

Clinical Neurophysiology 111 (2000) 457-463



www.elsevier.com/locate/clinph

# Test-retest reliability of cognitive EEG

L.K. McEvoy\*, M.E. Smith, A. Gevins

EEG Systems Laboratory & SAM Technology, 425 Bush Street, San Francisco, CA 94108-3708, USA

Accepted 20 September 1999

#### Abstract

**Objective:** Task-related EEG is sensitive to changes in cognitive state produced by increased task difficulty and by transient impairment. If task-related EEG has high test-retest reliability, it could be used as part of a clinical test to assess changes in cognitive function. The aim of this study was to determine the reliability of the EEG recorded during the performance of a working memory (WM) task and a psychomotor vigilance task (PVT).

**Methods**: EEG was recorded while subjects rested quietly and while they performed the tasks. Within session (test–retest interval of  $\sim$ 1 h) and between session (test–retest interval of  $\sim$ 7 days) reliability was calculated for four EEG components: frontal midline theta at Fz, posterior theta at Pz, and slow and fast alpha at Pz.

**Results**: Task-related EEG was highly reliable within and between sessions (r > 0.9 for all components in WM task, and r > 0.8 for all components in the PVT). Resting EEG also showed high reliability, although the magnitude of the correlation was somewhat smaller than that of the task-related EEG (r > 0.7 for all 4 components).

Conclusions: These results suggest that under appropriate conditions, task-related EEG has sufficient retest reliability for use in assessing clinical changes in cognitive status. © 2000 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Resting EEG; Working memory; Psychomotor vigilance; Frontal midline theta; Alpha

## 1. Introduction

Certain components of the ongoing EEG, recorded while subjects perform a cognitive task, are sensitive to task difficulty and to changes in an individual's cognitive state. For example, frontal midline theta increases with increased task difficulty, whereas signals in the alpha band decrease as task difficulty increases (Gevins et al., 1979, 1997, 1998; McCallum et al., 1988; Miyata et al., 1990; Yamamoto and Matsuoka, 1990; Gundel and Wilson, 1992; Smith et al., 1999). Transient mental impairment, such as that associated with alcohol intoxication or fatigue also affects EEG signals. Acute alcohol intoxication increases the amplitude of slow alpha (Davis et al., 1941; Lukas et al., 1986; Cohen et al., 1993), whereas fatigue is associated with increased diffuse theta (Davis et al., 1937; Gevins et al., 1977; Matoušek and Petersen, 1983; Makeig and Jung, 1995) and decreased fast posterior alpha (Gevins et al., 1977; Gevins and Smith, 1999). These findings have led to the suggestion that task-related EEG measures could be used to assess temporary changes in cognitive status due to such factors

E-mail address: linda@eeg.com (L.K. McEvoy)

PII: S1388-2457(99)00258-8

as fatigue, intoxication, illness, injury, or drug consumption (Gevins and Smith, 1999).

A necessary first step in the development of an EEGbased method to assess changes in cognitive function is to demonstrate that such measures have high test-retest reliability. A number of studies have shown that the EEG is relatively stable when measured during resting states or during the performance of low load cognitive tasks, such as oddball stimulus detection. For example, in a young, healthy adult population, Burgress and Gruzelier (1993) reported average reliabilities of 0.81 and 0.86 for theta and alpha features in resting, eyes open EEG with a testretest interval of about 1 h. Salinsky et al. (1991) reported high correlations (above 0.9) during the performance of an auditory oddball task with a test-retest interval of 5 min and somewhat lower correlations (of about 0.8) for an interval of 12-16 weeks. Pollock et al. (1991) found correlations between 0.8 and 0.84 for theta and alpha features in resting, eyes open and eyes closed data from a middle-aged population with a test-retest interval of 4.5 months.

The purpose of the current study was to assess the reliability of task-recorded EEG during the performance of tasks that could be used as part of a test battery to assess changes in cognitive state. Two tasks were used: a psychomotor

<sup>\*</sup> Corresponding author. Tel.: +1-415-837-1699; fax: +1-415-274-9575.

vigilance task (PVT, Dinges and Powell, 1985) and a working memory (WM) task. The PVT requires subjects to remain vigilant and make a simple response to the onset of an infrequent stimulus. Performance in this task has been shown to be highly sensitive to lapses in alertness due to sleepiness (Lisper and Kjellerberg, 1972; Glenville et al., 1978; Dinges et al., 1987, 1990; Kribbs et al., 1993; Kribbs and Dinges, 1994; Rosekind et al., 1994). The WM task requires subjects to attend to stimuli and make matching judgments on each stimulus relative to stimuli presented on previous trials. Two difficulty levels were used that differed in the amount of information subjects were required to remember. This task has been shown to induce changes in theta and alpha signals proportional to task-difficulty (Gevins et al., 1997, 1998; Smith et al., 1999). For comparison purposes, we also assessed the reliability of EEG recorded while subjects rested quietly with their eyes open.

## 2. Materials and methods

## 2.1. Subjects

Twenty healthy adults (18–29 years, mean age  $22.9 \pm 3.3$  years; 11 females) were paid to participate in the study. The use of human subjects in this study was approved by institutional review and all participation was fully informed and voluntary.

## 2.2. Tasks

Subjects performed two difficulty levels of a continuous performance WM task (Fig. 1) that required them to compare the spatial location of a current stimulus with the location of one presented previously (for full details, see Gevins et al., 1996). Briefly, single capital letter stimuli, drawn randomly from a set of 12, were presented on a computer monitor. At the beginning of each trial a warning cue, a small 'x', appeared in the center of the screen for 200 ms. The letter stimulus occurred 1.3 s after the onset of the warning cue in one of 12 locations on an imaginary circular grid. Stimuli were presented for 200 ms once every 4.5 s. The identity of the letter and its spatial position varied randomly from trial to trial. A small fixation dot was continuously present at the center of the screen.

In a difficult version of the task, subjects compared the spatial position of the current stimulus with the position of the stimulus presented two trials previously. Subjects were thus required to remember two positions (and their sequential order) for the duration of two trials (9 s), and to update that information on each subsequent trial. In an easy version of the task, subjects were required to match the position of the current stimulus with the position of the first stimulus in the block. In both versions of the task, stimuli were presented in blocks of 53 trials (the first 3 trials were warm-up trials and were discarded from analysis), with

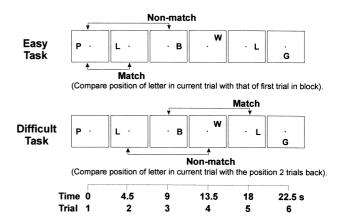


Fig. 1. Schematic diagram of the easy and difficult version of the WM task. Every 4.5 s, one of 12 possible capital letter stimuli appeared in one of 12 possible locations on a computer monitor. In the easy version of the task, subjects compared the location of the current stimulus, regardless of its identity as a letter, with the location of the very first stimulus in the block. In the difficult version of the task, subjects compared the location of current stimulus with the location of the stimulus presented two trials previously.

50% matches. Subjects were instructed to respond as quickly and accurately as possible.

The psychomotor vigilance task (PVT) required subjects to respond as quickly as possible to the onset of a stimulus. Subjects fixated on a box in the center of the computer screen. At varying intervals (between 2 and 10 s, average inter-stimulus interval of 6 s), a stimulus appeared in the center of the box. The stimulus consisted of a rapidly ascending series of numbers (increasing in milliseconds). Subjects were required to press a response key with the right index finger as quickly as possible after stimulus onset. The button press froze the stimulus briefly on the screen, providing subjects with immediate reaction time feedback. The task was presented in 10-min blocks.

## 2.3. Procedure

Each subject participated in 3 sessions. The first session was a practice session in which subjects learned to perform the WM tasks until accuracy and reaction time stabilized. Subjects also performed one block of the PVT. EEG data were not recorded during this session. Subjects then participated in two sessions in which the EEG was recorded continuously during task performance. The two sessions were scheduled for the same time of day, and occurred, on average, 7 days apart.

Each session consisted of a 'warm-up' block followed by two test blocks. The warm up block consisted of two repetitions each of the WM tasks and a short (3 min) version of the PVT. Subjects then performed two test blocks. Each test block consisted of a 4 min recording of resting, eyes open EEG, two 53-trial blocks of each of the WM tasks (order counterbalanced across subjects) and a 10 min block of the PVT. Subjects were given a rest break between the two test

blocks, so that the second test block began approximately 1 h after the start of the first test block.

# 2.4. Electrophysiological recordings

In both resting and task conditions, EEG was recorded continuously from 28 scalp locations using an electronically-linked mastoids reference. EOG was recorded from electrodes placed above and below the left eye, and at the outer canthi of each eye. Physiological signals were recorded with a bandpass filter of 0.01–100 Hz and sampled at 256 Hz. Data were digitally filtered off-line with a 2 Hz high pass band prior to performing automated artifact detection. Eye artifacts were removed using adaptive eye movement artifact removal filters (cf. Du et al., 1994). Data were then visually inspected and segments containing residual artifacts were discarded. Approximately 5–10% of the data were rejected due to artifacts, leaving about 3.5 min of data for each block of the WM task and resting conditions, and about 9 min of data per block of the PVT task.

## 2.5. Data analysis

Fast Fourier transforms were computed on 50% overlapped, 512 sample (i.e. 2 s) Hanning windows for all artifact-free trials. Average power spectra were then computed across all trials in each block for each subject, and a log transform was applied to the power spectra prior to statistical analysis.

The power spectra were computed across 4 s trials occurring around stimulus presentation, and thus may have included energy related to the stimulus event-related responses (ERPs) in the WM tasks and in the PVT. The 2 Hz digital high pass filter removed any slow ERPs elicited in these tasks and attenuated P300 responses. In previous analyses, we have not found differences when power spectra were computed across the entire trial, or when the half-second period following the stimulus has been omitted.

The power of four EEG components known to be sensitive to task difficulty or drowsiness was measured. Frontal midline theta was measured as the peak frequency between 4 and 7 Hz (average of 6 Hz) at Fz. Posterior theta was measured as the average power between 4 and 6 Hz at Pz. Slow and fast alpha were measured in 1-Hz bands at Pz.

Slow alpha was identified as the peak frequency over parietal and central channels between 8 and 12 Hz (mean 9.5 Hz). Fast alpha was defined based on peak frequency between 8 and 12 Hz (mean 11.0 Hz) at occipital electrodes.

Pearson correlation coefficients were used to assess within-session (block 1 vs. block 2) and between-session (session 1 vs. session 2) test-retest reliability. Repeated measures analyses of variance (ANOVAs) were used to examine the significance of any changes in power within or between sessions, and to examine difficulty effects in the task-related EEG. Prior to these ANOVAs, data were converted to standard scores to reduce between-subject variability.

#### 3. Results

#### 3.1. Behavioral results

#### 3.1.1. WM task

Mean accuracy and reaction time (RT) in the WM task are shown in Table 1. Both accuracy (assessed using d'scores) and RT (evaluated after normalizing with a  $\log_{10}$  transform) were affected by task difficulty. Responses were significantly faster and more accurate in the easy task than in the difficult task (d': F(1,19) = 12.62, P < 0.01; RT: F(1,19) = 34.08, P < 0.001). Accuracy did not differ within or between sessions. RT showed a significant main effect of block (F(1,19) = 5.07, P < 0.05), and a significant block by session interaction (F(1,19) = 7.13, P < 0.05). Post hoc analysis indicated that this interaction was due to significantly faster RTs in Block 2 than in Block 1 in the second test session only.

Accuracy scores showed relatively low reliability both within and between sessions (average within session correlation of 0.45, average between session correlation of 0.47; Table 2). RTs were highly correlated both within and between sessions (average within session correlation of 0.93, average between session correlation of 0.86).

# 3.1.2. PVT

Behavioral measures obtained in the PVT included mean RT (evaluated after normalizing with a log<sub>10</sub> transform), the

Table 1
Average accuracy (SEM) and average reaction time (SEM) in the easy and difficult versions of the WM task

	Easy WM Task		Difficult WM Task		
	First block	Second block	First block	Second block	
Accuracy (d')					
First session	4.43 (0.34)	4.46 (0.33)	4.00 (0.52)	3.81 (0.68)	
Second session	4.35 (0.45)	4.47 (0.39)	4.10 (0.68)	3.98 (0.62)	
RT (ms)					
First session	494 (9.9)	487 (10.2)	607 (14.3)	614 (14.7)	
Second session	504 (8.1)	469 (8.1)	616 (12.3)	585 (11)	

Table 2 Within-session and between-session test–retest reliability coefficients for the behavioral measures in the easy and difficult WM tasks. To assess within-session reliability, Pearson correlations were computed separately for the first and second sessions. Between-session reliability was computed separately for the first and second block. All correlations are significant at P < 0.001 unless otherwise indicated (one-tailed significance reported)<sup>a</sup>

	Within-session reliability		Between-session reliability		
	Session 1	Session 2	Block 1	Block 2	
Accuracy					
Easy	0.76	0.37 (n.s.)	0.53**	0.40*	
Difficult	0.53**	0.12 (n.s.)	0.41*	0.55**	
RT					
Easy	0.95	0.89	0.90	0.90	
Difficult	0.96	0.91	0.78	0.87	

<sup>&</sup>lt;sup>a</sup> n.s., not significant; \*\*P < 0.00; \*P < 0.05.

mean of the fastest 10% of the responses (fast response coefficient), number of lapses (missed responses, or responses occurring more than 500 ms after stimulus onset) and number of response errors (false responses). The mean behavioral scores for these measures are shown in Table 3. Since subjects performed the tasks in an alert state, there were few lapses or response errors. The behavioral measures did not show significant effects of test block or test session, nor did they show any significant interactions.

The across session and within session correlation coefficients for the behavioral measures are shown in Table 4. Although mean RT showed moderate within-session and between-session correlations (average within session correlation of 0.60 and average between session correlation of 0.52), the fast response coefficient showed high reliability (average within session correlation of 0.83 and average between session correlation of 0.80). Lapses showed low, nonsignificant within-session reliability, and moderate between session reliability (average of 0.49). Response errors showed higher correlations (average within session correlation of 0.75 and average between session correlation of 0.63).

# 3.2. Electrophysiological results

The average spectral power density for each EEG compo-

Table 3
Mean (SEM) reaction time, fast response coefficient (fast coeff.), number of lapses, and number of responses errors in the PVT

	First session		Second session	
	First block	Second block	First block	Second block
Mean RT (ms)	280 (7)	280 (7)	301 (16)	279 (7)
Fast coeff. (ms)	213 (3.5)	209 (5)	212 (4.1)	211 (4.2)
Lapses	2.16 (0.36)	2.66 (0.33)	3.30 (0.60)	2.91 (0.36)
Errors	0.90 (0.29)	0.79 (0.25)	1.42 (0.53)	1.16 (0.31)

nent in each condition (averaged across all testing intervals) is shown in Table 5.

# 3.2.1. WM load effects

Frontal midline theta was significantly larger (by an average of 0.76 dB) in the difficult task than in the easy task  $(F(1,19)=10.47,\ P<0.01)$ . In contrast, both slow and fast alpha were significantly smaller in the difficult task than in the easy task (slow alpha F(1,19)=59.53, P<0.001; average difference of 1.18 dB; fast alpha:  $(F(1,19)=52.29,\ P<0.001$ ; average difference of 1.17 dB). Posterior theta was not significantly affected by task difficulty.

# 3.2.2. Power changes within and between sessions

In the WM task posterior theta showed a significant 3 way interaction involving task version, block and session (F(1,19) = 5.99, P < 0.05). In the easy task, posterior theta power did not differ across test blocks or across test sessions. However, in the difficult task there was a significant decrease in power (by an average of 0.32 dB) between the first and second block in the second test session only (F(1,19) = 7.76, P < 0.05). None of the other components showed significant changes within or between sessions.

In the PVT, frontal midline theta was larger in the second test session than in the first (by an average of 0.53 dB (F(1, 18) = 5.63, P < 0.05)). No other component differed between the two sessions, nor did any component significantly differ between the first and second block within a test session.

In the resting EEG, fast alpha power was smaller in the second block than in the first, by an average of 0.7 dB (F(1,18) = 7.62; P < 0.05). No other component showed significant changes within or between sessions.

# 3.2.3. Reliability measures

Within-session and between-session correlation coefficients for each of the EEG components are shown in Table 6. In the task-related EEG, all components showed very high reliability both within and between sessions. For

Table 4 Within-session and between-session test–retest reliability coefficients for PVT behavioral measures. Within-session reliability was computed separately for the first and second session, between-session reliability was computed separately for the first and second test block. All correlations are significant at P < 0.001 unless otherwise indicated (one-tailed significance reported)<sup>a</sup>

	Within session reliability		Between session reliability		
	Session 1	Session 2	Block 1	Block 2	
Mean RT	0.62**	0.58**	0.53*	0.51*	
Fast coeff.	0.72	0.93	0.82	0.78	
Lapses	0.29 (n.s.)	0.36 (n.s.)	0.48*	0.50*	
Errors	0.73	0.78	0.63**	0.63**	

<sup>&</sup>lt;sup>a</sup> n.s., not significant; \*\*P < 0.01; \*P < 0.05.

Table 5
Average power spectral density (SEM), in dB, for frontal midline theta, posterior theta, slow alpha and fast alpha in each condition

	Frontal midline theta	Posterior theta	Slow alpha	Fast alpha
Easy WM	29.68 (0.29)	28.19 (0.20)	29.06 (0.50)	29.78 (0.48)
Difficult WM	30.44 (0.35)	28.12 (0.20)	27.88 (0.65)	28.62 (0.46)
PVT	29.6 (0.28)	28.35 (0.20)	29.79 (0.50)	30.03 (0.48)
Resting	29.0 (0.26)	27.58 (0.21)	29.42 (0.51)	30.12 (0.52)

the WM task, *r*-values ranged from 0.95 to 0.99 within sessions and 0.87 to 0.97 between sessions. For the PVT, *r*-values ranged from 0.92 to 0.97 within sessions and from 0.83 to 0.94 between sessions. Reliability was somewhat lower in the resting data: within-session *r*-values ranged from 0.74 to 0.92; between-session *r*-values from 0.76 to 0.95. Within subjects, the average change in the amplitude of the power spectra across the four testing sessions was quite small, with smallest changes seen in the WM task and largest changes seen in the resting data. Across the 4 components, the average absolute change (and standard error) within subjects across the four sessions was 1.44 (0.13) dB for data in the WM task, 1.78 (0.22) dB for data in the PVT, and 2.40 (0.33) dB for resting data.

Reliability of the task-difficulty related modulation of frontal midline theta and alpha signals was also examined. The difference in theta and alpha power between the easy and difficult tasks was computed, and the reliability coefficients of these differences were calculated (Table 7). Since these difference values have a much-restricted range relative to the range of the raw (log-normalized) power spectra, the reliability coefficients will be necessarily smaller. Nevertheless, the task-difficulty effect of the frontal midline theta signal showed good reliability, with an average within-session reliability of 0.77 and an average between session reliability of 0.78. The task-difficulty effect for the slow alpha component showed moderately good reliability, with an average within-session reliability of 0.68 and an

average between session reliability of 0.69. The task-modulation of the fast alpha signal was less reliable, with an average within-session reliability of 0.51 and a between session reliability of 0.43.

#### 4. Discussion

This study showed that the power spectral density of four EEG components recorded during performance of PVT and WM tasks from alert, healthy young adults, well practiced in the task performance, was highly reliable both within and between sessions. Average within session reliability, in which the test-retest interval was approximately 1 h, was above 0.9 in the task data and above 0.8 in the resting data. Similar reliability was observed between sessions, in which the test–retest interval was approximately seven days. These values are in agreement with previously reported reliability for resting EEG and EEG associated with the performance of an oddball task. (Pollock et al., 1991; Salinsky et al., 1991; Burgress and Gruzelier, 1993). The greater reliability of task-related EEG than of resting EEG is likely due to larger variations in attention and alertness levels in resting states than during task performance. Task performance imposes a more uniform level of alertness and mentation, and thus has a stabilizing effect on the EEG. Interestingly, the within and between session reliability for the different task types (vigilance; easy and difficulty working memory

Table 6 Within-session and between-session test–retest reliability coefficients for EEG measures in the easy and difficult WM tasks, the PVT, and resting condition. Within-session reliability was computed separately for the first and second session; between-session reliability was computed separately for the first and second test block. All correlations are significant at P < 0.001 (one-tailed significance reported)<sup>a</sup>

	Within-session reliability for first session				Within-session reliability for second session			
	FM theta	Post. Theta	Slow alpha	Fast alpha	FM theta	Post. Theta	Slow alpha	Fast alpha
Easy	0.97	0.95	0.97	0.98	0.95	0.97	0.99	0.98
Difficult	0.96	0.97	0.97	0.96	0.97	0.96	0.98	0.97
PVT	0.95	0.94	0.96	0.97	0.92	0.94	0.96	0.93
Resting	0.87	0.84	0.92	0.82	0.74	0.86	0.93	0.93
	Between-session reliability for first block		Between-session reliability for second block					
Easy	0.92	0.93	0.95	0.95	0.87	0.92	0.96	0.97
Difficult	0.91	0.91	0.97	0.97	0.93	0.92	0.97	0.94
PVT	0.92	0.92	0.93	0.93	0.83	0.86	0.90	0.94
Resting	0.76	0.86	0.90	0.89	0.81	0.78	0.95	0.77

<sup>&</sup>lt;sup>a</sup> FM theta, frontal midline theta; post. theta, posterior theta.

Table 7 Within-session and between-session test–retest reliability coefficients for task-difficulty-related modulation of frontal midline theta, slow alpha and fast alpha signals. Within-session reliability was computed separately for the first and second session; between-session reliability was computed separately for the first and second test block. All correlations are significant at P < 0.001 (one-tailed significance reported) unless otherwise indicated a

	Within session reliability		Between session reliability		
	Session 1	Session 2	Block 1	Block 2	
FM theta	0.71	0.83	0.67	0.89	
Slow alpha	0.61**	0.75	0.61**	0.78	
Fast alpha	0.48*	0.54**	0.55**	0.32 (n.s.)	

<sup>&</sup>lt;sup>a</sup> n.s., not significant; \*\*P < 0.01; \*P < 0.05.

tasks) was very similar. All these tasks required sustained visual attention, however, the working memory task also contained a judgment component, and the more difficult task level required constant updating of information in memory. Despite this variation in task complexity, the EEG was highly stable during the performance of all tasks.

EEG during the WM tasks showed difficulty effects consistent with those previously reported for this task (Gevins et al., 1997, 1998). Frontal midline theta was significantly larger in the more difficult version than in the easier version whereas slow and fast alpha were significantly smaller in the more difficult version than in the easier version. The difficulty-related modulation in the frontal midline theta signal and the slow alpha signal were quite reliable, with average correlation coefficients of 0.77 and 0.68, respectively. The difficulty-related effect in the fast alpha signal was less reliable, with an average correlation of 0.47. Consistent with previous results, posterior theta was not found to vary with task difficulty. This component is affected more by arousal manipulations than by task difficulty manipulations (Gevins and Smith, 1999).

Many of the behavioral measures in this study showed generally lower reliability than did the task-related EEG. This was likely due to ceiling and floor effects. In these well-practiced subjects, accuracy in the WM task was at near perfect levels for both task versions. Similarly, these alert, healthy subjects produced few erroneous responses in the PVT. RT proved a more stable measure. High reliability was observed for average RT in the WM task and for the fastest 10% of responses in the PVT. The magnitude of these correlations was similar to that of the task-related EEG measures.

It is important to note that EEG reliability was assessed using measurements from midline sites, Fz and Pz, that are less likely to be affected by muscle contamination than are more peripheral sites. Muscle activity can produce large broadband differences in the power spectra as a function of posture and tension level, and these differences are most apparent at electrode sites overlying the large scalp muscles. In our data, fast alpha measured at electrode site Oz (a position vulnerable to artifacts from neck muscles)

showed lower reliability than did the corresponding measure at Pz. In the worst case, a between-session reliability of 0.58 was obtained for resting data. Visual inspection of the data showed that this was due to large broadband differences between sessions at occipital sites for a few subjects. Thus to obtain stable measurements of the EEG, it is important to control for the effects of muscle contamination by choosing sites that are less prone to muscle artifacts, by carefully eliminating data segments with muscle contamination, or by developing algorithms to selectively remove muscle contamination.

This study has shown that under certain conditions, EEG associated with the performance of a working memory task and of a psychomotor vigilance task is highly reliable. These findings suggest that both the WM task and the PVT have the necessary reliability for use in assessing transient changes in the neurophysiological signals of cognition due to fatigue, medication, or illness.

# Acknowledgements

Supported by the National Aeronautics and Space Administration and the National Institute of Mental Health. We thank Devon Boyle and Caroline Prioleau for help with data collection and analyses. We are grateful to Dr. David Dinges for supplying us with the psychomotor vigilance task.

## References

Burgress A, Gruzelier J. Individual reliability of amplitude distribution in topographical mapping of EEG. Electroenceph clin Neurophysiol 1993;86:210–223.

Cohen HL, Porjesz B, Begleiter H. The effects of ethanol on EEG activity in males at risk for alcoholism. Electroenceph clin Neurophysiol 1993;86:368–376.

Davis H, Davis PA, Loomis AL, Harvey EN, Hobart G. Human brain potentials during the onset of sleep. J Neurophysiol 1937;1:24–37.

Davis PA, Gibbs FA, Davis H, Jetter WW, Trowbridge LS. The effects of alcohol upon the electroencephalogram (brain waves). Q J Stud Alc 1941;1:626-637.

Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav Res Methods Instrum Comput 1985;17:652–655.

Dinges DF, Orne MT, Whithouse WG, Orne EC. Temporal placement of a nap for alertness: contributions of circadian phase and prior wakefulness. Sleep 1987;10:313–329.

Dinges DF, Graeber RC, Connell LJ, Rosekind MR, Powell JW. Fatiguerelated reaction time performance in long-haul flight crews. Sleep Res 1990: 19:117

Du W, Leong H, Gevins AS. Ocular artifact reduction by adaptive filtering, Quebec City, CA: Seventh IEEE SP Workshop on Statistical Signal and Array Processing, 1994.

Gevins A, Smith ME. Detecting transient cognitive impairment with EEG pattern recognition methods. Aviat Space Environ Med 1999;70:1018– 1024.

Gevins AS, Zeitlin GM, Ancoli S, Yeager CL. Computer rejection of EEG artifact. II: contamination by drowsiness. Electroenceph clin Neurophysiol 1977;42:31–42.

Gevins AS, Zeitlin GM, Doyle JC, Yingling CD, Schaffer RE, Callaway E,

- Yeager CL. EEG correlates of higher cortical functions. Science 1979;203:665-668.
- Gevins AS, Smith ME, Le J, Leong H, Bennett J, Martin N, McEvoy L, Du R, Whitfield S. High resolution evoked potential imaging of the cortical dynamics of human working memory. Electroenceph clin Neurophysiol 1996:98:327–348
- Gevins A, Smith ME, McEvoy L, Yu D. High resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb Cortex 1997;7:374–385.
- Gevins A, Smith ME, Leong H, McEvoy L, Whitfield S, Du R, Rush G. Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. Human Factors 1998;40:79–91.
- Glenville M, Broughton R, Wing AM, Wilkinson RT. Effects of sleep deprivation on short duration performance compared to the Wilkinson auditory vigilance task. Sleep 1978;1:169–176.
- Gundel A, Wilson GF. Topographical changes in the ongoing EEG related to the difficulty of mental tasks. Brain Topogr 1992;5:17–25.
- Kribbs NB, Dinges D. Vigilance decrement and sleepiness. In: Harsh JR, Ogilvie RD, editors. Sleep onset mechanisms, Washington, DC: American Psychological Association, 1994. pp. 113–125.
- Kribbs NB, Pack AI, Kline LR, Getsy JE, Schuett JS, Henry JN, Dinges DF. The effects of one night without nasal CPAP treatment on sleep and sleepiness in patients with obstructive sleep apnea. Am Rev Respir Dis 1993;147:1162–1168.
- Lisper HO, Kjellerberg A. Effects of 24-hour sleep deprivation on rate of decrement in a 10-minute auditory reaction time task. J Exp Psychol 1972;96:287–290.

- Lukas SE, Mendleson JH, Benedilt RA, Jones B. EEG alpha activity increases during transient episodes of ethanol-induced euphoria. Pharmacol Biochem Behav 1986;25:889–895.
- Makeig S, Jung TP. Changes in alertness are a principal component of variance in the EEG spectrum. NeuroReport. 1995;7:213–216.
- Matoušek M, Petersen IA. A method for assessing alertness fluctuations in vigilance and the EEG spectrum. Electroenceph clin Neurophysiol 1983;55:108–113.
- McCallum WC, Cooper R, Pocock PV. Brain slow potential and ERP changes associated with operator load in a visual tracking task. Electroenceph clin Neurophysiol 1988;69:453–468.
- Miyata Y, Tanaka Y, Hono T. Long term observation on Fm-theta during mental effort. Neuroscience 1990;16:145–148.
- Pollock VE, Schneider LS, Lyness SA. Reliability of topographic quantitative EEG amplitude in healthy late-middle-aged and elderly subjects. Electroenceph clin Neurophysiol 1991;79:20–26.
- Rosekind MR, Graeber RC, Dinges DF, Connell LJ, Rountree M, Spinweber CL, Gillen KA. Crew factors in flight operations: IX. Effects of cockpit rest on crew performance and alertness in long-haul operations. NASA Technol Memorand 108839. 1994.
- Salinsky MC, Oken BS, Morehead L. Test-retest reliability in EEG frequency analysis. Electroenceph clin Neurophysiol 1991;79:383–392.
- Smith ME, McEvoy LK, Gevins A. Neurophysiological indices of strategy development and skill acquisition. Cogn Brain Res 1999;7:389–404.
- Yamamoto S, Matsuoka S. Topographic EEG study of visual display terminal VDT performance with special reference to frontal midline theta waves. Brain Topogr 1990;2:257–267.