° (light blue arrow) in counter-clockwise direction with starting points reassigned to exit points and vice versa, resulting in the more difficult maze for block 2. Lower panel: Example for the procedure to quantify the difficulty level of two maze versions. Arms leading away from the correct trace (left) are schematically displayed on the right. The arrow indicates the median number of the arms leading away from the correct trace serving as boundary value for the two different tracing directions. The number of arms below the median specifies the starting area of the original tracing direction of block 1 and the number of arms above the median specifies the starting area of the opposite tracing direction of block 2. The total number of secondary (green) and tertiary (pink) branching-off points in the starting area of each maze version (in this example 2 for block 1 and 11 for block 2) was counted with the aim to compare the mean number of branching-off points of all mazes per block (for statistics see Methods).

Scoring

The study participants were instructed to find the only correct trace as fast as possible without entering any blind alleys. In order to ensure that the participants truly planned, the briefing emphasized that accuracy was more important than speed. In general, there was no time limit, unless the participants took excessive time to find the exit. Performance time (from the moment on, when the subjects were in sight of the maze until

they finished tracing at the exit point) and number of errors (entering a blind alley) were recorded. Subsequently, the isolated correct trace of the respective maze had to be retraced by the participant on a second sheet of paper in order to measure motor execution time (Fig. 1 upper left and middle left panel). A grid of squared areas of the same dimensions as the mazes' path width allowed measuring the length of performed trace and of errors. Thus, an error and motor-corrected time measure could be calculated as follows:

Number of squares of covered distance yields the measure for performed trace length + error length (1)

overall performance time in s/(performed trace length + error length) = time in s per square (2)

time in s per square \times performed trace length
– motor execution in s = mental planning performance (3)

Study participants

Fourteen healthy older volunteers (six women, 65 ± 6.3 years of age, range 57–74, and eight men, 65.6 ± 5.7 years of age, range 57–73) participated in the study. The rationale for using elderly participants was to compare these healthy elderly with younger depressives patients in a large scale study, to test the concept that many physiological and maybe psychological factors in depression have similarities to ageing. The data presented here stem exclusively from the older healthy probands. Therefore the description does not include any comparative considerations. All participants were non-smokers, free from medical psychiatric, neurological and sleep disorders, and average chronotypes as assessed by screening questionnaires. Each participant went through a physical examination and a polysomnographically recorded screening night. Volunteers with more than 10 periodic leg movements per hour, an Apnea/Hypopnea Index higher than 10 and sleep efficiency lower than 80% were excluded. In order to ensure that none of the volunteers suffered motor, attentional or memory impairments, an additional neuropsychological assessment was carried out with *Motor Screening, Intra/Extradimensional Set Shifting, Pattern Recognition Memory* (CANTAB®, Cognition Ltd, Cambridge, UK) and the *Stroop Test*. Other exclusion criteria were: shift work within 3 months and transmeridian flights within 1 month prior to the study, excessive caffeine and alcohol consumption and excessive physical activity. All the study participants gave signed informed consent. The local Ethical Committee approved the

study protocol, screening questionnaires and consent form. All procedures conformed to the Declaration of Helsinki.

Protocol

The entire study consisted of two blocks of 5 days each separated by 2 weeks in between (the term ‘block’ refers always to the respective part of the study in the laboratory). During the week prior to each block the participants were instructed to maintain a regular sleep–wake cycle (bed- and wake-times within ± 30 min of self-selected target time), which was verified by a wrist activity monitor (Cambridge Neurotechnologies®, Cambridge, UK) and sleep logs.

In each study block the participants underwent one of two conditions in a balanced crossover design: SD and sleep satiation (NAP). A block consisted of one adaptation night (first night) and one baseline night (second night), followed by 40 h of either SD or sleep satiation, as well as one recovery night (Fig. 2). The sleep–wake schedules were calculated by centring the 8-h sleep episodes at the midpoint of each individual’s habitual sleep episode as assessed by actigraphy and sleep logs during the baseline week.

Sleep satiation was realized by 10 alternating cycles of 75 min of scheduled sleep (naps) and 150 min of scheduled wakefulness (Fig. 2), which in total comprised 40 h, the same time span the subjects spent in the SD protocol. In both protocols scheduled wake episodes were spent under constant routine conditions (constant dim light levels < 8 lx, semi-recumbent posture in bed, food and liquid intake at regular intervals, no time cues; for further details see Cajochen *et al.* 2001). During scheduled sleep episodes a minor shift to a supine posture was allowed, and the lights were off (0 lx).

Subjective sleepiness was rated every 30 min on the Karolinska Sleepiness Scale (KSS) (Gillberg *et al.*, 1994) and on a 100 mm visual analogue scale (VAS). Sustained attention was measured by the psychomotor vigilance task (PVT) (Dinges and Powell, 1985), which preceded each Maze Tracing Task by 60 min.

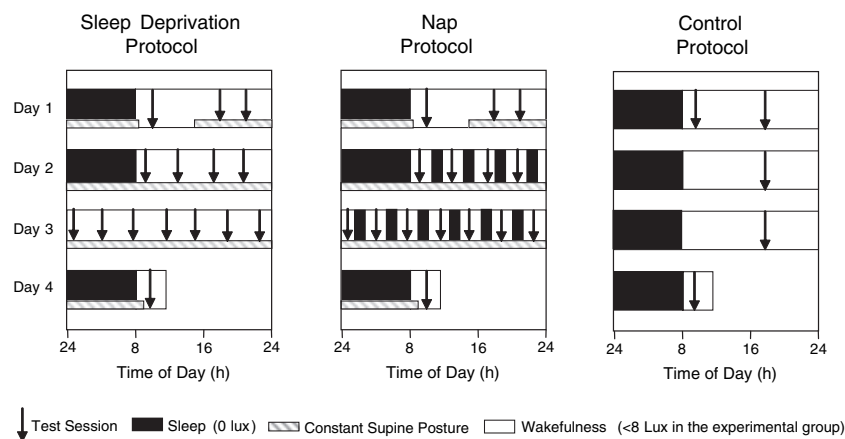


Figure 2. Overview of the protocol design. After two nights and a day in the laboratory, a 40-h sleep deprivation (left panel) or a short sleep–wake cycle paradigm (75/150 min) (middle) under constant routine conditions was carried out, followed by an 8-h recovery night. The control group was tested in an ambulatory setting under normal light and posture conditions (right panel).

Procedure

In the morning after the adaptation night, the subjects were tested with the 10 mazes of the Wechsler Intelligence Scale for Children (Wechsler, 1991), followed by the first two generated mazes in the afternoon at an interval of 4 h in order to familiarize the subjects with the task. Throughout the NAP protocol the remaining mazes were employed approximately 15–40 min before each nap and after the recovery night. During the SD protocol each maze was presented at times corresponding to those in the NAP condition.

Practice effect and ‘scaled effect subtraction’

The mazes’ degree of difficulty and potential practice effect were investigated in a separate control group. Five women (64 ± 3.2 years of age, range 61–68) and five men (63.8 ± 4.55 years of age, range 59–69) were recruited from qualified applicants for the laboratory study, who had been excluded solely because of hormone replacement therapy, varices or allergies. Ten ambulatory test sessions took place in two blocks of 5 days each, separated by 2 weeks in between. As in the experimental group the first block comprised the easy mazes and accordingly the second block the difficult ones. Timing was based on the original laboratory test protocol (Fig. 2). For the week before each test block the control group participants were instructed to maintain a regular sleep–wake cycle (verified by sleep logs).

We assumed that if a participant of the laboratory experiment (experimental group) had been in the control group, she or he would show similar performance (mental planning) as the control group’s mean performance. Hence, we scaled the mean control group performance curve by the ratio of the experimental group participant’s planning time for maze 3 to the mean time of the control group for the same maze (eqns 4 and 5, see below). Maze 3 was chosen, because it corresponds to test session in the experimental group at the beginning of the experimental conditions. The resulting curve served as a benchmark, simulating control group behaviour on the performance level of the experimental group participant. Each deviation of the experimental group participant’s performance curve from this benchmark curve could now be calculated by subtraction and attributed to an experimental effect (eqn 6):

$$\frac{\text{Experimental group-participant's performance in maze 3}}{\text{mean performance of the control group in the same maze}} \quad (4)$$

$$\text{Mean control group performance curve} \times (4) \\ = \text{scaled control group performance curve} \quad (5)$$

$$\text{Individual's performance curve} - (5) \\ = \text{net individual's performance curve} \quad (6)$$

Thus, this method of ‘scaled effect subtraction’ allows to discern the experimental effect from the mazes’ variation in

difficulty level, the differences between the two blocks, as well as the practice effect in the experimental group.

Testing for a possible interference effect

Sleep conditions were counterbalanced (see Protocol) but maze difficulties were not, as easy mazes were always employed in the first study block and difficult ones in the second study block. Referring to a possible interference effect, 10 additional subjects performed the differently difficult maze versions in a reversed order in an ambulatory setting: first the difficult versions and 2 weeks later the easy ones. The performance scores were error and motor corrected (see Procedure and Scoring). Analyses of variance for repeated measures (F_{ANOVA}) with the factor ‘difficulty level’ (difficult versus easy) and ‘session’ (1–11) revealed a significant difference between the two versions of mazes ($F_{1,9} = 53.9$; $P = 0.000$) with longer performance latencies in the difficult mazes. Therefore, we think that possible interference or order effects are negligible in our study. We can thus proceed from the assumption that block 2 mazes are indeed the difficult ones in contrast to block 1 mazes.

Statistical analysis

A F_{ANOVA} with the repeated factors ‘difficulty level’ (easy versus difficult mazes), ‘condition’ (NAP versus SD) and ‘session’ (sessions 4–14) were performed and are based on the multivariate test Pillai’s Trace. In two separate F_{ANOVAS} , one for each difficulty level, the factor ‘condition’ (NAP, SD) was used as between-subject factor. For the easy mazes the NAP condition contained data from six participants and the SD condition from seven participants, for the difficult ones *vice versa*. For *posthoc* comparisons the Duncan’s multiple range test was employed. In order to compare the time courses of subjective sleepiness (KSS, VAS) and of psychomotor vigilance performance (PVT) with maze tracing performance, dummy-coded linear regression analyses were conducted. The statistical packages SPSS® (SPSS for Windows, Version 11.0; SPSS Inc., Chicago, IL, USA) and Statistica® (STATISTICA for Windows, version 5.5, 1999; StatSoft Inc., Tulsa, OK, USA) were used.

The error rate was rather low in both the experimental and control group and therefore not further analysed. Gender differences were not analysed because the low number per subgroup (men: $n = 8$; women: $n = 5$) did not provide enough statistical power.

In the control group, one participant was excluded because of his outlying performance times (exceeding the range between 25th and 75th percentile). One participant of the experimental group was excluded because she showed a paradoxical time course in subjective sleepiness: in the NAP condition sleepiness ratings were above group average, whereas in the SD condition they were below group average.

RESULTS

Task difficulty

Figure 3 shows error and motor-corrected planning time for the control and the experimental group at both the task difficulty levels. In the control group a F_{ANOVA} with the repeated factors 'difficulty level' (easy and difficult) and 'session' (sessions 4–11) revealed significantly longer performance times for the difficult mazes ($F_{1,18} = 20.2$, $P < 0.001$). In the experimental group sleep conditions were counter-balanced (see Protocol) but maze difficulties were not (easy mazes always in study block 1 and difficult mazes always in study block 2). In order to test task difficulty in the experimental group the classification of factors in a F_{ANOVA} was therefore slightly different: repeated factors were 'condition' (NAP, SD) and 'session' (sessions 4–11). Additionally, the between subject factor 'order of condition' was introduced to code for difficulty level, as six participants had the easy mazes in the NAP condition and the difficult ones in the SD condition, and the other seven participants accordingly in reversed order. Thus, it was the significant interaction between the factors 'order of condition' and 'condition', which provided evidence that the performance times were significantly longer in the difficult mazes ($F_{1,11} = 15.1$, $P = 0.003$) (see also Table 1).

Motor performance

In order to test possible differences between the control and the experimental group regarding motor execution time of the maze trace, a F_{ANOVA} with 'session' as repeated factor and the between subject factor 'group' (experimental versus control group) was conducted. As the high sleep pressure in the SD protocol might cause psychomotor slowing, for the experimental group only data from the NAP condition was included. No significant differences between control and experimental group were obtained ($F_{1,29} = 1.6$, $P = 0.218$). Furthermore, to analyse motor performance in the experimental group only, a F_{ANOVA} with the repeated factors 'session' and 'condition' was conducted, but no significant effects were found except the main effect of 'session' (see also Table 1), which we expected because of the different path lengths of the mazes. This indicates that despite the fact of

differential sleep pressure condition, motor execution time in our maze task was not significantly different in the NAP versus SD protocol.

Planning performance

Planning performance in the experimental group was determined by two components: the experimental effect and a practice effect, whereas only the latter influenced task performance in the control group. In order to extract the experimental effect, the practice effect estimated from the control group was subtracted by means of 'scaled effect subtraction' (see Methods). The resulting measure was therefore described as the 'control group-corrected planning time'. Figure 4 illustrates the time course of this measure in the experimental group during the two conditions NAP and SD for both task difficulty levels. Visual inspection reveals different time courses for the low and high difficulty level. In the easier maze tasks of block 1, neither a circadian modulation nor any sleep pressure-dependent modulation could be detected (factors 'session', 'condition' or their interaction were not significant). In the more difficult tasks of block 2, performance times during the first 16 h into the protocols were not clearly distinct in the two conditions. Thereafter, during the biological night, a time-dependent increase of performance latency in the SD condition was manifest, superimposed on a circadian modulation. This was demonstrated by a F_{ANOVA} for the difficult mazes, where the 'session' \times 'condition' interaction was significant ($F_{10,2} = 20.6$ and $P = 0.047$) (see also Table 1). *Posthoc* analyses yielded significant differences only in the SD condition between the second (13:30 hours) and sixth (4:30 hours), as well as seventh (8:00 hours) measurement and also between each of the two latter designated ones and the fourth (21:00 hours) measurement respectively ($P < 0.05$, Duncan's multiple range test). A *t*-test with data pooled over day 1 and over day 2 revealed longer planning times in the SD protocol than in the NAP protocol ($P < 0.01$).

Sleepiness, reaction time and planning performance

Surprisingly, none of the subjective sleepiness scales (KSS, VAS) nor any of the PVT performance measures showed any predictive power in relation to planning performance in the

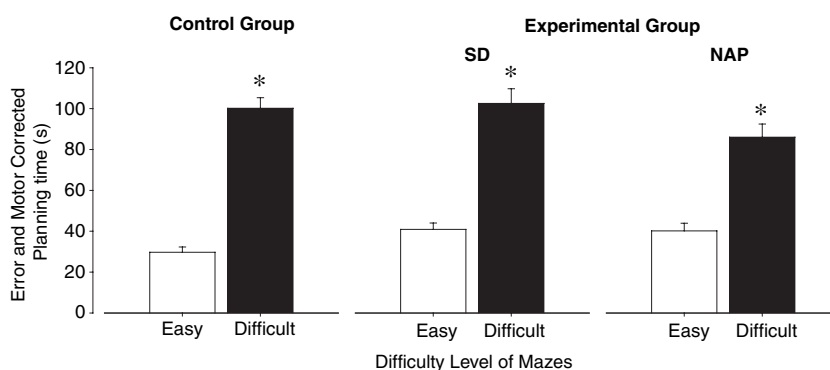


Figure 3. The different levels of task difficulty in block 1 and 2 exemplified by mean error and motor-corrected planning time in seconds for the control (left panel) and the experimental group (middle and right panel). Bars indicate SEM, asterisks indicate a significant difference between block 1 and block 2.

Table 1 Results of the FANOVAS of the measures 'error and motor corrected planning performance' of control group and experimental group, 'motor performance' of both these groups, and the 'control group adjusted planning performance' of the experimental group

Effect	F-value	d.f.	P-value
Task difficulty			
Control group*			
Session	1.9	10,9	0.168
Difficulty level	10.2	1,18	<0.001
Session × difficulty level	1.4	10,9	0.307
Experimental group			
Session	5.9	10,2	0.153
Condition	0.4	1,11	0.546
Order of condition	0.2	1,11	0.655
Session × condition	3.5	10,2	0.245
Session × order of condition	1.4	10,2	0.487
Condition × order of condition	15.1	1,11	0.003
Session × condition × order of condition	0.4	10,2	0.885
Motor performance			
Experimental and control group			
Session	14.8	10,20	<0.001
Group	1.6	1,29	0.218
Session × group	1.2	10,20	0.334
Experimental group			
Session	5.1	10,3	0.103
Condition	1.3	1,12	0.286
Session × condition	0.8	10,3	0.668
Planning Performance[†]			
Easy mazes			
Session	1.2	10,2	0.542
Condition	0.1	1,11	0.719
Session × condition	2.1	10,2	0.366
Difficult mazes			
Session	5.5	10,2	0.162
Condition	1.2	1,11	0.293
Session × condition	20.6	10,2	0.047

*The easy maze versions were always presented in the first study block and the difficult ones in the second study block. As the control group did not undergo experimental conditions, the task difficulty is manifested only in the variable 'difficulty level'. Whereas in the experimental group the difficult maze versions could be in the NAP or the SD condition, according to the classification of conditions to the respective study block.

[†]'Planning performance' comprises the control group adjusted mental planning time of the experimental group.

more difficult mazes in block 2 (regression coefficients all <0.1, predictive power had to be considered as negligible).

DISCUSSION

We aimed at testing to what extent the circadian timing system and the sleep homeostat influence PFC-related performance in differentially difficult versions of a single task under the carefully controlled conditions of a constant routine protocol. Eleven parallel versions of a paper and pencil-planning task were used and confounding practice effects were partialled out by means of an age-matched control group in an ambulatory setting. Circadian rhythmicity and sleep pressure significantly influenced planning performance only when the task was sufficiently difficult. In the difficult maze versions, volunteers

showed longer performance times in the SD condition than in the NAP condition. Planning time slowed down as soon as the waking period exceeded the normal daily amount of hours spent awake (16 h, at 1:00 hours), and was worst in the morning at 8:00 hours (circadian performance trough). There was no evidence for a pronounced diurnal modulation during the NAP condition, probably because of an outlier of a single participant at 21:00 hours on day 1. Assuming an average value at this point in time in Fig. 4, visual inspection reveals a clear diurnal modulation. In the easy mazes neither 'time of day' nor 'SD' affected planning performance.

We used control group-adjusted planning time as performance measure. Error rate was not a meaningful dimension, which has earlier been observed in self-paced as well as in short tasks (Bonnet, 2000). This has led to the assumption that during sleep loss individuals counteract imminent qualitative performance decrements by a more accurate and time-consuming processing (De Gennaro *et al.*, 2001; Dinges and Kribbs, 1991). By additionally assessing motor execution time, we were able to dissect out the cognitive aspects of performance, allowing statements about cognitive functioning under sleep loss without any confounding motor components. Although tracing movements became slower under SD, this observation was not statistically significant.

Our results of sleep loss and time of day effects in the more difficult maze versions are inconsistent with the results obtained by Drummond *et al.* (2004) and van Eekelen and Kerkhof (2003), who found no behavioural decrements in more complex versions of the logical reasoning task and the dual task respectively. The different cognitive domains of the reasoning (language), of the dual (working memory) and of the maze task (visuo-spatial) may have led to the dissimilar results. However, Drummond's finding of increased regional cerebral blood flow most likely reflected task difficulty and the effects of sleep loss. This leads us to interpret the lack of performance decrements in the above-mentioned studies as possibly due to an uncontrolled practice effect. Therefore, evaluating the learning curve in designs with repeated testing is of absolute importance, especially in the light of the often-formulated claim for a profile of task demands sensitive to sleep loss and circadian modulation. We also regard the results of Bonnefond *et al.* (2003) as a consequence of behavioural performance data confounded by a practice effect. They found deteriorated performance during the night only in a less complex task of visual discrimination but not in the more complex descending subtraction task. Visual discrimination has a natural limit in the perceptual system, whereas subtraction is a cognitive operation that may be trained to a greater extent. The complex task is susceptible to a practice effect, which – if not controlled – masks the time of day effect. Therefore, the fact that two different tasks were employed reduces the comparability with respect to task complexity.

Subjective sleepiness (KSS and VAS) and PVT performance (lapses and 10% slowest reaction times) did not show any predictive power for planning performance. This is in accordance with results attained by Owens *et al.* (1998), who found that although alertness was a good predictor for the circadian rhythm

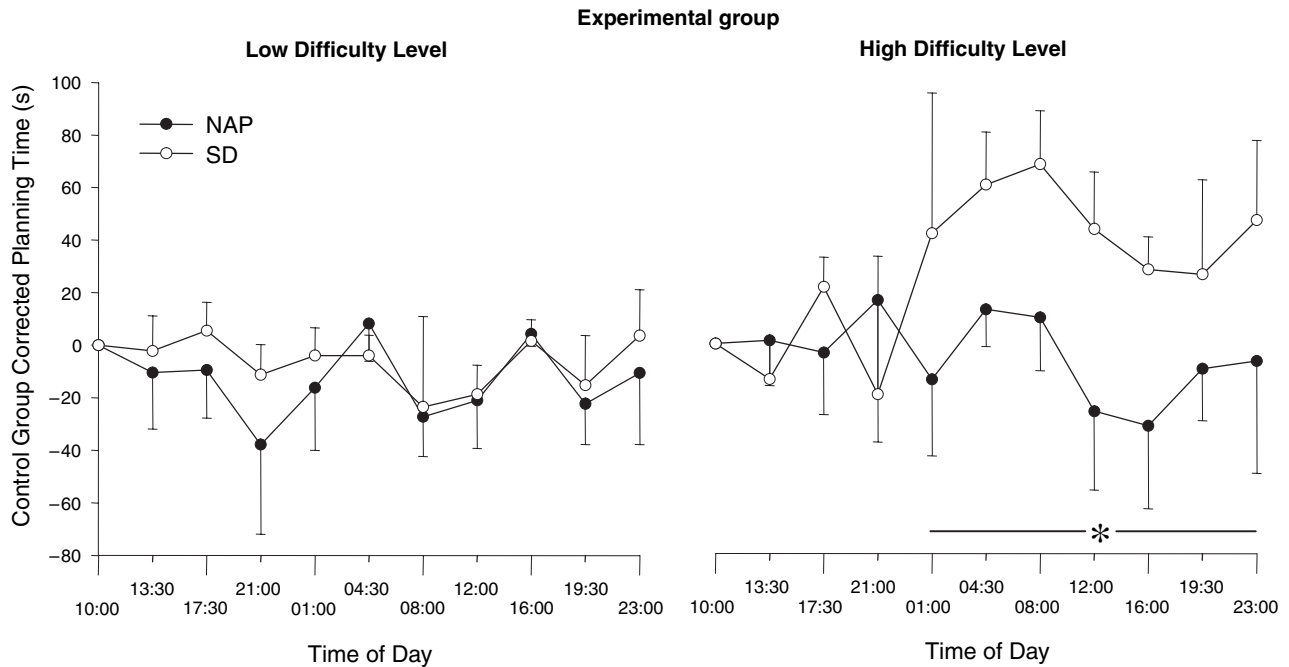


Figure 4. Maze tracing performance (mean control group adjusted planning time in seconds) in low (left panel) and high (right panel) difficulty task versions during sleep satiation (NAP) and sleep deprivation (SD). Only for the high difficulty level statistically longer performance times in the SD condition than in the NAP condition were present ($P < 0.05$). The line with asterisk indicates the significant difference between the two conditions on day 2 (t -test with data pooled for each day). Bars indicate SEM.

of simple perceptual motor speed, it could not be extrapolated to other performance measures. Atkinson and Reilly (1996) stated that cognitive performance is less sensitive to fatigue than subjective ratings. Moreover, the missing relationship between subjective ratings and cognitive performance points to the possibility described by Folkard *et al.* (1983), that each task performance might have its own circadian rhythm.

Considering the underlying mechanisms leading to behavioural decrements after sleep loss, we encounter attention as a global function. No matter which cognitive operation is carried out, it cannot be optimally processed if attention is dysfunctional. Therefore, attention is a prerequisite of any specific cognitive performance and covaries with increasing cognitive load (Stipacek *et al.*, 2003). Interestingly, this aspect has not yet been exhaustively scrutinized in chronobiological studies on performance, although sustained attention has been proven to be a very sensitive measure for sleep loss and circadian rhythmicity (Dinges *et al.*, 1987; Jewett *et al.*, 1999; Kribbs and Dinges, 1994; Monk *et al.*, 1997; Van Dongen *et al.*, 2003).

The study reported here was conducted with older volunteers. Following the appraisal of Anderson and Horne (2003), we assume that ageing effects within our sample are negligible, as the neuropsychological literature reports only minor age-related changes in healthy people within this relatively small age range. However, in the course of the age and memory debate the importance of optimal timing for performance assessments in the elderly has been stated (Hasher *et al.*, 1999). Several studies have concluded that older adults are more susceptible to time of day effects concerning memory performance than young study participants (Intons-Peterson *et al.*,

1999). If this is also true for PFC-related functions, we have to be aware that the effect we have demonstrated might be more articulated in the older sample than it would be in a younger one. However, there are indications that older participants who are aware of an age-dependent decrease in performance seem to be more motivated to deliberately counteract deficiencies than younger subjects (see also Bonnefond *et al.* 2003).

Our method of testing PFC-related planning function with parallel versions of a maze task as well as a control group allowed us to partial out the practice effect. We were able to distinguish between cognitive and motor slowing during SD and could show that a time of day and sleep pressure sensitivity is only manifested when the task is sufficiently difficult. Maze planning parallels driving proficiency, which makes it a function of high everyday relevance. As the maze-tracing task is a non-verbal test, it is rather independent of a person's educational background, and because of its uncomplicated administration it is easily applicable under constant routine and sleep laboratory conditions. Nevertheless, we have to keep in mind, that due to PFC-immanent characteristic traits, the explanatory power decreases with excessive testing. Our approach shows that apparently insensitive tasks can turn out to be sensitive to sleep loss and circadian variation, if thoroughly controlled for superimposed secondary effects such as practice or sequence effects.

ACKNOWLEDGEMENTS

We thank Claudia Renz, Marie-France Dattler, and Giovanni Balestrieri for their help in data acquisition, Dr Marc Blatter

for valuable contribution of mathematical knowledge, Martina Schacher for critical opinion on the task, Dr Carmen Schröder and Corina Schnitzler for the medical examinations, Christina Schmidt for having run the control experiment, and the subjects for participating. This research was supported by Swiss National Foundation Grants START No. 3130-054991.98 and No. 3100-055385.98 to CC, and the Velux Foundation (Switzerland).

REFERENCES

- Anderson, C. and Horne, J. A. Prefrontal cortex: links between low frequency delta EEG in sleep and neuropsychological performance in healthy, older people. *Psychophysiology* 2003, 40: 349–357.
- Atkinson, G. and Reilly, T. Circadian variation in sports performance. *Sports Med.* 1996, 21: 292–312.
- Bonnefond, A., Rohmer, O., Hoefl, A., Muzet, A. and Tassi, P. Interaction of age with time of day and mental load in different cognitive tasks. *Percept. Mot. Skills* 2003, 96: 1223–1236.
- Bonnet, M. H. Sleep deprivation. In: W. C. Dement (Ed.) *Principles and Practice of Sleep Medicine*, 3rd edn. Saunders, Philadelphia, PA, 2000: 53–71.
- Burgess, P. W. Theory and methodology in executive function research. In: P. Rabbitt (Ed.) *Methodology of Frontal and Executive Function*. Psychology Press, Hove, 1997: 81–116.
- Cajochen, C., Knoblauch, V., Kräuchi, K., Renz, C. and Wirz-Justice, A. Dynamics of frontal EEG activity, sleepiness and body temperature under high and low sleep pressure. *NeuroReport* 2001, 12: 2277–2281.
- Cajochen, C., Blatter, K. and Wallach, D. Circadian and sleep-wake dependent impact on neurobehavioral function. *Psychol. Belg.* 2004, 44: 59–80.
- De Gennaro, L., Ferrara, M., Curcio, G. and Bertini, M. Visual search performance across 40 h of continuous wakefulness: measures of speed and accuracy and relation with oculomotor performance. *Physiol. Behav.* 2001, 74: 197–204.
- Dinges, D. F. and Kribbs, N. B. Performing while sleepy: effects of experimentally induced sleepiness. In: T. H. Monk (Ed.) *Sleep, Sleepiness and Performance*. John Wiley & Sons, Chichester, 1991, 97–123.
- Dinges, D. F. and Powell, J. W. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behav. Res. Methods Instrum. Comput.* 1985, 17: 625–655.
- Dinges, D. F., Orne, M. T., Whitehouse, W. G. and Orne, E. C. Temporal placement of a nap for alertness: contributions of circadian phase and prior wakefulness. *Sleep* 1987, 10: 313–329.
- Drummond, S. P., Brown, G. G., Salamat, J. S. and Gillin, J. C. Increasing task difficulty facilitates the cerebral compensatory response to total sleep deprivation. *Sleep* 2004, 27: 445–451.
- van Eekelen, P. J. and Kerkhof, G. A. No interference of task complexity with circadian rhythmicity in a constant routine protocol. *Ergonomics* 2003, 46: 1578–1593.
- Folkard, S., Wever, R. A. and Wildgruber, C. M. Multi-oscillatory control of circadian rhythms in human performance. *Nature* 1983, 305: 223–226.
- Folkard, S., Totterdell, P., Minors, D. and Waterhouse, J. Dissecting circadian performance rhythms: implications for shiftwork. *Ergonomics* 1993, 36: 283–288.
- Freivalds, A., Chaffin, D. B. and Langolf, G. D. Quantification of human performance circadian rhythms. *Am. Ind. Hyg. Assoc. J.* 1983, 44: 643–648.
- Fuster, J. *The Prefrontal Cortex*, 3rd edn. Lippincott-Raven, Philadelphia, PA, 1997.
- Gillberg, M., Kecklund, G. and Åkerstedt, T. Relations between performance and subjective ratings of sleepiness during a night awake. *Sleep* 1994, 17: 236–241.
- Gillooly, P. B., Smolensky, M. H., Albright, D. L., Hsi, B. and Thorne, D. R. Circadian variation in human performance evaluated by the Walter Reed performance assessment battery. *Chronobiol. Int.* 1990, 7: 143–153.
- Harrison, Y. and Horne, J. A. Sleep loss impairs short and novel language tasks having a prefrontal focus. *J. Sleep Res.* 1998, 7: 95–100.
- Harrison, Y., Horne, J. A. and Rothwell, A. Prefrontal neuropsychological effects of sleep deprivation in young adults – a model for healthy aging? *Sleep* 2000, 23: 1067–1073.
- Hasher, L., Zacks, R. T. and Rahhal, T. A. Timing, instructions, and inhibitory control: some missing factors in the age and memory debate. *Gerontology* 1999, 45: 355–357.
- Horne, J. A. Human sleep, sleep loss and behaviour. Implications for the prefrontal cortex and psychiatric disorder. *Br. J. Psychiatry* 1993, 162: 413–419.
- Intons-Peterson, M. J., Rocchi, P., West, T., McLellan, K. and Hackney, A. Age, testing at preferred or nonpreferred times (testing optimality), and false memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 1999, 25: 23–40.
- Jewett, M. E., Dijk, D. J., Kronauer, R. E. and Dinges, D. F. Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. *Sleep* 1999, 22: 171–179.
- Jewett, M. E., Wright, K. P., Duffy, J. F., Rodriguez, D. M. and Czeisler, C. A. Practice effects observed over a month-long 28-hour forced desynchrony protocol in a cognitive throughput task are well described by a saturating exponential function. *Sleep* 2001, 24(Abtract Supplement):A4.
- Jones, K. and Harrison, Y. Frontal lobe function, sleep loss and fragmented sleep. *Sleep Med. Rev.* 2001, 5: 463–475.
- Kribbs, N. B. and Dinges, D. F. Vigilance decrement and sleepiness. In: R. Ogilvie (Ed.) *Sleep Onset: Normal and Abnormal Processes*. American Psychological Association, Washington, 1994: 113–125.
- Monk, T. H. and Leng, V. C. Interactions between inter-individual and inter-task differences in the diurnal variation of human performance. *Chronobiol. Int.* 1986, 3: 171–177.
- Monk, T. H., 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.
- Owens, D. S., Macdonald, I., Tucker, P., Sytnik, N., Minors, D., Waterhouse, J., Totterdell, P. and Folkard, S. Diurnal trends in mood and performance do not all parallel alertness. *Scand. J. Work Environ. Health* 1998, 24(Suppl. 3):109–114.
- Porteus, S. D. *Porteus Maze Test. Fifty years' application*. Psychological Corporation, New York, 1965.
- Rogers, N. L., Dorrian, J. and Dinges, D. F. Sleep, waking and neurobehavioural performance. *Front. Biosci.* 2003, 8: s1056–1067.
- Sivak, M., Olson, P. L., Kewman, D. G., Won, H. and Henson, D. L. Driving and perceptual/cognitive skills: behavioral consequences of brain damage. *Arch. Phys. Med. Rehabil.* 1981, 62: 476–483.
- Stipacek, A., Grabner, R. H., Neuper, C., Fink, A. and Neubauer, A. C. Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neurosci. Lett.* 2003, 353: 193–196.
- Van Dongen, H. P. A., Maislin, G., Mullington, J. M. and Dinges, D. F. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep* 2003, 26: 117–126.
- Wechsler, D. *WISC-III: Wechsler Intelligence Scale for Children*, 3rd edn. The Psychological Corporation, San Antonio, TX, 1991.