USAARL Report No. 89-19

Auditory Evoked Potentials as a Function of Sleep Deprivation

(Reprint)

By

John Harsh
University of Southern Mississippi

and

Pietro Badia
Bowling Green University

September 1989

Approved for public release; distribution unlimited.

United States Army Aeromedical Research Laboratory
Fort Rucker, Alabama 36362-5292

Notice

Qualified requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Cameron Station, Alexandria, Virginia 22314. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this report when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Human use

Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRDC Reg 70-25 on Use of Volunteers in Research.

Reviewed:

GERALD P. KRUEGER, Ph.D.,
LTC, MS
Director, Biomedical Applications
Research Division

J.D. LaMOTHE, Ph.D.
COL, MS
Chairman, Scientific
Review Committee

Released for publication:

DAVID H. KARNEY
Colonel, MC
Commanding

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): USAARL Report No. 89-19	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: University of Southern Mississippi; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: U.S. Army Aeromedical Research Laboratory
6c. ADDRESS (City, State, and ZIP Code): Southern Station Box 9371, Hattiesburg, Mississippi 3906-9371	7b. ADDRESS (City, State, and ZIP Code): P.O. Box 577, Fort Rucker, Alabama 36362-5292
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: U. S. Army Medical Research & Development Command; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code): Fort Detrick, Frederick, Maryland 21701-5012	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
63787A	3E162-787A879	BH	165

11. TITLE (Include Security Classification)

(U) Auditory evoked potentials as a function of sleep deprivation

12. PERSONAL AUTHOR(S)

John Harsh, Pietro Badia

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	1989 September	13

16. SUPPLEMENTARY NOTATION

Originally published in Work & Stress Vol 3, No. 1, 79-91

17. COSATI CODES FIELD	GROUP	SUB-GROUP
01	03	
06	05	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Evoked potentials, sleep deprivation, performance, tracking, monitoring

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Event-related brain potentials (ERPs) were studied in subjects deprived of sleep over a 48-hour test period to assess the effects of different durations of continuous wakefulness on ERP components and to determine whether changes in the ERP components were related to changes in performance. Forty subjects were randomly assigned to either an experimental (sleep deprived) group (n=30) or a control (not sleep deprived) group (n=10). For the experimental subjects, ERP and performance measures were obtained in 4-hour test blocks throughout the 48-hour period. Performance was assessed using the Walter Reed performance assessment battery. The control subjects were tested at the same times except during designated sleep periods. Both performance and evoked potential measures showed systematic changes over the experimental test period in association with sleep deprivation, time of day, and repeated testing. The latency of the N2 component of the evoked potential covaried with throughput measures on the performance assessment battery across the 12 4-hour test (Continued)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Chief, Scientific Information Center	22b. TELEPHONE (Include Area Code): (205) 255-6907; 22c. OFFICE SYMBOL: SGRD-UAX-SI

10. Abstract (Continued)

blocks of the experiment. These data suggest that ERPs reflect central processes that change across the sleep deprivation period and that ERP measures might be useful in assessment and prediction of performance degradation under adverse conditions such as sleep loss.

WORK & STRESS, 1989, VOL. 3, NO. 1, 79–91

Auditory evoked potentials as a function of sleep deprivation

JOHN HARSH*

Department of Psychology, Southern Station Box 9371,
University of Southern Mississippi, Hattiesburg, MS 39406-9371, USA

PIETRO BADIA

Department of Psychology, Bowling Green University,
Bowling Green OH 43403, USA

Event-related brain potentials (ERPs) were studied in subjects deprived of sleep over a 48-h test period to assess the effects of different durations of continuous wakefulness on ERP components and to determine whether changes in the ERP components were related to changes in performance. Forty subjects were randomly assigned to either an experimental (sleep deprived) group ($n = 30$) or a control (not sleep deprived) group ($n = 10$). For the experimental subjects, ERP and performance measures were obtained in four-h test blocks throughout the 48-h period. Performance was assessed using the Walter Reed performance assessment battery. The control subjects were tested at the same times except during designated sleep periods. Both performance and evoked potential measures showed systematic changes over the experimental test period in association with sleep deprivation, time of day, and repeated testing. The latency of the N2 component of the evoked potential covaried with throughput measures on the performance assessment battery across the 12 four-h test blocks of the experiment. These data suggest that ERPs reflect central processes that change across the sleep deprivation period and that ERP measures might be useful in assessment and prediction of performance degradation under adverse conditions such as sleep loss.

Keywords: Evoked potentials; Sleep deprivation; Performance; Tracking; Monitoring.

1. Introduction

More than ever before, machines exceed human ability to process and respond to information in complex stimulus environments. These machines, however, require surveillance by alert human monitors and increases in technological sophistiation have resulted in greater demands on the human operators of human–machine systems in civilian and military work settings. Because of the greater demand on human monitors, there is a greater need to understand the factors that affect their performance. There is a correspondingly greater need for assessing and predicting changes in performance readiness, particularly under conditions such as sleep loss, fatigue, and boredom, that are common in many work environments.

Researchers have studied a variety of measures in their efforts to monitor performance readiness. Peripheral psychophysiological measures have included blood and urine composition, heart rate, galvanic skin response, electromyography, and others. It has been a concern for some time, however, that inferences based on peripheral measures will be of limited value in predicting and understanding performance changes (e.g., Malmo 1959). That is, it is likely these measures deal with physiological systems too distant from the

* To whom reprint requests should be sent.

central processes underlying readiness. The purpose of this study was to assess the usefulness in this regard of event-related brain potentials (ERPs). ERP measures are considered to be more closely related to central processes.

Event-related potentials consist of postive- and negative-going waveforms, often referred to as components, that are evoked by discrete stimulus events. They can be recorded using scalp electrodes. A distinction is usually made between 'exogenous' and 'endogenous' components. Exogenous components are viewed as obligatory, short-latency (< 100 ms) responses whose amplitude, latency, and scalp distribution are determined by the physical properties of the eliciting stimulus. Endogenous components are viewed as nonobligatory longer-latency (> 100 ms) components whose characteristics are determined by the psychological properties of the eliciting stimulus. The components of interest in the present experiment have latencies close to or exceeding 100 ms and are refereed to as P1, N1, P2, N2, and P3 (see Hillyard and Hansen 1986). The neuroanatomical orgins of these components are unclear; however, the relationship between the components and a host of independent variables of interest to human performance researchers suggests that they reflect central states and processes such as attention (see Hillyard and Hansen 1986) and information processing (see Donchin *et al.* 1986).

The goal of our research was to examine whether ERP variables provide a reliable and practical way of predicting performance changes resulting from adverse environmental, task, and field conditions. In addition to being closely related to central processes, ERP measures can be obtained rapidly, inexpensively, and with minimal intrusion. As a first step towards our goal, we focused on identifying fundamental relationships between ERP measures, different levels of sleep deprivation, and performance on a variety of tasks. Sleep deprivation was chosen as a laboratory manipulation because (1) it has been intensively researched; (2) it is an important, yet simple variable to quantify and vary systematically; (3) it enters into relationships with many other variables; and (4) it has high inherent interest in military and civilian work settings.

Two main questions were addressed. The first concerned the effects of sleep deprivation on event-related potentials. Previous studies have shown that ERP variables differentiate awake, sleepy, and sleep states (e.g., Fruhstorfer and Bergstrom 1969, Hakkinen and Fruhstorfer 1967, Weitzman and Kremen 1965, Williams *et al.* 1962) and also normal and pathologically sleepy individuals (e.g., Broughton *et al.* 1982). It has been established that sleep deprivation alters evoked potentials (e.g., Gauthier and Gottesman 1983, Peeke *et al.* 1980, Pressman *et al.* 1982). ERP measures were obtained only at the end of the sleep deprivation period in the latter studies, however, and information about the time course of the evoked potential changes was not provided. In the present experiment, evoked potentials were measured every four hours over a 48-h test period during which half of the subjects were deprived of sleep while the other half were not.

The second question of the present experiment focused on the relationship between changes in the characteristics of ERPs and changes in performance. Performance was assessed using the Walter Reed performance assessment battery (Thorne *et al.* 1985) which was designed to investigate performance change across time, treatment conditions, etc. The battery was not intended as a diagnostic instrument or to reflect specific neuropsychological processes, but to sample a broad range of behavioral functioning.

2. Methods

2.1. *Subjects*

Subjects were 40 male undergraduates between the ages of 18 and 25. Each subject was examined by a physician prior to the experiment. None of the potential subjects were excluded for medical reasons and all subjects reported being in good health. None of the subjects reported using prescription or nonprescription drugs at the time of the experiment. Twenty subjects were recruited and tested at the University of Southern Mississippi and 20 were recruited and tested at Bowling Green State University. All subjects were fully informed concerning details of the research project. Subjects were told that they were free to terminate their participation at any time. To encourage subjects to complete the experiment, they were paid for each day of their participation and told they would receive a $30 bonus for full-term participation. All subjects completed the experiment and were each paid $120.

2.2. *Apparatus and recording procedures*

2.2.1. *Recording of evoked potentials*

Evoked potentials were elicited using an 'odd ball' task which involved asking subjects to count the least probable of two regularly occurring stimuli. Auditory stimuli were presented in Bernoulli series of low-pitched (1000 Hz; $p = 0{\cdot}80$) and high-pitched (1500 Hz; $p = 0{\cdot}20$) 0·5-s tones (65 dB), that were spaced 1·1 s apart and delivered binaurally through earphones. The subjects were instructed to count the number of high-pitched tones. The subjects were also asked to 'tap your foot' upon hearing a high-pitched tone. The purpose of the latter instructions was to allow the experimenter to determine that the subject was attending to the tones.

The odd ball task was either presented alone or concurrently with an 'easy' or a 'hard' version of a tracking task developed in the laboratory. The tracking task was implemented on a microprocessor and involved the subject manipulating a control stick to keep a cursor either on a moving target ('chase') or away from a moving target ('run'). The chase and run modes alternated unpredictably and the speed of the target was varied to make the task 'easy' or 'hard'.

Event-related potentials were recorded to the presentation of the less probable (high-pitched) tone. The recording epoch extended from the onset of the tone for a period of 800 ms. A trial was defined as the period during which tone presentations continued until data from 35 contamination-free (see below) recording epochs were obtained.

The ERPs were recorded for four consecutive trials during each 4-h test block. During the first two trials, the oddball task alone was presented (ERP only). During the last two trials the oddball task was superimposed over the easy and hard versions (in random order) of the tracking task (ERP/tracking).

The ERPs were recorded from Cz-A1 with silver/silver chloride electrodes and with impedance values less than 10 K ohms. Signal averaging was performed by an Apple II plus microprocessor equipped with an RC Electronic Computerscope signal averager. The ERP signals were amplified by a Grass Model 7P122 low-level DC amplifier (TC 0·8, sensitivity at 0·05, high pass filter at 35). When movement or other sources of contamination during a tone presentation resulted in voltages above an adjustable threshold (the threshold level was set at the lowest possible value for each subject), the data from that presentation were rejected.

Visual analysis was used to identify components (P1, N1, P2, N2, P3) of the evoked responses. The latency of each component was obtained by finding the time from stimulus onset to the peak of the waveform. Baseline to peak amplitudes were obtained for each component. Peak-to-peak amplitudes were also calculated by finding the voltage difference between P1–N1, N1–P2, P2–N2, and N2–P3.

2.2.2. *Behavioral measures*

The Walter Reed performance assessment battery (Thorne *et al.* 1985) is a computer controlled multi-task array. The version of the performance assessment battery (PAB) used in this study consisted of several tasks including visual search and recognition (MAST 6); syllogistic reasoning (LOGICAL); short-term memory recall (PROBE-MEM); mental addition and subtraction (SERIAL ADD/SUB); spatial memory (MATRIX 2); and visual/motor coordination (WILKINSON). Completion of the battery required about 25 min.

Although reaction time and accuracy of performance are traditional measures used in sleep deprivation research, theoretically, either of these measures alone may be insufficient to describe performance decrements during sleep deprivation. For example, the subject may choose to work at a slower rate in order to increase accuracy, or to increase speed and thereby sacrificing accuracy. Because of this tradeoff between speed and accuracy, these two measures were combined into a third measure called 'throughput' (see Thorne *et al.* 1983). Throughput is a measure which gives the rate of successes per unit of time. Throughput is derived numerically by calculating per cent correct and dividing by the mean reaction time, and multiplying by a constant.

2.3. *Procedure*

Twenty subjects were tested at each of the laboratories in groups of four on the same days of the week (Wednesday through Sunday) during five consecutive weeks. The subjects were asked to report to the laboratory at 2200 h on Day 1 (Wednesday) and were given final release from the laboratory at no later than 1330 h on Day 5 (Sunday). The subjects slept in the laboratory Wednesday and Thursday night from approximately 2300 h to 0700 h to ensure equal initial levels of sleep and also for adaptation to laboratory conditions. They received practice sessions with the behavioral tasks from 2200 to 2300 h on both nights. During the first two days subjects were free to leave the laboratory during the daytime to attend classes etc. The experiment proper began at 0700 on Day 3 (Friday). Data collection began at 0800 h and proceeded in four-h blocks. During each four-h block, the subjects rotated among four test stations. Data were collected using PAB at one test station. Evoked potential measures were obtained at a separate test station. The remaining stations were used to collect data on other performance tasks (data from these tasks will be reported elsewhere). Testing at each station required from 30–45 min. The order of the different tests varied across subjects but did not vary across test blocks. All test orders were equally represented in each group of subjects.

The experimental subjects ($n = 30$) remained awake from 0700 h Friday until the conclusion of the experiment on Sunday. The control subjects ($n = 10$) were permitted to sleep during each of the test nights (Friday and Saturday night) from 2400 h to 0700 h with the exception of an awakening at 0400 h (25–30 min) for the recording of ERPs.

The subjects spent approximately three of each 4-h block in testing. During their free time they were permitted to read, study, play video games, etc., but were not allowed to leave the laboratory area. Meals were provided and snacks and beverages were available during free periods. Caffeinated beverages and smoking were permitted but only at levels the subjects described as normal before they began the experiment.

3. Results

3.1. *ERP measures*

Evoked potentials were scored following visual inspection of the overall waveform to identify the location of the components referred to as P1, N1, P2, N2, and P3. Measures of

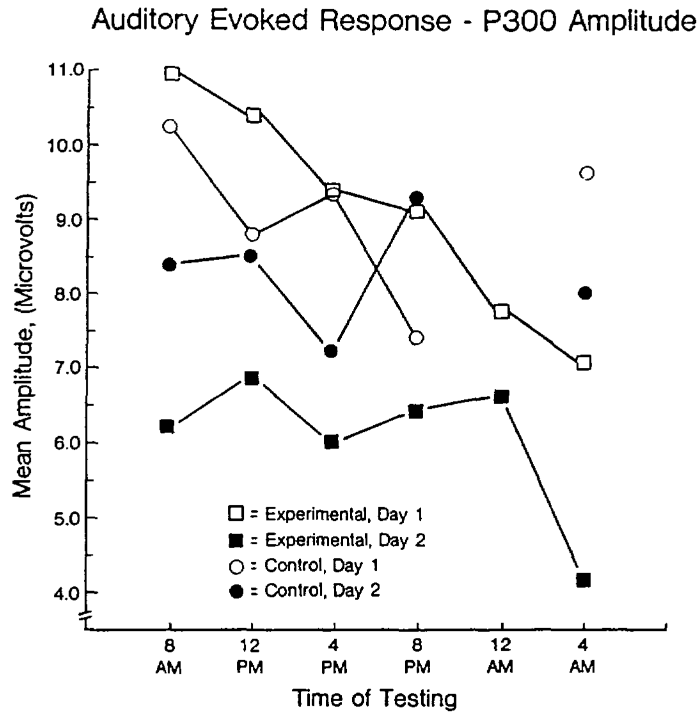

Figure 1. P300 amplitude (in microvolts) for the Experimental and Control groups for each time of day on both test days.

the components for the first two trials (no concurrent tasks; ERP only) of each test session were then averaged together. Similarly, the measures of the components from trials three and four (concurrent tracking task; ERP/tracking) were averaged together. The recordings from three of the subjects (two experimental and one control subject) were contaminated by excessive artifact and were unscorable.

Analysis of the changes in the amplitude of the P3 component suggested that both sleep deprivation and repeated testing were determining factors. Figure 1 illustrates data obtained from ERP Only trials. A sleep deprivation effect is indicated by the lower amplitudes for the Experimental group relative to the Control group on Day 2 (sleep deprivation) but not Day 1 (no sleep deprivation). This observation was supported by a group (2) by days (2) by time of day (5) analysis of variance which yielded a group by day interaction, $F(1,35)=6{\cdot}87$, $p<0{\cdot}05$. The Geisser-Greenhouse conservative F test (Kirk 1982) was used here and in other analyses with repeated measures. An effect of repeated testing is suggested by the observation that P3 amplitudes of both the Experimental and Control group dropped sharply across test blocks on the first day of testing.

Event-related potential latencies were also affected by the test conditions. Figure 2 presents latency data obtained from the experimental subjects for each of the components (P1, N1, P2, N2, P3) across the 12 four-h test blocks for the ERP Only trials (no concurrent task) as well as the ERP/tracking trials (concurrent task). Inspection of these data reveal that latencies of the later components (P2, N2, and P3) increased across test blocks and that the increases were more pronounced when ERPs were obtained while subjects were performing the tracking task (left side of figure 2). P3 latencies under these conditions

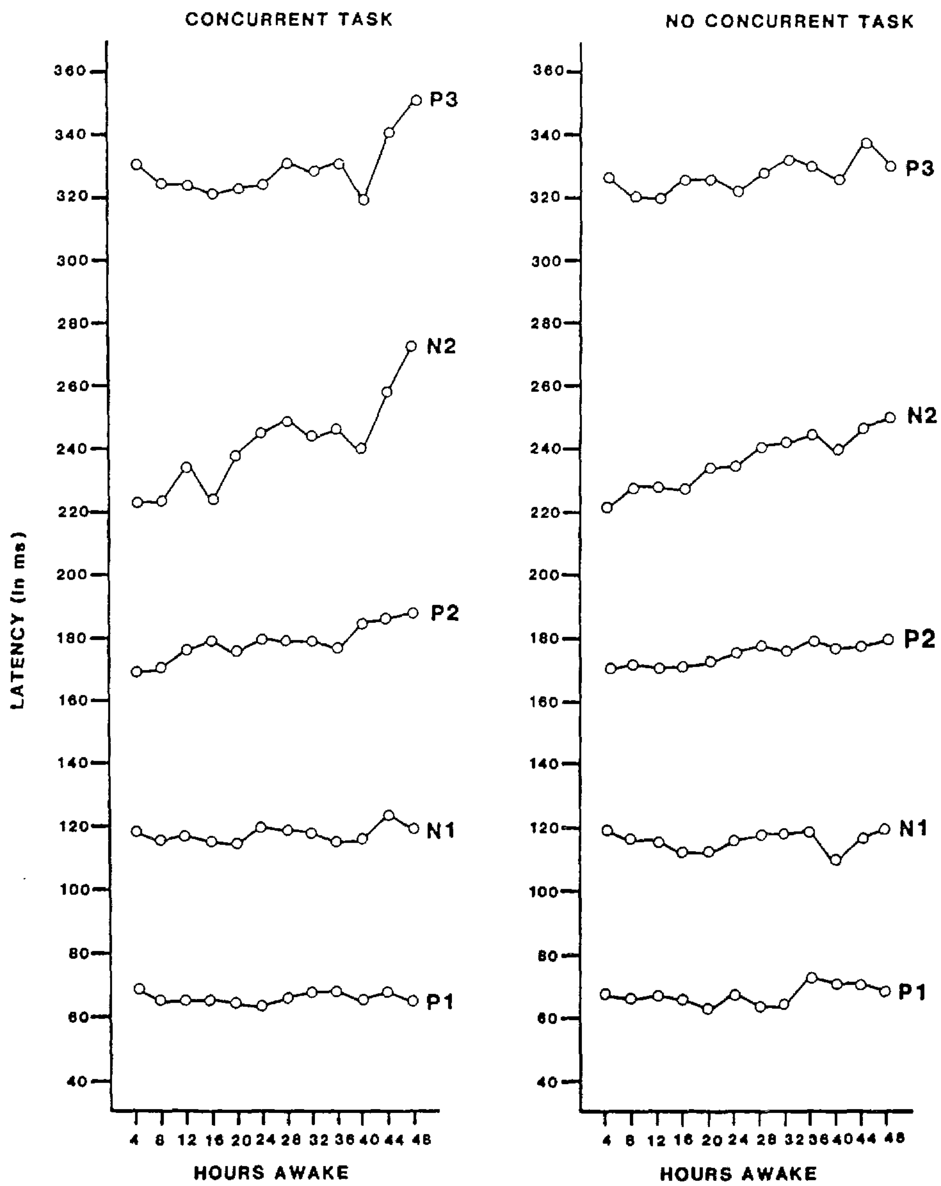

Figure 2. Mean latency (in ms) for each test block for each ERP component for the Experimental group.

remained between 320 and 340 ms through all of the first test day (first six test blocks) and most of the second. There was a 30 ms increase in latencies, however, during the last two test blocks of the 48-h session. A day (2) by block (6) analysis of variance revealed a significant effect of days, $F(1, 27) = 13{\cdot}06$, $p < 0{\cdot}01$; blocks, $F(5, 135) = 4{\cdot}43$, $p < 0{\cdot}01$; and a day by block interaction, $F(5, 135) = 2{\cdot}83$, $p < 0{\cdot}05$. The interaction effect can be attributed to the sharp increase in latencies during the last two test blocks of day 2. The Control group mean latencies (not shown) remained between 320 and 340 ms across all test blocks.

A sleep deprivation effect is more apparent in the N2 latencies. It can be seen that N2 latencies for the Experimental group were relatively stable during the first four test blocks

(223 ms on three of the four blocks) of the first day, and then increased approximately 20 ms over the last two test blocks. The same pattern is evident for Day 2 except, in addition to generally longer latencies, there is a 30 ms increase during the last two test blocks. This pattern suggests both a sleep deprivation effect and a time-of-day effect. The same pattern was not seen in the control group data (not shown). The latencies for the control group were stable (usually between 240 and 250 ms) until the very last test block of Day 2 at which time the mean increased approximately 15 ms. A day (2) × block (6) analysis of variance of Experimental group N2 latencies revealed a significant day effect $F(1,27) = 74{\cdot}75$, $p < 0{\cdot}001$, confirming the longer latencies on Day 2; and a significant block effect, $F(5,135) = 11{\cdot}18$, $p < 0{\cdot}001$, confirming the time-of-day effect. Statistical analysis of the N2 latencies obtained from the Control group provided no evidence of systematic latency changes.

3.2. *Behavioral measures*

3.2.2. *Performance Assessment Battery (PAB)*

In the analysis of the throughput measure for the performance tests of the PAB, a predeprivation baseline was established for each subject by finding the mean of the first four test blocks. Statistical analysis was then performed on the percentage change from baseline of the scores from each test block. Percentage change values for each of the tasks for each of the 12 test blocks for the Experimental group only are shown in figure 3.

Consistent with the expected deprivation effect there was, on the average, a marked deterioration of performance for five of the six tasks (all except the MAST6), especially toward the end of the second day of deprivation. For the same five tasks, performance was poorest during the last block of sleep deprivation (hours 44–48) with performance decrements ranging from 15% to 35% below baseline.

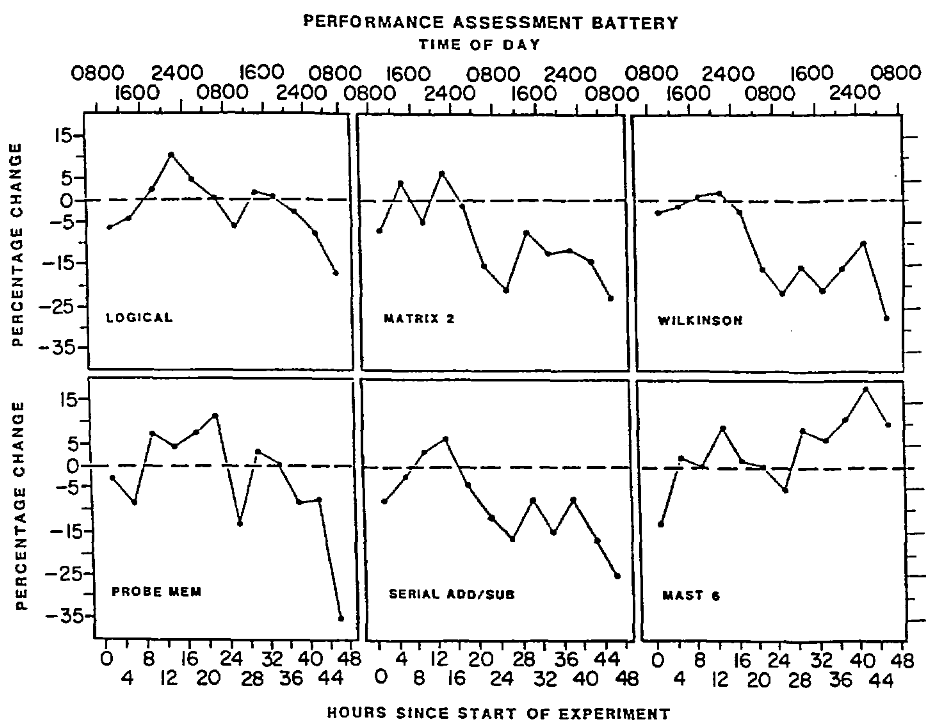

Figure 3. Mean throughput (percentage change from baseline) for each of the tasks of the Walter Reed performance assessment battery across the 12 4-hr test blocks. Experimental group only.

It is apparent from figure 3, that responding on the five tasks was related to time of day. The lowest level of performance on both days tended to occur during the late night (0400–0800) and early morning (0800–1200) hours. There was also a tendency for peak performance to occur during the evening hours (2000–2400) on the first day and the early afternoon hours (1200–1600) on the second day.

An exception to the above was performance on the MAST6 task. Although there appeared to be a time-of-day effect consistent with that seen for the other tasks, performance on Day 2 was higher rather than lower than on Day 1. The unexpected increase in performance on Day 2 was found for both the Experimental and Control groups.

A Day (2) by block (6) analysis of variance of the throughput measure for each of the tasks for the Experimental group supported the observations made above with significant or marginally-significant day effects for all except the MAST6 [Logical – $F(1, 29) = 3{\cdot}51$, $p = 0{\cdot}07$; Probe-Mem – $F(1, 29) = 15{\cdot}47$, $p = 0{\cdot}01$; Matrix 2 – $F(1, 29) = 6{\cdot}20$, $p = 0{\cdot}02$; Add/Sub – $F(1, 29) = 14{\cdot}72$, $p = 0{\cdot}01$; Wilkinson – $F(1, 29) = 58{\cdot}31$, $p = 0{\cdot}01$] due to lower performance on the second day and significant block effects for all but the Probe-Mem task [Logical – $F(1,145) = 2{\cdot}62$, $p = 0{\cdot}05$; Matrix 2 – $F(1, 145) = 3{\cdot}69$, $p = 0{\cdot}01$; Add/Sub – $F(1, 145) = 5{\cdot}17$, $p = 0{\cdot}01$; Wilkinson – $F(1, 145) = 13{\cdot}15$, $p = 0{\cdot}01$] indicating performance changes across time of day. *Post hoc* tests revealed that significant day by block interactions for the Logical task, $F(1, 145) = 2{\cdot}99$, $p < 0{\cdot}03$, and Wilkinson task, $F(1, 145) = 2{\cdot}73$, $p < 0{\cdot}05$, can be attributed to an increase in performance from the 2000–2400 h to the 0400–0800 h test block on one of the days and a decrease on the other. For the Probe-Mem task, the interaction effect $F(1, 145) = 4{\cdot}06$, $p < 0{\cdot}01$. was due to performance increasing during the three test blocks from 1200 h to 2400 h on the first day and decreasing during these same test blocks on the second day.

A group (2) by day (2) by block (4) analysis of variance of throughput was used to compare the performance on the PAB of the experimental subjects with the control subjects. Only blocks 1 through 4 of Days 1 and 2 were used in the analysis because the control group subjects slept during blocks five and six. Of interest in this analysis was whether a single night of sleep deprivation would result in the Experimental group showing lower daytime performance relative to a Control group which was repeatedly tested but allowed to sleep. A groups X days interaction was found for the Logical task, $F(1, 38) = 14{\cdot}07$, $p < 0{\cdot}001$, the Wilkinson task, $F(1, 38) = 20{\cdot}23$, $p < 0{\cdot}0001$, and the Add/Sub task, $F(1,38) = 3{\cdot}56$, $p < 0{\cdot}07$. Additional analyses showed that, for these three tasks, the performance of the groups differed on the second day but not the first. No significant effects were found on the remaining tasks.

3.3. *Correspondence between the ERP and behavioral measures*

The foregoing analyses showed changes in the ERP and performance measures in association with sleep deprivation and time of day. To assess whether the ERP and behavioral measures changed similarly as a function of sleep deprivation and time of day, changes in the 12 block means for the throughput measures of the PAB (data from the MAST6 were excluded) were studied in relation to changes in the 12 block means for each of the ERP components using correlation coefficients. Table 1 presents the results of the analyses using the ERP components of the ERP/tracking trials. Although no corrections were made for conducting multiple comparisons, several patterns can be observed. First, the statistically significant correlations involve the later components (P2, N2, P3) as opposed to the earlier ones (P1 and N1). Second, the signs of the coefficients are negative for all of the latency measures. This suggests longer evoked potential latencies are associated with lower

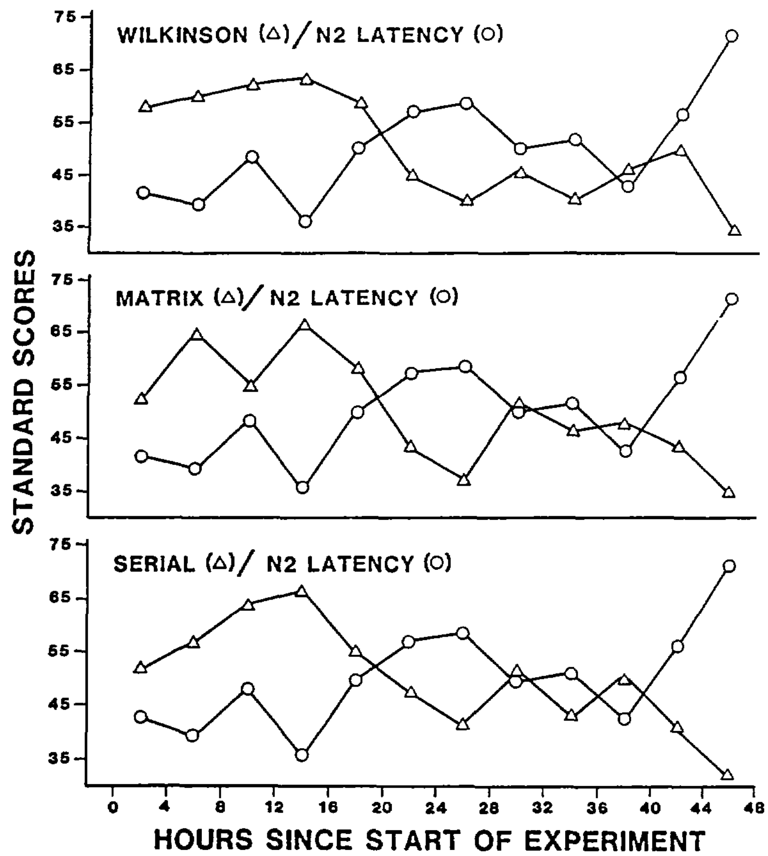

Figure 4. N2 latency and throughput measures for the Wilkinson, Matrix, and Serial tasks of the PAB. The scores are expressed as standard scores with a mean of 50 and a standard deviation of 10.

levels of performance. Finally, higher P3 amplitudes are significantly related to higher levels of performance for each of the tasks.

A pictorial representation of the evoked potential/performance relationship is presented in figure 4. The latency of N2 and the throughput measures on the Wilkinson, Matrix, and Add/Sub tasks of the PAB were converted to standard scores and then plotted on the same scale (figure 4). The standard score formula was:

$$Z_i = 50 + 10(X_i - \bar{X})/s$$

where

Z_i = standard score
Z_i = raw score (a block mean)
$\bar{X}$ = mean of raw scores (mean of block means)
s = standard deviation of raw scores (block means)

The figure shows quite clearly the extent of correspondence between N2 latency and performance; longer N2 latencies are associated with lower performance.

Table 1. Correlation coefficients describing the relationship between the 12 block means for the ERP components (latency and amplitude) and the 12 block means (% from baseline) for the PAB tasks.

	Absolute amplitudes				
	P1	N1	P2	N2	P3
Matrix	−0·10	−0·16	−0·42	0·04	*0·68*
Wilk	−0·09	−0·08	−0·54	0·09	*0·65*
Probe	0·19	0·47	−0·25	0·06	*0·69*
Serial	0·20	0·20	−0·42	−0·10	*0·70*
Logic	0·29	0·34	−0·19	−0·16	*0·63*

	Peak-to-peak amplitudes			
	P1N1	N1P2	P2N2	N2P3
Matrix	−0·13	−0·40	−0·51	*0·59*
Wilk	−0·09	−0·45	*−0·63*	*0·59*
Probe	0·36	0·05	−0·25	0·53
Serial	0·24	−0·21	*−0·66*	0·53
Logic	0·36	0·08	−0·32	0·39

	Latencies				
	P1	N1	P2	N2	P3
Matrix	−0·27	*−0·62*	*−0·62*	*−0·86*	*−0·67*
Wilk	−0·23	−0·32	*−0·61*	*−0·80*	*−0·58*
Probe	−0·32	−0·22	−0·46	*−0·58*	*−0·77*
Serial	−0·40	−0·51	−0·54	*−0·86*	*−0·80*
Logic	−0·42	−0·51	−0·34	*−0·63*	*−0·83*

Note: All italicized values have $p<0{\cdot}05$ (df=10).
ERP data are from the ERP/tracking trials.

4. Discussion

4.1. *Event-related potentials and sleep loss*

Several changes in evoked potentials were observed across the 48-h testing period of the present experiment. These included decreases in the amplitide of the P3 component and increased in the latency of P2, N2, and P3. The findings are generally in accordance with the existing literature. Other investigators have found increased latency and amplitude changes of ERP components in association with sleep loss and/or excessive sleepiness (e.g., Broughton *et al.* 1982, Peeke *et al.* 1980, Pressman *et al.* 1982). It should be noted that the pattern of changes described here does not exactly replicate that described by others. The increase in P3 latency with sleep loss, for example, has not been previously described. Also, changes in the N1 component described by others (e.g., Peeke *et al.* 1980, Gauthier and Gottesman 1983) were not found here. These differences may in large part be due to procedural differences (cf., Pressman *et al.* 1982).

A time-of-day effect was apparent in the measures of N2 latency, P3 amplitude, and to a lesser extent, the measures of other components. There was a strong tendency for P3 amplitudes to decrease and N2 latencies to increase during the night-time and/or morning test blocks on both test days. Changes in the opposite direction tended to occur during the afternoon and early evening hours. Diurnal variation in ERPs has been described in several studies (e.g., Kerkhof 1982, Wesensten *et al.* 1989). As with the sleep loss effects, the pattern of ERP changes with time of day has varied from study to study – perhaps as a result of procedural factors.

An important finding of the present study was that, in general, the evoked potentials obtained while subjects performed the tracking task were more sensitive to the test conditions than the ERPs recorded without the tracking task. That is, as sleep deprivation increased the changes observed with the ERP/tracking trials tended to parallel more closely the changes observed in performance. Researchers have shown in other contexts that evoked potential measures are more sensitive to environmental and task manipulations when subjects are concurrently performing two or more tasks (e.g., Kramer *et al.* 1981). The present findings suggest that a concurrent-task paradigm be used in future studies of the relationship between ERPs and performance degradation.

The present findings also suggest using a more extended sleep loss period in future continuous wakefulness studies involving evoked potential measures. That is, for most of the measures, it was apparent from inspection of the data that the greatest changes occurred in the last four to eight hours of the 48-h test period. Thus, even more dramatic ERP changes and greater ERP/performance correspondence may have been observed if subjects had been deprived of sleep for an additional 24 hours.

It is important to note that although several of the changes observed were unambiguously related to sleep deprivation, not all changes observes in evoked potentials could be attributed to sleep loss. That is, control subjects who were allowed to sleep, showed changes across the test conditions which were similar, but usually smaller, to those seen with experimental subjects. One possible explanation is that control subjects may have experienced some sleep loss due to sleep disruption when awakened for testing purposes. Although the control subjects were permitted sleep, their sleep was interrupted at 0400 h for recording of evoked potentials. There were no apparent effects of this disruption, however, on the performance measures of the control subjects.

It is also possible the ERP changes were, in part, due to habituation resulting from repeated testing. For example, although sleep-deprived subjects had lower P3 amplitudes than the control subjects, it was apparent that reductions in amplitudes began on the first day (no sleep loss) for both groups. A 'repeated testing' effect on P3 amplitudes has been more clearly shown in Wesensten *et al.* (1989).

4.2. *Performance and sleep loss*

As in earlier studies, the present study found that sleep loss resulted in performance degradation. This was evident with all of the PAB tasks used in the present experiment with the exception of the MAST6 task. There was also evidence of a strong time-of-day effect on performance across the two days and nights of the experiment. Generally, performance was lower during the night and morning tests on both Day 1 and Day 2. Further, there was a tendency for performance to be higher in the afternoon and early evening tests. We should note the performance trough occurred during the 0400 h testing block. Others have reported similar findings. It is difficult to attribute the marked performance decrement to any one factor since body temperature, sleepiness, and other circadian factors may be involved (cf., Rutenfranz *et al.* 1972).

There was no clear evidence that any of the tasks were more or less sensitive to sleep deprivation. Although the amount of change at a given time period may have been greater for one task than another, the time point at which changes began to appear was typically the same from task to task (cf., Thorne *et al.* 1983).

4.3. *Correspondence between ERPs and performance*

The systematic correspondence between certain ERP components and performance measures may hold promise for predicting performance degradation. Finding that the changes in performance across the 12 test blocks were related to changes in evoked potentials across the test blocks (i.e., both co-vary with sleep deprivation) suggest that knowledge of evoked potentials can be used to make predictions about performance. For example, it was found that during the 48-h test period, relatively long N2 latencies, whether due to sleep deprivation, circadian rhythms, or other factors, could be used to predict relatively low performance. The findings regarding the N2 component were surprising. This component of the wave form appeared to be a better predictor of performance than did P3. Others have also investigated the relationship of the N2 component to the P3 component (Michalewski *et al.* 1986). These researchers found a strong co-varying relationship between N2 and P3 components with N2 accounting for 61% of the variance of the P3 latency. They also report that the highest correlation between peak latency and reaction time was found for N2; the next highest was P3.

The evoked response/performance relationship should be interpreted cautiously. Other factors which degrade performance may not systematically affect ERP measures. It is important that relationships between ERPs and performance degradation be explored under other conditions of stress.

Acknowledgement

This research was supported by US Army contract DAMD17-84-C-4084.

References

BROUGHTON, R., LOW, R., VALLEY, V., DACOSTA, B. and LIDDIARD, S. 1982, Auditory evoked potentials compared to performance measures and EEG in assessing excessive daytime sleepiness in narcolepsy-cataplexy, *Electroencephalography and Clinical Neurophysiology*, **54**, 579–582.

DONCHIN, E., KARIS, D., BASHORE, T. R., COLES, G. H. and GRATTON, G. 1986, Cognitive psychophysiology and human information processing. In M. G. H. Coles, E. Donchin and S. W. Porges (eds), *Psychophysiology: Systems, Processes, and Applications* (Guilford Press, New York), pp. 244–267.

FRUHSTORFER, H. and BERGSTROM, R. 1969, Human vigilance and auditory evoked responses, *Electroencephalography and Clinical Neurophysiology*, **27**, 346–355.

GAUTHIER, P. and GOTTESMANN, C. 1983, Influence of total sleep deprivation on event-related potentials in man, *Psychophysiology*, **20**, 351–355.

HAKKINEN, V. and FRUHSTORFER, H. 1967, Correlation between spontaneous activity and auditory evoked responses in the human EEG, *Acta Neurology Scandinavia*, **43**, 160–161.

HILLYARD, S. and HANSEN, J. C. 1986, Attention: Electrophysiological approaches. In M. G. H. Coles, E. Donchin and S. W. Porges (eds), *Psychophysiology: Systems, Processes, and Applications* (Guilford Press, New York), pp. 227–243.

KERKHOF, F. 1982, Event-related potentials and auditory signal detection: Their diurnal variation for morning-type and evening-type subjects, *Psychophysiology*, **19**, 94–103.

KIRK, R. E. 1982, *Experimental Design: Procedures for the Behavioral Sciences* (2nd edn)(Brooks Cole Publishing Co, Belmont, CA).

KRAMER, A., WICKENS, C., VANASSE, L., HEFFLEY, E. and DONCHIN, E. 1981, Primary and secondary task analysis of step tracking: An event-related potentials approach. In R. C. Sugarman (ed.), *Proceedings of the 25th Annual Meeting of the Human Factors Society* (Rochester, New York).

MALMO, R. 1959, Activation: A neurophysiological dimension, *Psychological Review*, **66**, 367–386.

MICHALEWSKI, H., PRASHER, D. and STARR, A. 1986, Latency variability and temporal interrelationships of the auditory event-related potentials (N1, P2, N2, and P3) in normal subjects, *Electroencephalography and Clinical Neurophysiology*, **65**, 59–71.

PEEKE, S., CALLAWAY, E., JONES, R., STOVE, G. and DOYLE, J. 1980, Combined effects of alcohol and sleep deprivation in normal young adults, *Psychopharmacology*, **67,** 279–287.

PRESSMAN, M., SPIELMAN, A., POLLAK, C. and WEITZMAN, E. 1982, Long-latency auditory evoked responses during sleep deprivation and in narcolepsy, *Sleep*, **5,** 147–156.

RUTENFRANZ, J., ASCHOFF, J. and MANN, H. 1972, The effects of a cumulative sleep deficit, duration of preceding sleep period and body temperature on multiple-choice reaction time. In W. P. Colquhoun (ed.), *Aspects of Human Efficiency: Diurnal Rhythm and Loss of Sleep* (English Universities Press, London), pp. 217–229.

THORNE, D., GENSER, S., SING, H. and HEGGE, F. 1983, Plumbing human performance limits during 72 hours of high task load. In *Proceedings of the 24th DRG Seminar on the Human as a Limiting Element in Military Systems* (Defence and Civil Institute of Environmental Medicine, Toronto), pp. 17–40.

THORME, D., GENSER, S., SING, H. and HEGGE, F. 1985, The Walter Reed performance assessment battery, *Neurobehavioral Toxicology and Tetatology*, **7,** 415–418.

WEITZMAN, E. and KREMEN, H. 1965, Auditory evoked responses during different stages of sleep, *Electroencephalography and Clinical Neurophysiology*, **18,** 65–70.

WESENSTEN, N., BADIA, P. and HARSH, J. 1989, Repeated testing, interblock intervals, and time-of-day effects on P300 amplitude (manuscript submitted for publication).

WICKENS, C., ISREAL, J. and DONCHIN, E. 1977, The event-related cortical potential as an index of task workload. In A. S. Neal and R. F. Palasek (eds), *Proceedings of the Human Factors Society 21st Annual Meeting* (Human Factors Society, Santa Monica, California).

WILLIAMS, H., TEPAS, D. and MORLOCK, H. 1962, Evoked responses to clicks and electroencephalographic stages of sleep in man, *Science*, **138,** 685–686.

Manuscript received 21 July 1987
Manuscript accepted 28 October 1988