SLEEP DEPRIVATION EFFECTS ON COGNITIVE PERFORMANCE

Hoermann, H.-J.
Steinmann, M., Uken, T. & Elmenhorst, E.-M.
German Aerospace Center (DLR)
Institute of Aerospace Medicine
Hamburg and Cologne, Germany

An experimental sleep deprivation study has been launched at the German Aerospace Center (DLR) in order to determine effects of varying degrees of sleepiness and alcohol on cognitive performance. A total of 48 subjects in cohorts of eight subjects each will stay for twelve consecutive days and nights in the AMSAN sleep laboratory in Cologne. During their stay in the laboratory, subjects are deprived of sleep in a successive manner totally and partially. In addition, on one day they are exposed to moderate alcohol levels. In between the interventions two recovery days are provided per design. A short test-battery of pilot's and air traffic controller's aptitudes is administered to the subjects including spatial orientation, perceptual speed and control of attention tests. In addition, self-concept of mental fitness is measured via a questionnaire prior to and after the cognitive tasks in order to examine whether potential performance decrements due to fatigue are recognized by the subjects. The study design and some preliminary data of the first two cohorts (N = 16) are presented in the paper.

After a continuous time of wakefulness or insufficient rest within 24 hours, human performance can be impaired significantly. Scientific research showing adverse effects of acute fatigue on different cognitive processes is well established and summarized for example in a recent meta-analysis by Lim & Dinges (2010). However, many laboratory studies did administer primarily simple reaction time tasks and therefore it is less apparent whether and how more complex cognitive functions are affected by sleep deprivation. Although, fatigue has remained in aviation among the top 10 safety concerns of the NTSB for over 20 twenty years, still it appears to be a severe issue in accident and incident investigations (National Academy of Sciences, 2011). In the US, crew fatigue has been linked to at least 10 major accidents and even more incidents since 1990. FAA has recently announced to address the fatigue issue among commercial pilots by proposing new flight time, duty and rest requirements based on latest scientific findings in the field.

Based on earlier research at the DLR sleep laboratory of combined effects of sleep deprivation and alcohol on performance (Elmenhorst et al., 2008, 2009) a new study is currently ongoing with extended periods of wakefulness up to 38 hours and a wider scale of performance tasks including cognitive, psychomotor tests and a number of self-assessments. This research aims at identifying potential combined and differential effects of alcohol and sleep deprivation on a variety of psycho-physiological measurements. Some initial thoughts and findings are described below.

Method

Sample

A total of 16 healthy subjects (8 female, 8 male) with an average age of 26 years participated in the study until the time this paper has been written. By end of this year 48 subjects will have been gathered. All subjects are mentally and physically in good conditions with no recent history of sleep disorders, medication and drug or alcohol abuse. In advance of the study all subjects received a comprehensive briefing and training in all performance measures. They signed an informed consent form and were paid for participation. The study was approved by the Ethics Committee of the North Rhine Medical Board.

Procedure

Two cohorts (out of finally six), each with eight subjects stayed for twelve consecutive days and nights in the DLR sleep laboratory AMSAN (Samel et al., 1997). All subjects adjusted their sleep-wake cycle to the laboratory conditions one week in advance of the study. Additionally, they were trained intensively on all performance tests with at least 20 trials per test distributed across the two preceding weeks of the experiment.

The first one and a half experimental days and the first night were used for adaptation to the study environment (Adapt). All measurements throughout the following 48 hours period (i.e. until day four) served as a baseline (Base) for the psycho-physiological parameters. On night four subjects were deprived of sleep totally (TSD, no sleep) or partially (PSD, four hours of sleep). Night five and the following days six and seven served as recovery period (Reco). During night seven again all subjects were deprived of sleep totally or partially with two more days and nights for recovery. In the evening of day ten all subjects consumed alcohol up to a calculated blood alcohol concentration of 0.1 % (ALC). During the following night ten, additionally they were deprived of four hours of sleep (PSA, four hours of sleep). Two more days and nights provided sufficient time for recovery before the end of the study on day thirteen. The experimental interventions were rotated systematically from cohort to cohort to avoid sequential effects. However, at the time of writing this paper only two cohorts had completed the experiment.

Support personnel were present in the facility day and night. Medical doctors were available throughout the experiment to provide medical care if needed. Subjects did not leave the facility for the entire time of the experiment. In order to avoid potential influences on alertness, they were asked not to engage in physical exercise, thrilling games or to take nicotine or caffeine.

As per experimental protocol performance testing took place every three hours except during sleeping periods. All performance measures were administered by computers in the private rooms of the subjects and supervised remotely by the investigating team. Each testing session lasted about 45 minutes. Altogether each subject completed 63 testing sessions during the twelve days in the laboratory. No explicit performance feedback was provided.

Performance Measurement

Three tests from the DLR test battery of air traffic controller and pilot selection tests (Goeters, 2004) were chosen to assess the impact of TSD, PSD, PSA, and ALC on important performance factors of aviators.

Aircraft-Position Test (FPT): The Aircraft-Position Test is a test of spatial orientation. Each item consists of two equal aircraft silhouettes with different headings displayed on a computer screen. The task is to mentally rotate one of the aircraft silhouettes according to a number of heading and direction instructions (e.g. 90°R, 270°L, 180°, 90°R). The letters “L” (left) or “R” (right) indicate the direction of the rotation. At the end of these rotations, subjects have to compare the two silhouettes in order to determine a final heading change needed to reach a complete match between the two aircraft. The FPT has a maximum of 112 items in 5 minutes. The error rate and the number of correct answers are counted.

Attention Control Test (KBT): The Mental Concentration Test involves a combination of different cognitive functions such as visual search, working memory, decoding speed, and simple arithmetic under time pressure. On the screen a search area is presented with a set of 19 symbols connected to numbers. For each task, two symbols are displayed which have to be “translated” into the corresponding digits by reference to the search area. As soon as the sum of the two digits has been entered via a numeric keypad the next task appears. After every ten tasks a new search area is presented. In ten minutes subjects can solve a maximum of 200 tasks. The error rate and the number of correct answers are counted.

Visual Perceptual Speed (OWT): The Visual Perceptual Speed test measures the ability to quickly grasp certain details of visually presented information. Images are presented for less than two seconds depicting two instruments and two sets of objects. Subjects have to enter the perceived two instrument readings and afterwards to respond to the question how many objects of either set A or B they have seen. The error rate and the number of correct answers is counted. The maximum score is 63 points.

In addition to these three aviation tests, one test from the AGARD STRES battery, the Unstable Tracking Task (Santucci et al., 1989), the Psychomotor Vigilance Task (Dinges & Powell, 1985), and the Lane Change Task (Mattes, 2003) for measuring driving performance were administered. However, data of these tests are not included in this paper.

Subjective Measurement

Subjective Fatigue Checkcard (FAT): Subjective fatigue levels (FAT) were assessed before and after each performance testing session using the Samn-Perelli Subjective Fatigue Checkcard (Samn & Perelli, 1982). Subjects reported their current levels of fatigue by ratings of ten different mental states. FAT results in total scores ranging from 0 to 20, which can be categorized into four different fatigue levels related to performance capabilities: class I - severe fatigue; class II – moderate to severe fatigue; class III – mild fatigue; class IV – sufficiently alert. A lower FAT total score indicates a higher level of subjective fatigue. Perceived quality of sleep was rated once per day (except for TSD) by using a sleep diary.

Self-concept of Performance (SOP): Subjects were asked to assess their perceived level of performance on anchored 6-point Likert scales ranging from 1 = Min to 6 = Max. Each self-rating scale was directly related to one of the administered performance tests as described above. Subjects assessed their expected performance level immediately before each testing session and also retrospectively after each testing.

Additional Data

A number of other psycho-physiological parameters were also collected at nights and during some of the testing sessions. The results of these psycho-physiological data will be described elsewhere.

Results

In general, the motivation of the subjects to participate in the study was very high. They arrived well prepared with their training and sleep logs for the adaptation day. Though they shared the facility the entire time together with people they had not seen before, we did not encounter controversies or any tense situations in the first cohorts. Just in a few cases medication was needed against headache. A few data had been lost because one subject left a day earlier due to family reasons and one subject arrived four days later because of a prior illness.

The treatment effects were clearly visible in the acute subjective fatigue scores (FAT). According to the Fatigue Checkcard scoring procedure, subjects remained on average in the highest class IV (= alert) during the majority of measurements on baseline and recovery days. During PSD/PSA and TSD they dropped down to class II (= moderate to severe fatigue). It remains to be shown if cumulative sleep debt was present. In fact, the recovery periods for the second and third treatment seemed to be slightly longer than for the first. The mean course of subjective fatigue is depicted in Figure 1 for the first eight subjects. The scores for the second cohort appeared to be equivalent but cannot be included in the chart because of the different sequence of treatments. In summary, these data show that the treatment had substantial effects as expected on the subjective levels of fatigue of the participants.

The mean performance course for the KBT for cohort 1 is shown in Figure 2. In spite of the intensive training before the experiment the data revealed a positive performance gradient especially during the adaptation and baseline days. This is confirmed by the data of the second cohort and also for the FPT and OWT that subjects obviously continued improving their average performance by about 5% to 10% over the 63 repeated measurements. The magnitude of this effect has to be taken into account for the evaluation of the degree of performance impairment due to sleep loss and alcohol. Therefore, we have decided to use the average of the baseline measurement (Base2) and the measurement on the recovery day following the treatment (second recovery day for TSD) at the corresponding point of time as baseline scores in order to provide a preview of effects for the purpose of this paper.

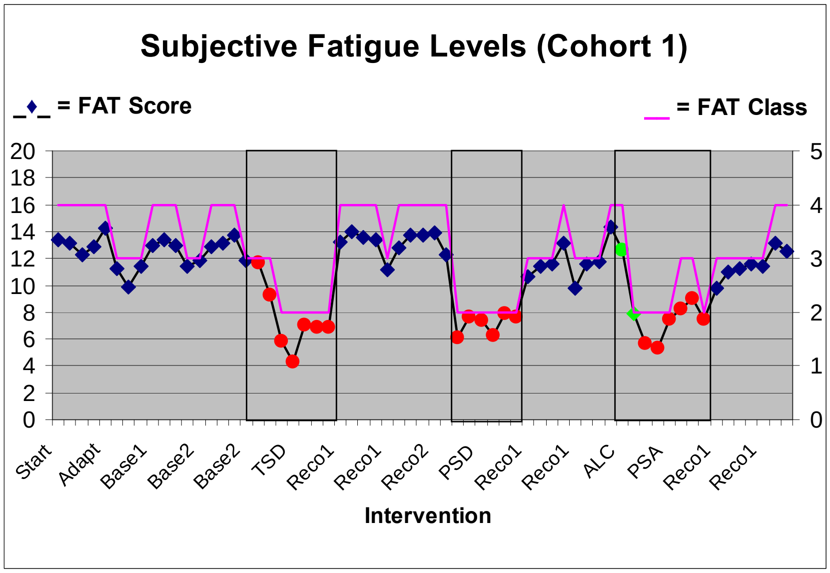

Figure 1. Subjective levels of fatigue throughout the experiment (Cohort 1, N=8). Colour and boxes indicate periods of treatment (blue = no treatment, red = TSD/PSD/PSA, green = ALC)

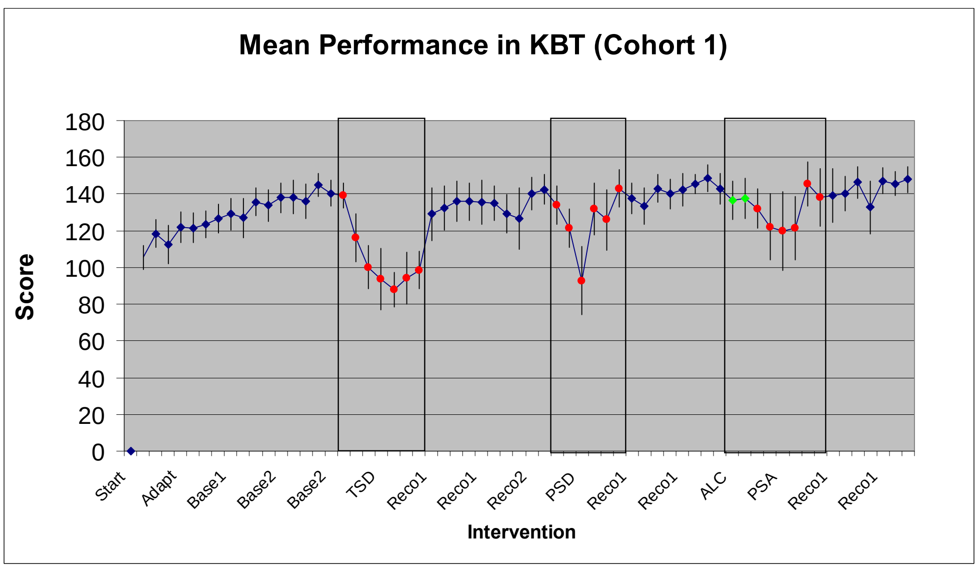

Figure 2. Mean performance scores in KBT (Cohort 1, N=8). The error bars indicate the corresponding standard deviations.

At the first glance the effects of sleep deprivation on levels of performance as shown in Figure 2 seem to be massive especially for TSD. Statistical significance of systematic treatment effects was tested with t-tests with dependent samples. The results of the 9am measurements for TSD, PSD, and PSA and the 6pm measurement for ALC are shown in Table 1. Each of the three tests offered total scores for aspects of performance quantity and accuracy (error rate). TSD at 9am was equivalent to 26 hours of wakefulness. For PSD and PSA subjects had 4 hours of sleep in the preceding 26 hours. Alcohol intake was scheduled 2 hours in advance of the 6pm test session on day ten, which means a calculated blood alcohol concentration of 0.07% to 0.08% at the time of the respective performance measurement.

Though our results are only preliminary and shown here primarily for the purpose of illustrating the effectiveness of the experimental design, a rather consistent pattern of effects seemed worth to be noted. While sleep deprivation impaired performance quantity significantly for most of the tests, alcohol had significant effects primarily on performance accuracy but less on performance quantity. Under the influence of TSD for example, performance was slowed down by about 40%, while the error rates remained rather unaffected. Statistical effect sizes varied from -0.6 standard deviations of performance decrements in the partial sleep deprivation and alcohol conditions (moderate effects) up to -1.5 standard deviations of performance decrements for the condition of total sleep deprivation (large effects).

Table 1. *Effects of sleep deprivation and alcohol on performance. Results of t-tests with two-tailed significance levels. Significant effects (p < 0.05) are in boldface characters. The effect size d was calculated according to Morris & DeShon (2002).*

	Performance Test	Quantity	Error Rate
TSD			
	FPT	**t(14) = 5.8, p = 0.00, d = -1.5**	t(14) = -1.3, p = 0.22
	KBT	**t(14) = 4.8, p = 0.00, d = -1.4**	**t(14) = -3.1, p = 0.01, d = 0.8**
	OWT	**t(14) = 3.0, p = 0.01, d = -0.8**	t(14) = -1.9, p = 0.08
PSD			
	FPT	**t(15) = 2.2, p = 0.04, d = -0.6**	t(15) = 1.1, p = 0.30
	KBT	**t(15) = 3.2, p = 0.01, d = -0.8**	t(15) = -2.0, p = 0.07
	OWT	t(15) = 0.3, p = 0.80	t(15) = 0.5, p = 0.61
PSA			
	FPT	**t(13) = 2.4, p = 0.03, d = -0.6**	t(13) = -0.9, p = 0.39
	KBT	**t(13) = 2.1, p = 0.05, d = -0.7**	t(13) = -1.4, p = 0.18
	OWT	t(13) = -0.9, p = 0.39	t(13) = 2.0, p = 0.07
ALC			
	FPT	t(15) = 1.6, p = 0.13	**t(15) = -2.4, p = 0.03, d = 0.6**
	KBT	t(15) = 1.2, p = 0.23	**t(15) = -2.6, p = 0.02, d = 0.6**
	OWT	t(15) = -1.1, p = 0.28	t(15) = 1.4, p = 0.18

Discussion

According to the initial analyses massive performance decrements were identified for all three tests after a period of 26 hours of wakefulness. So far our data have shown that sleep deprivation can affect performance on spatial orientation and complex attention tasks by a reduction of work quantity. The amount of deterioration is significantly higher for TSD compared to PSD and PSA. Performance in the perceptual speed test was also impaired but to a lesser extent. When being debriefed subjects reported having had more difficulties with the more complex tasks because of “interruptions” of their cognitive processes due to fatigue. When loosing the thread of thought they had to go back and restart the task from the beginning, which cost them time. Compared to FPT and KBT performance in the OWT seemed to have a higher degree of automaticity, which was less likely to be interrupted.

It is interesting to compare the potential influences of alcohol and fatigue for the current subsample of 16 subjects. The data indicated that while fatigue seemed to reduce the level of productivity in general, under the influence of alcohol work quantity remained unchanged. However, working speed seemed to be on cost of accuracy because the error rate was deteriorating significantly after alcohol consumption. Since our study is still ongoing, what is reported here are only preliminary findings, which need to be verified when all subjects have completed the experiment.

With respect to operational crewmembers, lower performance does not necessarily imply less safety. Final outcomes depend for example on how pilots decide on their course of action ahead. It was shown in other research that fatigued pilots tend to avoid riskier options and require more time to collect necessary information before finalising a decision. However, this can lead to higher time pressure during high workload phases as well. And also the freedom of choice can be limited by contextual factors. In further analyses of this study we will analyse how the measures of self-awareness of performance are related to differential effects on the performance tests. Self-efficacy is an important factor in decision making. If confidence is not realistic decisions can lead to undesired outcomes.

References

Dinges D.F., Powell J.W. (1985). Microcomputer analysis of performance on a portable, simple visual RT task during sustained operations. Behavior Research Methods. Instruments and Computers, 17, 652–655.

Elmenhorst D. , Elmenhorst E.-M., Luks N., Maass H., Mueller E.W., Vejvoda M., Wenzel J., Samel A. (2009). Performance impairment during four days partial sleep deprivation compared with the acute effects of alcohol and hypoxia. Sleep Medicine, 10, 189–197.

Elmenhorst, E.-M., Elmenhorst D., Luks N., Maass H., Vejvoda M., Samel A. (2008). Partial sleep deprivation: Impact on the architecture and quality of sleep. Sleep Medicine, 9, 840–850.

Goeters, K.M. (Ed.) (2004). Aviation psychology: Practice and research. Ashgate, Aldershot.

Lim, J. & Dinges, D.F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. Psychological Bulletin, 136, 375-389.

Mattes, S. (2003). The lane change task a tool for driver distraction evaluation. In H. Strasser & H. Bubb (Eds.), Quality of work and products in enterprises of the future. Ergonomia, Stuttgart.

Morris, S.B. & DeShon, R.P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychological Methods, 7, 105-125.

National Research Council of the National Academies (2011). Issues in commuting and pilot fatigue: Interim report. The National Academies Press, Washington, D.C., www.nap.edu .

Samel, A., Diedrich, A., Drescher, J. Lorenz, B., Plath, G., Vejvoda, M. & Wenzel, J. (1997). Langzeitmonitoring psychophysiologischer Größen in der Flugphysiologie. Internist, 38(8), 755-769.

Samn, S.W., Perelli, L.P., 1982. Estimating aircrew fatigue: a technique with application to airlift operations. Brooks AFB, USAF School of Aerospace Medicine. Technical report SAM-TR-82-21.

Santucci, G., Boer, L., Farmer, E., Goeters, K.M., Grisset, J., Schwartz, E., Wetherall, A., Wilson, G. & Yates, R. (1989). Human performance assessment methods. AGARDograph N° 309. NATO-AGARD, Neuilly-sur-Seine.