U.S. Department of Transportation
Federal Highway Administration

EYE-ACTIVITY MEASURES OF FATIGUE AND NAPPING AS A FATIGUE COUNTERMEASURE

FINAL REPORT

FHWA-MC-99-028 **January 1999**

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-RD-		

4. Title and Subtitle	5. Report Date
Eye-Activity Measures of Fatigue and Napping as a Fatigue Countermeasure	
	6. Performing Organization Code
7. Author(s) SEE PREFACE	8. Performing Organization Report No.
9. Performing Organization Name and Address Trucking Research Institute 2200 Mill Road Alexandria, VA 22314	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. DTFH61-96-00022
12. Sponsoring Agency Name and Address Office of Motor Carrier Research and Standards Federal Highway Administration, U. S. D.O.T. 400 7th Street, S.W. Washington, DC 20590	13. Type of Report and Period Covered Final Report September 1997-February 1998
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This study investigated the potential use of an eye tracking system for detecting reduced driver alertness, and the impact of prophylactic napping on driver performance and alertness.

The study used traditional behavioral and physiological measures of alertness. In addition, an unobtrusive eye tracker attached to the simulator structure was used to measure eye and eyelid behavior. The results showed clear time-of-day (TOD) and time-on-task (TOT) effects for the following eye closure measures: Partial closures during fixations, speed of slow eyelid closure (SEC), blink duration, and blink frequency.

Eye closures during fixations exhibited the following alertness monitoring characteristics: 1) the cyclic phases of a driver experiencing brief lapses of alertness and recovery; 2) a continuous decline ultimately leading to an off-road simulator crash; 3) an early warning potential of 10 minutes or more; 4) a dramatic decline in the measure beginning 2-3 minutes before an off-road simulator crash. SEC events and blink duration showed sustained increases with TOT and TOD. A preliminary algorithm for detecting level of alertness was developed. This algorithm uses the eye closure measure in a way that includes partial eye closures during fixations, blink frequency, blink duration, and speed of eye closure effects.

The 3-hour afternoon nap increased the subjects' nighttime alertness and improved driving performance. Beneficial effects of the afternoon nap on nighttime driving performance included significantly fewer crashes, shorter run completion times, and smaller standard deviations of lane position. The results provide evidence that the 3-hour afternoon nap was effective in reducing sleepiness levels during the following night and suggest that prophylactic naps may be more beneficial than recuperative naps during all-night driving situations.

17. Key Words	18. Distribution Statement
Eye tracking, ocular dynamics, fatigue, alertness, simulator, napping	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	71	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

This document is the final report for a research project conducted under Contract No. DTFH61-96-00022 between the U.S. Department of Transportation's Federal Highway Administration (FHWA) and the Trucking Research Institute (TRI) of the ATA Foundation. The TRI engaged the services of several organizations to carry out the project. The project was organized and managed by Applied Science Group (ASG) Inc., Bedford, MA, which also provided and supported the eye tracking instrumentation and contributed to processing and analysis of eye tracker data. The ASG project manager was Mr. Thomas Dukes.

The Institute of Circadian Physiology (ICP), White Plains, NY, recruited subjects, provided and supported instrumentation for electroencephalogram (EEG) and electrooculogram (EOG) measurements, administered performance and fatigue assessment tests, and contributed to most aspects of data processing and analysis. The ICP project leader was Dr. Ziad Boulos.

The truck simulation facility and subject facility were provided and supported by the Liberty Mutual Research Center, Hopkinton, MA., which also contributed to processing and analysis of simulator data. The Liberty Mutual effort was led by Mr. Thomas Ranney.

Thanks are owed to Professor. John Stern, of Washington University, St. Louis, MI, for his specific contributions to processing and analysis of EOG and eye tracker data and for his general advice on all aspects of the study.

The project also had the benefit of support and advice from Dr. William Rogers, of the Trucking Research Institute, a research affiliate of the ATA, and Dr. Ronald Knipling and Mr. Robert Carroll of the Federal Highway Administration, Office of Motor Carriers.

TABLE OF CONTENTS

FIGURES and TABLES

LIST OF ABBREVIATIONS

A / D	Analog to Digital
ASG	Applied Science Group
ASL	Applied Science Laboratories
ATA	American Trucking Association
BL	Base Line Drive Session
CR	Corneal Reflection
BMI	Body Mass Index
CMV	Commercial Motor Vehicle
DATLOS	Data Loss
DFAS	Driver Fatigue and Alertness Study
DOT	Department of Transportation
ECG	Electrocardiograms
EEG	Electroencephalograms
EMG	Electromyograms
EOG	Electrooculograms
ePERCLOS	Estimate of Percent of Eye Closure
EYENAL	Eye Movement and Pupil Diameter Data Analysis software
EYEPOS	Eye Position and Pupil Diameter Computation software
FFT	Fast Fourier Transform
FHWA	Federal Highway Administration
FMO	Floor Mounted Optics
Hz	Hertz
IBM	International Business Machines
ICP	Institute for Circadian Physiology
LMRC	Liberty Mutual Research Center
MAST	Memory and Search Test
MHz	Megahertz
MPH	Miles Per Hour
MSPD	Mean Speed
NIH	National Institutes of Health
OMC	Office Of Motor Carriers
PAB	Walter Reed Performance Assessment Battery
PC	Personal Computer
PERCLOS	Percent Eye Closure
POBOT	Post-Blink Opening Time
POG	Point Of Gaze
PRBCT	Pre-Blink Closing Time
PSD	Psychiatric Sleep Disorder
PSG	Polysomnograph

REM	Rapid Eye Movement
RT	Response Time
R&T	Research And Technology
SA	Sleep Apnea
SDL	Scenario Definition Language
SEC	Slow Eye Closure
SLP	Standard Deviation Of Lane Speed
SSPD	Standard Deviation
SSTR	Standard Deviation Of Steering Wheel Position
STI	Systems Technology Incorporated
STISIM	Systems Technology Incorporated Simulation Software
TOD	Time-Of-Day
TOT	Time-On-Task
TRI	Trucking Research Institute
TST	Total Sleep Time
VAS	Visual Analog Scale
V / H	Vertical Pupil Diameter/Horizontal Pupil Diameter
XDAT	External Data Input/Output

EXECUTIVE SUMMARY

A six-month pilot study investigated the potential use of an eye tracking system for detecting reduced driver alertness, and the impact of prophylactic napping on driver performance and alertness. Nine professional truck or bus drivers with at least two years of long-haul driving experience participated in the study designed to examine nighttime driving performance using a truck simulator. The study used traditional behavioral and physiological measures of alertness. In addition, an unobtrusive eye tracker attached to the simulator structure was used to measure eye and eyelid behavior.

OCULAR DYNAMICS AS A PREDICTOR OF REDUCED ALERTNESS

Eye closure measures showed the most convincing relationship to reduced alertness and confirm the possibility of monitoring driver alertness with eye dynamics measures. The results show clear time-of-day (TOD) and time-on-task (TOT) effects for the following eye closure measures:

- Partial closures during fixations
- Speed of slow eyelid closure (SEC)
- Blink duration
- Blink frequency

The results of saccade frequency measures, although inconclusive, showed sufficient change in the minute before and after simulator crashes to warrant further analysis in a subsequent study. Additionally, extreme head and body movements, although not specifically measured, were also observed to be possible indicators of driver fatigue and deserve further study.

The most impressive measure of reduced alertness was the partial eye closures during fixations. This measure exhibited the following valuable alertness monitoring characteristics:

1) The cyclic phases of a driver experiencing brief lapses of alertness and recovery

2) A continuous decline ultimately leading to an off-road simulator crash

3) An early warning potential of 10 minutes or more

4) A dramatic decline in the measure beginning 2-3 minutes before an off-road simulator crash

SEC events and blink duration showed sustained increases with TOT and TOD. Both parameters demonstrate short-term utility in predicting fatigue, particularly in the final minute preceding an off-road simulator crash.

Blink frequency showed a favorable although less consistent TOT and TOD effect.

A preliminary algorithm for detecting level of alertness is proposed. This algorithm uses the eye closure measure in a way that includes partial eye closures during fixations, blink frequency, blink duration, and speed of eye closure effects. All of these are potentially measurable with an unobtrusive eye tracking system. The results suggest that this eye closure measure provides an effective starting point for the development of the optimal fatigue detection algorithm. However, while not presently included, it is anticipated that the optimal algorithm will also include point-of-gaze measures (as defined by future data analysis) for predicting reduced alertness.

Additional testing and analysis beyond the scope of this study is required to verify the alertness detection results and to verify, enhance, and implement the proposed detection algorithm. The study has provided excellent guidance for the design of follow-on efforts to further investigate these very promising alertness measurement parameters, and has provided valuable information for the design of eye tracking systems to be used for alertness monitoring in vehicles.

PROPHYLACTIC NAPPING AS A FATIGUE COUNTERMEASURE

The three-hour afternoon nap increased the subjects' nighttime alertness. Beneficial effects on nighttime performance were evident as follows:

- Lower subjective fatigue and sleepiness scores
- Improved results on the Performance Assessment Battery tests
- Improved performance on the driving simulator
- Lower spectral EEG amplitudes

Performance on the computerized tests was generally faster, more accurate, and/or less variable in the Nap condition. However, differences between conditions were only statistically significant ($0.05<p<0.1$) for one measure in the first night test session, and approached statistical significance for several other measures in the last two sessions.

Beneficial effects of the afternoon nap on nighttime driving performance included significantly fewer crashes, shorter run completion times, and smaller standard deviations of lane position. Performance on two secondary driving tasks was also generally better in the Nap condition, although the differences did not reach statistical significance.

Spectral amplitudes in the theta and 6-10 Hz electroencephalograph (EEG) frequency bands during the night driving runs were significantly lower in the Nap than in the No-Nap condition. A trend toward lower EEG activity was also observed in the alpha frequency band. Activity in these frequency bands is known to increase as a function of drowsiness, and theta activity is also one of the signs of sleep onset. Thus, these results provide additional evidence that the three-hour afternoon nap was effective in reducing sleepiness levels during the following night.

Three of the subjects were unable to start or complete the third night driving run and were allowed to take an unscheduled nap instead, both in the Nap and No-Nap conditions. Interestingly, the unscheduled naps, which lasted one-two hours, had few beneficial effects. For example, subjective sleepiness and fatigue levels remained almost as high as they were at the end of the second night driving run. More importantly, driving performance was worse during the fourth driving run, after the unscheduled nap, than it was during the second. In addition, even after the unscheduled naps, driving performance during Run 4 remained better and spectral EEG amplitudes remained generally lower in the Nap than in the No-Nap condition.

These latter results suggest that prophylactic naps may be more beneficial than recuperative naps during all-night driving situations. Other studies have shown that naps taken early within a long period of sustained wakefulness have more beneficial effects than do ones taken at later times. This question could be tested experimentally in future studies of scheduled napping as a fatigue countermeasure in long-haul driving operations.

I. INTRODUCTION

A. PROBLEM STATEMENT

The commercial motor vehicle (CMV) drivers may drive up to 10 hours continuously before taking a break, often drive at night, and sometimes have irregular and unpredictable work schedules. Much of their mileage is compiled during long trips on Interstate and other four-lane roadways. Because of their far greater mileage exposure and other factors, commercial drivers' risk of being involved in a fatigue-related crash is far greater than that of non-commercial drivers – even though CMV drivers represent only about 4% of the drivers involved in known fatigue-related crashes and their rate of involvement per mile traveled is no greater than that of non-commercial drivers.

Fatigue is directly implicated in other studies as accounting for between 5 and 10% of all truck accidents. It is for this reason that driver fatigue is identified as a top priority CMV safety issue by the Federal Highway Administration (FHWA) (Knipling and Wang, 1994).

Diminished driver attention may be caused by a variety of factors. These factors include emotional stress, boredom, environment, sleep history, food history, drugs, alcohol, or general fatigue. With the CMV industry providing 24-hour per day service, there are additional pressures on CMV drivers contributing to fatigue. CMV drivers may need to drive continuously for up to 10 hours, may need to drive at night, and may suffer from irregular and unpredictable work schedules. Their off-duty activities and sleep habits play a key role in determining what their alertness and performance level will be while on-duty driving.

Since 1990, a number of studies and research programs have been conducted, by both commercial and academic organizations, to address the issue of drowsiness/fatigue in drivers. The FHWA Office of Motor Carriers (OMC) has more than 20 Research and Technology (R&T) projects that were either recently completed, presently on-going, or planned, relating to driver drowsiness/fatigue and hours of service. This includes the comprehensive over the road Driver Fatigue and Alertness Study (DFAS) involving the Essex Corporation, Transport Canada, and the Trucking Research Institute (TRI) of the American Trucking Associations Foundation (ATA). Also of significance are the studies conducted by Dr. Walter W. Wierwille and associates at Virginia Tech. The goal of these efforts has been to identify countermeasures to combat driver fatigue and develop a vehicle based capability to continuously monitor driver fatigue.

B. RESEARCH GOALS

The FHWA contracted, through the TRI, Applied Science Group (ASG) and its subcontractor Institute for Circadian Physiology (ICP) to conduct this pilot study. The pilot study examined the feasibility of using eye tracking technology to create an algorithm that reliably predicts decrements in alertness. At the same time, the study provides preliminary data on the effectiveness of an afternoon nap in offsetting fatigue associated with nighttime hours of service. A counterbalanced, within subjects design was used to assess the effects on driver performance of an afternoon nap following a truncated night's sleep (midnight to 5:00 a.m.) versus a 5-hour sleep period with no nap. In both conditions, electroencephalograms (EEG) and four eye movement variables as well as outputs from the simulator were measured throughout a 10-hour nighttime driving task.

The specific objectives of this feasibility study were as follows:

1. To prove the concept of utilizing specific measures of ocular dynamics, singly and in combination, as a reliable method for predicting decrements in driver alertness and performance.
2. To assess the relative importance/effectiveness of parameters in four measurement categories (blink patterns, saccadic eye movements, pupil dynamics, and scan patterns) and to do preliminary development of a predictive algorithm
3. To assess the influence of an afternoon nap on alertness and performance following a night of truncated sleep.

The data for this study were collected between September 1997 and February 1998.

II. METHODS

A. SUBJECTS

Nine professional drivers (eight males, one female; ages 29 - 49) were recruited from the metropolitan Boston area. Subject characteristics are summarized in Table 1. Eight of those selected were truck drivers and one was a bus driver; all had at least 2 years long-haul driving experience (range 2-24 years). The subjects gave written informed consent prior to the start of the study, and were paid for their participation. The Institutional Review Boards of ICP and of Liberty Mutual Research Center (LMRC) approved the study protocol.

B. EXPERIMENTAL DESIGN

The study used a counterbalanced crossover design with two conditions, one with a scheduled 3-hour nap on the afternoon preceding a night of driving (Nap condition) and one without (No-Nap condition). The order of conditions was chosen at random, except for one subject whose initial condition was changed from No-Nap to Nap after the start of that condition (see Results). A 1-week interval separated the start of the first and second conditions, except for two subjects (Subjects 6 and 7) who began the second condition 4 weeks after the start of the first. One subject (Subject 8) withdrew from the study after completing one condition only (Nap condition); data from that subject were excluded from most analyses.

Following an initial telephone interview and screening, prospective subjects met with an experimenter and filled out a general health questionnaire (adapted from the Standard Shiftwork Index, Folkard et al. 1995), the Sleep Disorders Questionnaire (Douglass et al. 1986, 1994), and the Morningness-Eveningness Questionnaire (Horne & Ostberg 1976). During the week preceding each condition, the subjects were required to fill out a daily log indicating times spent asleep, awake, and at work. The study was conducted at LMRC in Hopkinton, MA, where the subjects remained for the duration of each condition (about 47 hours for the first condition and 43 hours for the second). In most cases, the subjects were driven from Boston to the laboratory in Hopkinton and back.

The study protocol is summarized in Figure 1 and Table 2. Day 1 of the first condition served to thoroughly familiarize the subjects with the testing facilities, and to give them extensive training on the performance test battery and the driving simulator. This included two uninterrupted 2-hour practice driving runs, the last of which was performed with the subjects wearing EEG and electrooculograms (EOG) electrodes.

Table 1. Subject Characteristics

No	Initials	First condition	Sex	Age	Truck /bus	CDL	Years driving	Years long-haul	Comments
1	DB	non	M	49	T	A	2	1994-96	
2	JG	nap	M	37	T	B	6	1987-89	
3	RA	non	M	46	T	B	10	1990-97	
4	ES	nap	M	47	B	A	5	1991-96	Night session 3 incomplete. Slept 4-5 AM both conditions.
5	CH	nap	M	41	T	A	10	1986-92	Night session 3 not done. Slept 3-5 AM both conditions.
6	HH	non	F	33	T	Class 1	10	1980-90	Night session 4 not done in nap condition due to back pain.
7	JA	non	M	40	T	B	24	1994-97	Night session 3 not done. Slept 2:25-4:25 AM (no-nap). Slept 4:15-5-15 AM (nap).
8	WM	nap	M	29	T	A	10	1987-97	Did nap condition only, slept 15 min. during night session 3, did not complete night session 4.
9	TH	nap	M	34	T	B	17	1985-98	

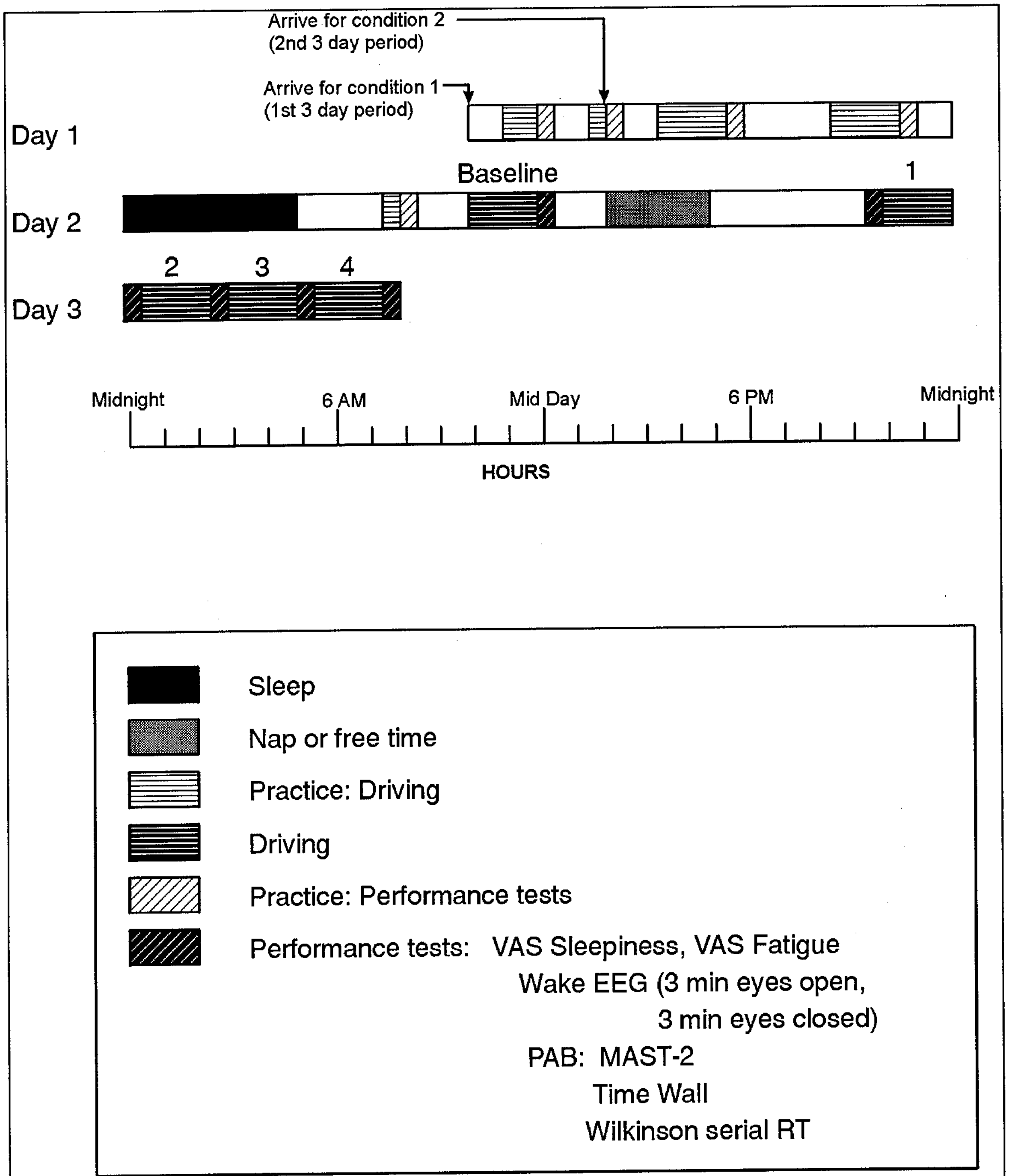


Figure 1. Experimental Protocol

Table 2. Experimental Protocol

DAY 1:		
	10:00	Arrival at Truck Simulator Laboratory in Hopkinton
	10:00-11:00	Familiarization with testing facilities
	11:00-12:30	1st training session: driving simulator (1 hour) and performance testing
	12:30-13:00	Lunch
	13:30-15:00	2nd training session: driving simulator (30 min.) and performance testing
	15:30-18:00	3rd training session: driving simulator (2 hours) and performance testing
	18:30-19:00	Dinner
	19:30-20:30	EEG wire-up
	20:30-23:00	4th training session: driving simulator (2 hours), EEG and performance
DAY 2		
	24:00-05:00	Sleep/darkness
	05:00	Fill out sleep quality questionnaire
	05:00-06:00	Electrodes removed, shower
	6:00-06:30	Breakfast
	07:30-08:30	5th training session: driving simulator (30 min.) and performance testing
	09:00-10:00	EEG wire-up
	10:00-12:00	Baseline session: driving simulator (2 hours)
	12:00-12:30	Baseline session: EEG and performance testing (20 min.)
	12:30-13:00	Lunch
	14:00-17:00	Nap/darkness or free time
	17:00	Fill out sleep quality questionnaire, electrodes removed (Nap condition
	18:30-19:00	Dinner
	20:30-21:30	EEG wire-up
	21:30-22:00	EEG and performance testing (20 min.)
	22:00-24:00	Driving simulator (2 hours)
	24:00-24:30	EEG and performance testing (20 min.)
	24:30-02:30	Driving simulator (2 hours)
	02:30-03:00	EEG testing (10 min.) and snack
	03:00-05:00	Driving simulator (2 hours)
	05:00-05:30	EEG and performance testing (20 min.)
	05:30-07:30	Driving simulator (2 hours)
	07:30-08:00	EEG and performance testing (20 min.)
	08:00	Electrodes removed, subject driven back to Boston

The subjects slept in a windowless bedroom from 24:00 until 05:00 of Day 2, when they were awakened by an experimenter. After an additional, short training session in the morning, the subjects completed a 2-hour baseline driving run starting at 10:00. Immediately after the driving run, while the subjects were still in the driving seat, they filled out two Visual Analog Scales (VAS) for assessing subjective sleepiness and fatigue. Electrode impedance was checked and electrodes were reapplied when necessary. Resting EEG recordings were then obtained while the subjects remained seated with eyes open and fixating on a blank screen for 3 minutes, followed by 3 minutes with eyes closed. They then moved to an adjacent room where they completed three computerized performance tests. Completion of the VAS, EEG and performance (see Section II.H) testing took 15 to 20 minutes.

In the Nap condition, the subjects were required to remain in bed, in darkness, from 14:00 until 17:00, whether or not they were able to sleep. In the No-Nap condition, they spent that time engaged in sedentary activities of their choice. The schedule then called for VAS, EEG and performance testing starting at 21:30, and four 2-hour driving runs (Night Runs 1-4) starting at 22:00, 00:30, 03:00 and 05:30, respectively. Each 2-hour driving run was followed by a 30-minute interval during which the subjects completed VAS, EEG and performance testing (performance testing was omitted from the 02:30-03:00 interval to allow the subjects sufficient time for a snack).

In the second condition, the subjects arrived at LMRC at 14:00 instead of 10:00 (less training was required in this condition than in the first), and the first training session took place from 15:00-18:00 (driving simulator for 2 hours, performance testing for 1 hour). From then on the schedule was the same as during the first condition.

Caffeinated beverages were only allowed before 18:00 on Day 1, and before 07:30 on Day 2. Smoking was allowed outdoors, when it did not interfere with scheduled sleep, driving or testing periods.

C. TRUCK DRIVING SIMULATOR

A fixed-base driving simulator, located at the LMRC for Safety and Health, was used to replicate the essentials of driving in a medium-size heavy truck. The cab included a standard air-suspension truck seat, throttle and brake pedals modified to provide the feel of heavy truck controls, and a heavy truck steering wheel. A turn signal lever and horn button, mounted on the dashboard, were used to record drivers' responses for target-detection tasks. Two rear-view truck mirrors (405 mm x 152 mm) were mounted on the outside of the cab, one on each side.

The simulation software (STISIM, v.7.03), developed by Systems Technology Inc. (STI), ran on an 80486DX2 computer. A Pro Audio Spectrum sound card created simulated engine and horn sounds. An optical encoder interpreted signals from the vehicle controls, and a ComputerBoards CIO-DIO24 controlled the mirror-image presentation.

The roadway scene was projected onto a 1.83 m x 2.43 m wall-mounted screen, using a ceiling-mounted NEC MultiSync projection monitor. The distance between the subject's eyes and the projection screen was 3.35 m, allowing the eyes to focus at infinity. The scene changed dynamically in response to steering and pedal inputs from the subject. The display consisted of a 2-lane roadway divided by a broken yellow line. A digital speedometer, indicating speed in miles per hour, was projected onto the bottom right corner of the roadway display. A trip odometer, indicating the number of miles driven since the beginning of each run, was presented at the bottom left corner of the roadway display.

Vehicle-control parameter settings and simulation scenarios, which consist of a sequence of discrete events (e.g. oncoming vehicles, curves), were created using Scenario Definition Language (SDL). Luminance of scene components and ambient lighting in the lab were selected to represent nighttime driving on 2-lane rural roads.

Images of simulated following vehicles were presented on 21 cm x 27.5 cm monitors (Commodore 1942). The images were created with two Amiga 1200 computers. The 'folded' distance from the subject's eyes to the mirror to the target image, was approximately equal to the distance between the subject's eyes and the projected roadway image (described below), allowing subjects to move their eyes between the two images without refocusing.

Closed-circuit video cameras, located above the projection screen and behind the participant's head, were used to view the participant and the roadway display, during the experiment.

Driving Scenarios

During each driving run, participants encountered a sequence of interactive driving events, which required control inputs, and non-interactive events, which did not require control inputs (see Appendix A, Table A1). Session length was determined by distance, which was set at approximately 110 miles and was generally completed in a little under 2 hours. Interactive events included straight and curved road segments, signalized and non-signalized intersections, work zones, oncoming vehicles, some of which encroached into the drivers' lane, vehicles stopped in the drivers' lane, cross winds, and 2 secondary target-detection tasks. The secondary tasks required participants to detect pedestrians appearing alongside the roadway and specially designed target images presented in the rear view mirrors. Non-interactive events, which included buildings in the distance, planes in the sky, and trees alongside the roadway, were included to provide variation in the general level of visual stimulation during different segments of the runs.

Run Structure

Each 2-hour run was divided into four 30-minute blocks, which were matched in terms of event frequency and driver workload. Each 30-minute block was divided into three 10-minute segments, in which the level of task demand was varied. High-demand segments contained a higher frequency of events than medium and low-demand segments, and the events themselves tended to be more demanding. For example, the curves were sharper and the intersections required stop/go decisions. In addition, work zones, which required slowing and driving through narrow gaps, were included in the high-demand segments. The 10-minute segments were presented in a different order during each 30-minute block. Frequencies of different events for the three levels of task demand are presented in Appendix A, Table A2.

Vehicle Control Measures

Multiple measures of vehicle control were sampled continuously throughout each drive. Measures included steering wheel, throttle, and brake positions, speed in miles per hour (mph), and lateral position within the lane. Vehicle control data were sampled approximately 30 times per second.

Pedestrian Response Task

Subjects were required to respond to gray stick-figure pedestrians by pressing the horn button located to the right of the steering wheel. Responses made within 3 seconds of the appearance of pedestrians were considered correct detections. Responses made after 3 seconds were considered misses.

Pedestrians varied in brightness and were presented at different positions alongside the roadway, at an apparent distance of 700 feet. As the vehicle neared the stationary pedestrian, the image increased in size. If the driver responded within 3 seconds following the appearance of the pedestrian on the screen, the horn sounded and the pedestrian remained stationary on the screen until it passed out of view as the driver passed it. If the driver failed to respond within 3 seconds, the pedestrian began walking toward the roadway. If the driver responded to the moving pedestrian, the horn sounded and the pedestrian reversed direction and moved away from the roadway. If the driver failed to respond, the pedestrian continued walking toward the roadway. Depending upon the pedestrian's initial position and the speed of the vehicle, the driver may be required to execute an evasive maneuver to avoid colliding with the pedestrian. Specifically, pedestrians that started at the "near" position were more likely than pedestrians starting at the "far" position to walk into the vehicle's path. Pedestrian targets were presented on straight roadway segments only, and only in the absence of other interactive events. Half of all pedestrians were presented in the presence of roadside trees, and half were presented in the absence of trees. Pedestrian task independent variables included side of road (left, right), lateral distance from the side of the road (near, far), target-background contrast (high, low), and trees (present, absent). This resulted in a total of sixteen combinations. This design was replicated once every hour of driving, resulting in a total of 32 pedestrian targets in each 2-hour run.

Mirror Response Task

Participants were required to identify target vehicle images presented on monitors located behind the truck cab, which were viewed in the rear-view mirrors. At the beginning of each event, a front-view vehicle image (11.5 cm x 10.5 cm) appeared in one of the mirrors. On the front of each vehicle where six rectangles (24 mm x 13 mm) displayed in a 3 x 2 arrangement. On target-absent trials, all rectangles were complete. Participants were required to respond to target-present vehicles (rectangular outline missing) by activating the turn signal indicator in the direction corresponding to the location of the mirror in which the target vehicle appeared. Participants were instructed to ignore distracter vehicles, for which all six rectangular outlines were complete.

Mirror targets were presented on straight roadway segments only, and only in the absence of other interactive events. Independent variables for the mirror task included the mirror in which the image was presented (left, right) and the location of the incomplete rectangle (six positions). This resulted in a total of twelve combinations of target trials, and six combinations of distracter trials. This design was replicated once every hour of driving, resulting in a total of twenty-four targets and twelve distracters in each 2-hour run.

Reward/Penalty
A monetary reward/penalty system was used to increase subject motivation and encourage good driving performance. The subjects could receive up to $15 per 2-hour run depending on their performance. Subjects earned $5.00 for completion of each 2-hour run. They were also rewarded for finishing a run in less than a reference time (160 minutes), and were penalized for completion times greater than the reference time. Specifically, a prorated reward of $1.00 per minute was given for early completion of a run, and a prorated penalty of $1.00 per minute was given for late completion of a run. Subjects earned a reward of $0.10 for each correct detection of a pedestrian or target vehicle in the mirror. Penalties were given for each collision (with another vehicle, pedestrian or sign) or off-road accident ($5.00), for each speeding ticket ($2.00), for each barrel hit ($0.25), and for each pedestrian or mirror task missed ($0.10). Penalties were also given for each incorrect response to a mirror task ($0.10), and for each false alarm to a pedestrian or mirror task ($0.10).

D. EYE TRACKER

An Applied Science Laboratories (ASL) Model 4250R+ eye tracker system was used to acquire eye measurement data on all test subjects. This system is a general purpose measurement device used primarily for research in visual sciences and medical diagnostics applications. The Model 4250R+ was specifically selected for this study because of its flexibility (supports all advanced optics options), remote (unobtrusive) optics, extended head tracking option, ability to accurately measure a subject's pupil diameter and point of gaze, and suitability for installation and operation in the LMRC driving simulator. While the system is not specifically designed for vehicle applications, it is considered an excellent research tool for measuring parameters of ocular dynamics.

The eye tracker uses the pupil to corneal reflection (CR) technique with bright pupil optics. A 2-axis servo tracking mirror is used to direct the telephoto eye camera view and the coaxial illumination beam as a subject moves about. The mirror is directed by commands from the eye tracker computer so as to continually attempt to re-center the pupil image within the eye camera field of view. A second wide angle camera (locating camera) is aimed along the same optical path as the primary telephoto eye camera. The wide angle image is used for initial manual location of the subject's eye, and to implement an automatic outer tracking loop which acts to re-center the eye image any time it escapes from the telephoto camera field of view. The system is also capable of using a magnetic head tracker to make mirror tracking more robust and to help account for geometric point of gaze measurement errors associated with changes in subject position. The magnetic tracker could not be used for this study, however, because it interfered with the EEG and EOG systems. The system displays a cursor indicating point of gaze superimposed on a video image of the scene being viewed by the subject, and also records digital data (on the eye tracker computer hard disk) indicating pupil diameter and point of gaze coordinates with respect to the scene image.

In order for the eye tracker to account for individual subject differences, an eye calibration procedure is conducted for each subject. For this purpose, a calibration point target display was developed such that it could be generated by the simulator and seen on the large screen display. This display consisted of nine calibration target points. The target points were black dots overlaid on a gray background. The points covered roughly about 80% of the display area and were positioned such that they were vertically and horizontally collinear and perpendicular about a center point. With the scene camera fixed, the location of the nine calibration points on the screen display could be entered into the eye tracker. The subject then proceeded to look (fixate) at each of the nine points as eye tracker measures were made. The raw data measured by the eye tracker consist of the separation between the pupil center and the corneal reflection. Since the relation between these raw values and the eye line of gaze differs for each subject, the calibration procedure was performed just prior to each 2-hour data collection session. The procedure typically takes about 30 seconds.

The eye tracker system components include a floor mounted optics (FMO) module, a scene camera assembly, a control unit, and an IBM compatible personal computer (eye tracker PC) as described below.

Optics Module
The remote optics module consists of a pupil camera, locating camera, illuminator (infrared source), and associated optics. The module is set on a floor mounted and adjustable pedestal. It is located in front of and slightly below the simulator dash board. The pupil camera is used to provide a telephoto image of the eye while the locating camera provides a wide angle image that includes about half of the subject's head. The pupil camera, locating camera, and illumination beam are aimed along the same optical path towards the subject's left eye.

Scene Camera Assembly
The scene camera assembly consists of a video camera and lens mounted on an adjustable tripod. This assembly is positioned behind and slightly above the driver's right shoulder. The camera is used to provide a video image of the same scene that the subject views. The image provides the reference frame for the eye point of gaze measurement.

Control Unit and Computer
The eye tracker control unit and PC provide operator controls and switches, three video monitors, data interfaces, data processing hardware/software, and the data archiving functions. The PC for this study was a 120 MHz Pentium. The control unit, with the help of the computer, processes the eye camera signal to extract the elements of interest (pupil and reflection of the light source on the cornea) and computes both pupil diameter and line of gaze. This data is displayed and relayed to external data ports. Both of these units are located on a table in an adjoining control room and provide the necessary operator interface functions.

In order to facilitate the use of the eye tracker for this application, several minor modifications or custom adjustments were done. First, the standard eye position and pupil diameter computation (EYEPOS)) software for the PC was modified to include vertical pupil diameter measures as well as the standard horizontal pupil measures. Second, the system's external data (XDAT) parallel interface port was configured to accept the digitally formatted event codes provided by the driving simulator. Third, instead of freezing point-of-gaze (POG) values during periods of pupil image loss (e.g., during blinks), POG values were set to an identifiable default position during these periods. In order to smooth measurement noise, the recorded POG value was normally the running average of the current value and the previous three values. When the pupil and CR were recognized following an image loss (e.g., blink), no values during or preceding the loss were included in the average. Fourth, an analog signal proportional to the horizontal POG position was relayed to the EEG equipment and recorded on an extra channel with EEG and EOG measures.

Two separate video recorders were connected to the system and used to record images for subsequent data analysis. The first video recorder was used to record a 4-window video image consisting of the scene camera, the driver's head, the driver's eye, and the simulator scene. The second video recorder was used for the scene with POG cursor superimposed. A time code generator was used to record elapsed time and time of day on the video images.

E. POLYSOMNOGRAPHY

Continuous polysomnographic recordings were obtained during all scheduled sleep episodes. The EEG (bilateral frontal, central, parietal and occipital leads cross-referenced to mastoids), EOG (electrodes positioned above and lateral to one eye and below and lateral to the other), EMG and ECG signals were amplified by a sixteen-channel Nihon-Kohden polygraph (model EEG-4317F) and recorded on a PC using a data acquisition system (Eclipse, Stellate Systems) which digitized the signals on-line at an effective sampling rate of 128 Hz, after low pass filtering with a cut-off frequency at 64 Hz.

F. DRIVING EEG/EOG

EEG, EOG, EMG and ECG were recorded throughout the baseline and nighttime driving sessions. Recording montages were the same as for sleep, except that the EOG was also recorded using separate horizontal and vertical bipolar derivations. Two additional polygraph channels recorded inputs from the horizontal eye position channel of the eye tracker, and from the driving simulator, which generated coded event stamps to allow accurate synchronization of EEG and EOG signals with driving performance.

The signals were recorded with the same equipment and parameters used during sleep, but with a different data acquisition system, better suited to quantitative EEG analysis (Rhythm, Stellate Systems).

G. RESTING EEG/EOG

Resting EEG, EOG, EMG and ECG were recorded before the start of nighttime driving and immediately after each 2-hour driving run. The subjects sat in the driving simulator seat with eyes open and fixating on a blank screen for 3 minutes, followed by 3 minutes with eyes closed. Recording equipment and parameters were the same as during driving.

H. PERFORMANCE ASSESSMENT BATTERY

Three self-paced performance tests, taken from the Walter Reed Performance Assessment Battery (PAB), were administered on a PC before the start of nighttime driving and following each 2-hour driving session, except for Night Session 2. These were: (1) Two-Letter Memory and Search Test, a visual search and recognition task (30 trials), (2) Time Wall, a time estimation task (20 trials), and (3) Wilkinson Four-Choice Serial Reaction Time, a visual-motor coordination task (100 trials).

I. SUBJECTIVE SLEEPINESS AND FATIGUE

These were assessed before the start of nighttime driving and immediately after each 2-hour driving run by means of two 100-millimeter VAS. The sleepiness scale was labeled "Very alert" at one end and "Very sleepy" at the other, while the corresponding labels on the fatigue scale were "Not tired at all" and "Extremely tired".

J. SUBJECTIVE SLEEP QUALITY

The subjects filled out a Sleep Quality Questionnaire (adapted from Gortelmeyer 1985) immediately after the end of each scheduled sleep episode (night sleep and afternoon nap).

III. RESULTS

A. GENERAL OBSERVATIONS

The study protocol proved to be more demanding than anticipated. Although the subjects were professional truck or bus drivers and had considerable all-night driving experience, they found it difficult to complete the four scheduled night driving sessions. They all commented that simulator driving was much less stimulating than real driving, and that the study protocol precluded such activities as drinking coffee, smoking, or listening to the radio, which help them to remain awake during real driving. Indeed, three subjects (Subjects 4, 5, and 7) were unable to start or to complete Driving Run 3 (03:00-05:00), and asked to take a nap instead (see Table 1). They were allowed to sleep for 1-2 hours, after which they returned to the simulator room and completed Driving Run 4. Even subjects who did complete all driving sessions noted that, had they felt as tired while driving a real truck or bus as they did during some of the late night/early morning sessions, they would not have continued driving without first getting some rest.

Two other subjects failed to start or to complete Driving Run 4, in one case (Subject 6) because of back pain from an injury sustained shortly before the start of the Nap condition, and in the other (Subject 8) because of nausea and severe headache, possibly the result of simulator sickness.

Subject 5 complained of insufficient sleep on the night before the start of his first condition (originally scheduled as No-Nap), and was concerned about being able to stay up all night after only 5 hours of scheduled sleep in the laboratory. To increase the chances of this subject completing the protocol, this condition was changed from No-Nap to Nap. Such a change introduces an obvious bias, but in this case, the bias would work against our hypothesis that prophylactic napping improves subsequent alertness and performance.

B. SUBJECT CHARACTERISTICS

Based on the results of the screening questionnaires, the subjects were generally in good physical and mental health, although, based on their Body Mass Index (BMI, listed in Table 3), several were clinically overweight (BMI of 25-29.9, according to recent NIH guidelines) or obese (BMI of 30 and above). Subject 1 was taking medication for borderline high blood pressure (which, however, was within acceptable limits for CDL certification). Eight of the subjects were smokers.

The Sleep Disorders Questionnaire provides a Psychiatric Sleep Disorders (PSD) Index and a Sleep Apnea (SA) Index. All scores on the PSD Index were within normal range, although Subject 2 admitted to occasional sleep disturbances in the preceding few months related to personal matters. Subjects 6 and 7 scored above the 90th percentile on the SA Index (Table 3), which places them at high risk for sleep apnea, while Subjects 2, 5 and 8 were at moderate risk, with scores between the 75th and 90th percentiles (note that the percentile ranking corresponding to a given score on the SA index depends on age and gender).

Scores on the Morningness-Eveningness Questionnaire are listed in Table 3. Four of the subjects were classified as Moderately Morning Type and one as Moderately Evening Type, while the remaining were Neither Type.

Although no single characteristic accurately predicts the subjects' ability to remain awake during the night driving sessions, the results suggest that Morning Type subjects and subjects at higher risk for sleep apnea may have more difficulty staying awake at night.

C. POLYSOMNOGRAPHY (PSG)

PSG recordings were obtained during the two-night sleep episodes and the scheduled daytime nap and analyzed using standard sleep stage scoring criteria (Rechtschaffen & Kales 1968). Table 4 shows sleep onset latency (latency to at least 1.5 consecutive minutes of stage 1 or 20 seconds of any other sleep stage), total sleep time (TST, stages 1, 2, 3, 4, and REM) and sleep efficiency (100 x TST/time in bed from lights out to lights on) for all subjects. Night sleep for Subject 4 in the Nap condition was not recorded due to equipment failure. Sleep was not recorded during unscheduled naps.

There were no significant differences in sleep onset latency or sleep efficiency between the three scheduled sleep episodes (ANOVA, $p > 0.05$). Sleep onset latencies were within normal limits (< 30 minutes) for all subjects except during the nap for Subject 2. Sleep efficiency, however, was frequently below the normal level (< 90%), probably due to the novelty of the sleeping environment. During the 14:00-17:00 naps, the subjects slept on average for 2.35 hours, or about 78% of the scheduled time.

As expected, combined TST in the Nap condition (night sleep plus nap) exceeded TST in the No-Nap condition (night sleep only), but in Subject 2 the difference was only 10 minutes.

The PSG recordings of Subject 7, who was a heavy snorer and at high risk for sleep apnea, showed a consistent pattern of 4-8 seconds of wake (microarousals) approximately once every 30 seconds. Although the present study did not include respiratory monitoring, this PSG pattern is strongly suggestive of sleep apnea. It should be noted that these microarousals, which would represent about 20% of total time in bed, are not taken into account in deriving sleep efficiency measures using standard sleep scoring criteria.

D. SUBJECTIVE SLEEP QUALITY

The Sleep Quality Questionnaire provides measures of subjective sleep quality and of "feeling refreshed" upon awakening. Scores on these two measures for the two scheduled night sleeps and the daytime nap are plotted in Figure 2. No significant differences were found between the two night sleep episodes or between night sleeps and naps for either measure (Wilcox on signed-rank, $p > .05$).

Table 3. Screening Questionnaires

No.	Initials	Sex	Age	M-E score	BMI	SA Index	SA percentile
1	DB	M	49	43.0	23.9	23	50-75th
2	JG	M	37	56.5	30.5	27	75-90th
3	RA	M	46	64.5	22.3	23	50-75th
4	ES	M	47	59.5	26.5	21	50th
5	CH	M	41	53.5	35.2	27	75-90th
6	HH	F	33	34.5	38.7	27	95th
7	JA	M	40	66.5	41.1	34	90-95th
8	WM	M	29	63.5	30.5	27	75-90th
9	TH	M	34	54.5	25.9	19	25-50th

Morningness-Eveningness Questionnaire (Home-Ostberg):

16-30	Definitely Evening Type
31-41	Moderately Evening Type
42-58	Neither Type
59-69	Moderately Morning Type
70-86	Definitely Morning Type

Table 4. Sleep Parameters

Subject	No Nap Condition			Nap Condition						
	Sleep			Sleep			Nap			Combined
	Latency	TST	Efficiency	Latency	TST	Efficiency	Latency	TST	Efficiency	TST
01	6.33	275.00	91.6%	12.00	266.67	88.6%	2.33	159.67	88.4%	426.34
02	26.33	217.33	72.4%	28.00	147.00	48.1%	42.67	80.33	44.4%	227.33
03	16.67	258.67	85.6%	15.33	248.00	82.5%	3.33	59.33	32.8%	307.33
04	4.00	296.00	97.3%	N/A	N/A	N/A	8.00	166.67	91.4%	N/A
05	2.00	281.00	93.6%	2.67	231.00	77.2%	3.00	167.00	92.4%	398.00
06	19.33	262.33	86.8%	11.67	273.67	91.1%	14.67	163.00	90.2%	436.67
07	3.67	285.67	94.9%	7.00	250.67	86.5%	2.33	131.67	73.0%	382.34
08	N/A	N/A	N/A	2.67	267.00	88.8%	1.33	171.33	95.0%	438.33
09	23.00	198.00	65.9%	17.33	272.00	90.6%	1.33	170.67	94.8%	442.67

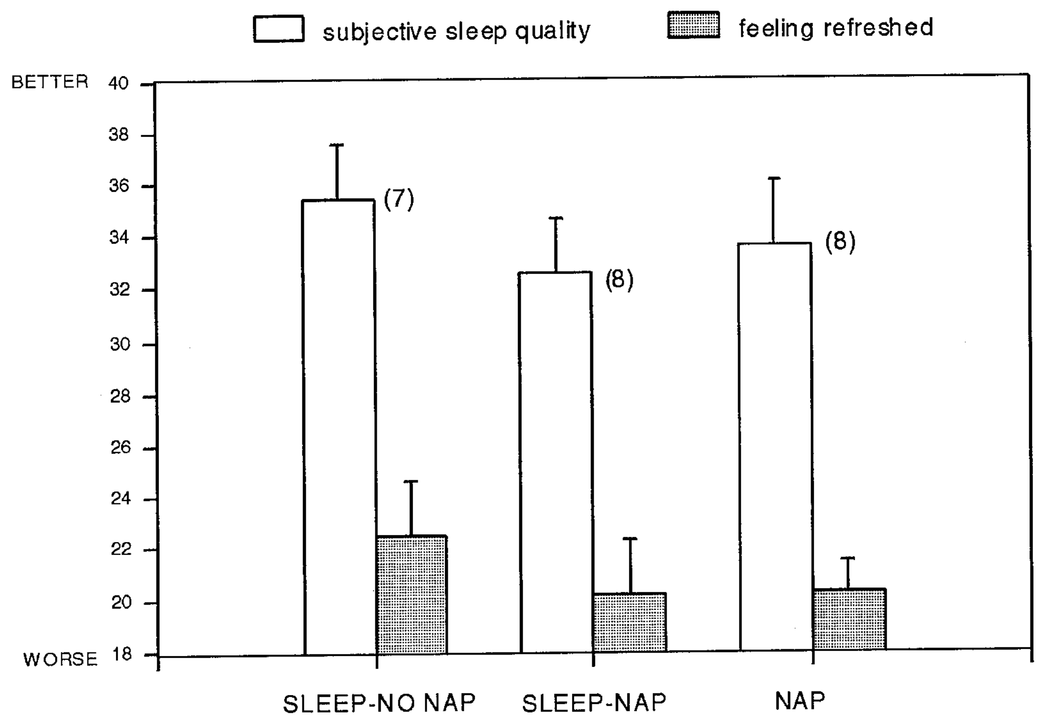


Figure 2. Sleep Quality Questionnaire scores

E. DRIVING PERFORMANCE

The results presented in this section provide a macroscopic analysis of driving performance over the course of the five 2-hour runs. For each measure, summaries of performance were created for each completed 2-hour run. In cases where drivers did not complete entire runs or when partial data were lost, extrapolation was done using the available data. When viewing the figures, the reader should recall that the first run, referred to as the baseline run, was completed during the morning (i.e. 10:00 and 12:00 hours) before the overnight runs. This was before administration of the Nap/No-Nap treatment. The following four runs, which are numbered 1 through 4 in the figures, were run consecutively, beginning 10 hours following completion of the baseline run. Specifically, these runs started at 22:00, 0:30, 3:00, and 5:30 hours, respectively. Half-hour breaks were used for alertness testing and rest. Some subjects did not complete all driving runs. This was most common for the third session, which was scheduled to run between 03:00 and 05:00 hours.

1. Crashes

The majority of crashes were run-off-road crashes, which are characteristic of loss of alertness. Figure 3 presents the mean crash frequencies (based on all crashes) across all subjects by run and condition. In this and the following figures, the standard error is included to represent the between-subject variation associated with the means. The general tendency was for crash frequency to increase over the course of the five 2-hour drives. In general, this was more apparent in the No-Nap condition than in the Nap condition, suggesting a possible benefit associated with the prophylactic nap. Crash frequencies greater than approximately ten per 2-hour run generally indicate significant impairment, most often when the driver was almost asleep.

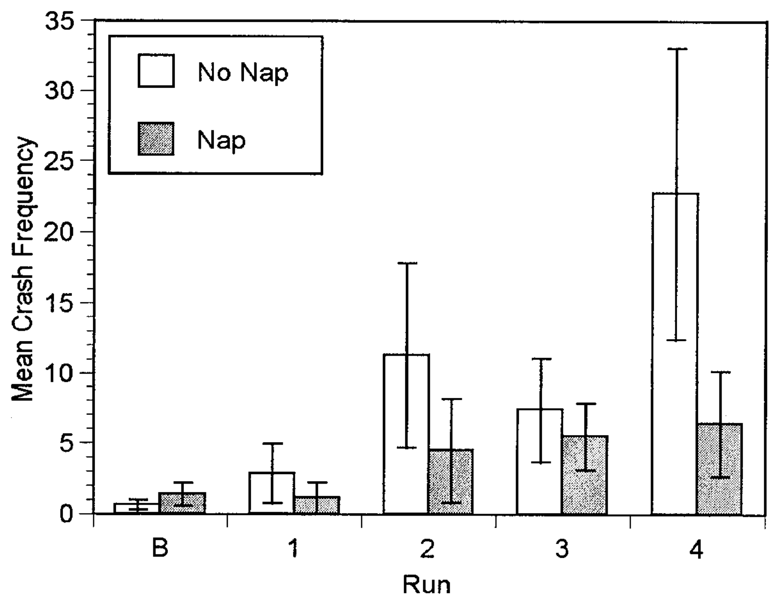


Figure 3. Mean Crash Frequency x Run x Condition (± SE)

Analyses of variance were computed on crash frequency and two transformations of crash frequency. Specifically, square root and log transforms were applied to each crash frequency in an attempt to reduce the skew of the distribution. Square-root transformations are generally recommended for analysis of frequency data (Kirk, 1982; Myers and Well, 1991). Independent variables were condition (two levels: Nap, No-Nap) and run (1,2,3,4). Baseline runs were eliminated from the ANOVAs because they were taken before the nap and thus would not be expected to contribute meaningfully to comparisons of the effects of the nap on driving performance. A summary of the analyses results is presented in Table 5.

Despite the large difference between mean crash frequency values for the Nap and No-Nap conditions (Nap: M = 4.24; SD = 8.10; No-nap: M = 11.78; SD = 19.90), ANOVA results using the untransformed frequencies showed the main effect of condition to be non-significant. As suggested by the large standard deviation values, there were relatively large differences between subjects. The transformations reduced the skew in the frequencies, however the resulting distributions were still formally not normal. As shown in Table 5, the main effect of condition was statistically significant for both analyses using either transformation.

Table 5. Summary of ANOVA Results for Crash Frequency

	Condition		Run		Cond x Run	
Dep. Measure	**F**	**P**	**F**	**p**	**F**	**p**
Crashes	4.14	.08	3.70	.03	3.34	.05
Sqrt (crashes)	6.51	.04	7.76	.00	3.80	.03
Log (crashes)	8.83	.02	12.79	.00	2.42	.11

Because the results of the analyses for the untransformed and transformed data were not consistent, two additional tests were conducted to evaluate the overall effect of the prophylactic nap on crash frequency. First, following Iman, Hora, and Conover (1984), the overall mean crash frequencies for each subject x condition combination were computed and ranked (1-16). These ranks were then subject to the one-way (S x A) ANOVA. Iman, Hora, and Conover (1984) have demonstrated the usefulness of this approach for one-way repeated-measures classifications only. The main effect of condition was statistically significant ($F(1,7) = 6.30$, $p = .04$). Finally, the direction of the difference was considered for each subject, individually. It was found that the overall mean frequency of crashes in the No-Nap condition was greater than the corresponding mean frequency in the Nap condition for seven of eight subjects. When a sign test is performed, this proportion ($p = .88$) was found to be significantly different from $p = .5$ ($p = .035$). The preponderance of evidence supports the conclusion that the subjects experienced more crashes in the No-Nap condition than in the Nap condition.

As shown in Table 5, the main effect of run was statistically significant for all analyses, reflecting the consistent increase observed across the four numbered runs. The Condition x Run interaction was statistically significant for both the raw and square-root transformed data, reflecting the increased crash frequencies in runs 2 and 4, observed in the No-Nap condition, but not in the Nap condition (see Figure 3).

2. Speeding Tickets

The mean frequencies of speeding tickets for all subjects combined, by run and condition are presented in Figure 4. Relative to the overall trend for crashes, it appears that drivers began on average to accumulate large numbers of speeding tickets during the third, rather than final run. However, three subjects did not complete the third run, and the elevated ticket frequencies in the third run, reflect the behavior of only two subjects.

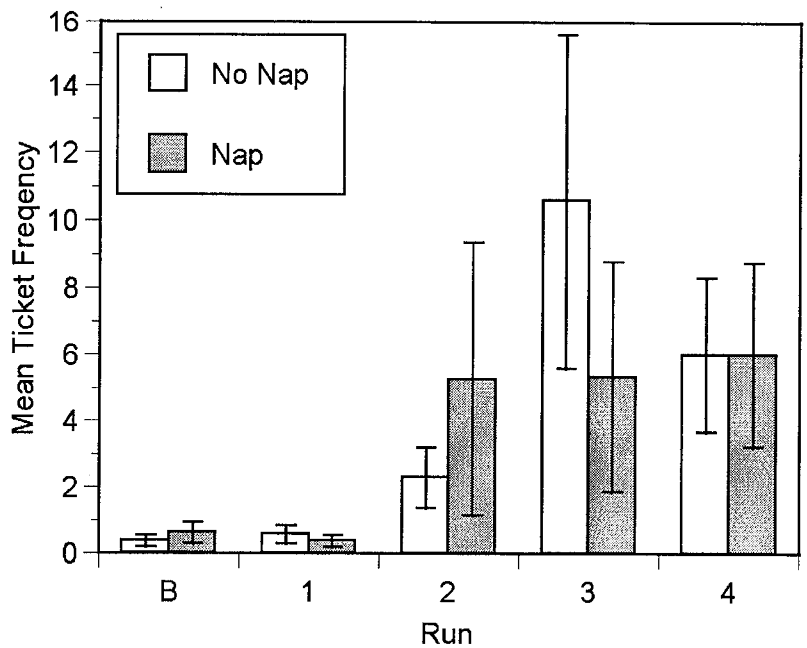


Figure 4. Mean Ticket Frequency x Run x Condition (±SE)

Analyses of variance were computed on ticket frequency and two transformations of ticket frequency. Specifically, square root and log transforms were applied to each ticket frequency in an attempt to reduce the skew of the distribution. Independent variables were condition and run. A summary of the analyses results is presented in Table 6.

The results consistently indicate no difference between the Nap and No-Nap conditions in the frequency of tickets (Nap: M = 4.10; SD = 8.35; No-Nap: M = 4.48; SD = 7.59).

The main effect of run was statistically significant only for the log transformed frequencies. The Condition x Run interaction was not statistically significant for any of the analyses.

Table 6. Summary of ANOVA Results for Ticket Frequency

	Condition		**Run**		**Cond x Run**	
Dep. Measure	**F**	**p**	**F**	**p**	**F**	**p**
Tickets	.01	.94	1.49	.25	.66	.59
Sqrt (tickets)	.11	.75	2.62	.08	0.27	.84
Log (tickets)	0.42	.54	4.28	.02	0.04	.99

As with crash frequencies, large numbers of speeding tickets generally reflect significant impairment. When a significant number of tickets occurred in the absence of crash increases, the subject may have intentionally decided to ignore the monetary penalties associated with the speeding tickets, in an attempt to complete the 2-hour drive as quickly as possible. However, instances of either large numbers of tickets or large numbers of crashes reflect a deviation from the intended responses to the driving scenario. These instances clearly reflect the point in time at which the driver would prefer not to be driving and hopefully the point at which, in the real world, the driver would have discontinued driving.

3. Run Completion Time

The mean adjusted completion time by run and condition is presented in Figure 5. Completion times faster than the pre-established reference time are reflected as negative numbers, while positive numbers reflect slower than expected performance. For interpretation of these data, it is important to note that the adjusted completion time is influenced by the crash frequency, since following each crash, there is a delay during which the vehicle is returned to the center of the lane and more importantly, the vehicle must start from a stopped position. This correlation is most apparent in the No-Nap condition in the fourth run. This is the only run by condition combination for which the mean adjusted completion time is positive and as shown in Figure 3, the mean crash frequency was greatest for this combination.

On average, the drivers were faster in completing the driving runs in the Nap condition than in the No-Nap condition. As shown in Figure 5, this trend was apparent for runs 2 through 4.

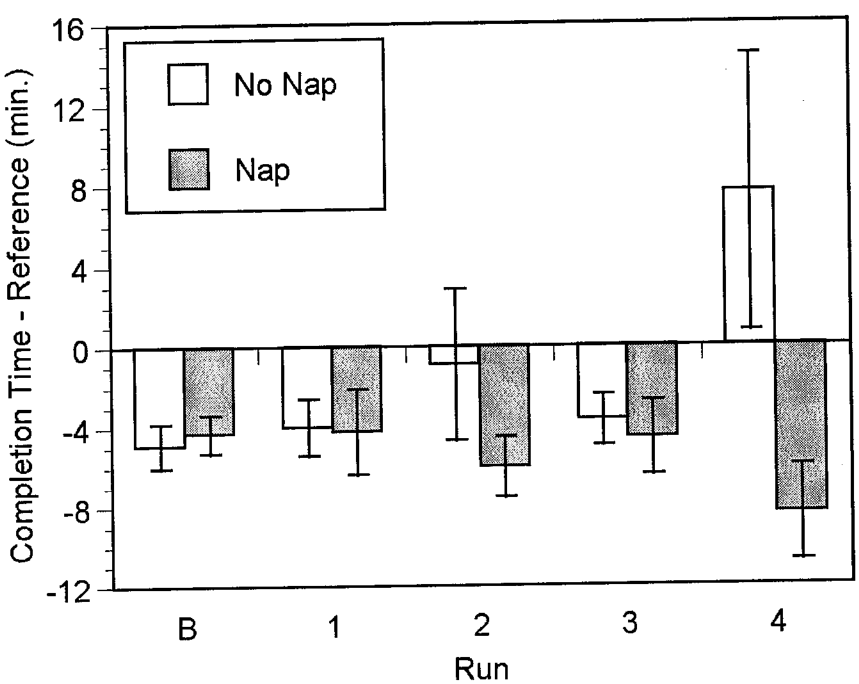


Figure 5. Mean Adjusted Completion Time x Run x Condition (±SE)

An analysis of variance was computed on completion times. Independent variables were condition and run. A summary of the analyses results is presented in Table 7.

Table 7. Summary of ANOVA results for Completion Time

	Condition		**Run**		**Cond x Run**	
Dep. Measure	**F**	**p**	**F**	**p**	**F**	**p**
Completion time	5.47	.05	2.85	.07	4.18	.02

The difference between the means for the two conditions is approximately 4 minutes (Nap: M = -3.6, SD = 7.3; No-Nap: M = 0.3, SD = 12.6), which represents approximately 3 % of the 118-minute reference time for completing a driving run. As shown in Table 7, the main effect of condition is marginally significant, reflecting the large individual variation between subjects. Consideration of the individual means for the two conditions revealed that seven of eight subject means were faster for the Nap condition relative to the No-Nap condition. This difference is statistically significant, when subjected to a sign test. The Condition x Run interaction was significant for this measure. Consideration of the means revealed a large increase in completion time between the third and fourth runs in the No-Nap condition (No-Nap Condition: Run 3: M = -3.7; Run 4: M = 7.6), which was not apparent in the Nap condition (Nap Condition: Run 3: M = -4.5; Run 4: M = -2.9). This result reveals a significant deterioration in performance in the fourth run in the No-Nap condition that was not apparent in the Nap condition.

4. Total Monetary Reward

The total of all monetary rewards and penalties is a summary measure of overall driving performance. The reward/penalty system was designed so that subjects could nominally earn approximately $10-$20 for good performance during a 2-hour run (the maximum they could actually receive was set at $15 per run). Poor overall performance, including a large number of tickets and/or crashes could result in a large negative value for the total monetary reward. Because the subjects realized that they would not actually be required to pay the negative amount, the negative values should be interpreted with caution. Large negative values generally indicate a disruption of normal driving. In the present context, large negative values are likely to indicate either significant impairment or a deliberate strategic shift, whereby the driver sacrificed one part of the driving task (e.g. speeding tickets) in order to complete the run as quickly as possible. It is likely that they reflect the point in time at which subjects would have chosen not to continue driving under real-world conditions.

Mean overall monetary rewards by condition and run are presented in Figure 6. The drivers were able to perform effectively enough to obtain positive rewards during the baseline condition. The average monetary reward was also positive during the first run in the Nap condition, but slightly negative in the No-Nap condition. The incidence of positive monetary rewards in runs 2 through 4 was minimal. For these three runs, the average monetary penalties were greater in the No-Nap condition than in the Nap condition.

An analysis of variance was computed on completion times. Independent variables were condition and run. A summary of the analysis results is presented in Table 8.

The main effect of condition did not reach statistical significance, despite the large difference between the condition means (Nap: M = -19.4, SD = 58.8; No-Nap: M = 60.4, SD = 114.0). The relatively large standard deviations reflect large differences among the subjects. Consideration of overall condition means for each subject revealed that seven of eight subjects received more total monetary compensation in the Nap condition than in the No-Nap condition. As indicated above, this proportion (.88) is significantly greater than .5 and, according to a sign test, indicates a significant difference between the conditions. The significant Condition x Run interaction reflects a large decrease in total monetary reward between the third and fourth runs in the No-Nap condition that was not apparent in the Nap condition. According to this measure, performance was generally much worse during the fourth run in the No-Nap condition than in the corresponding run in the Nap condition.

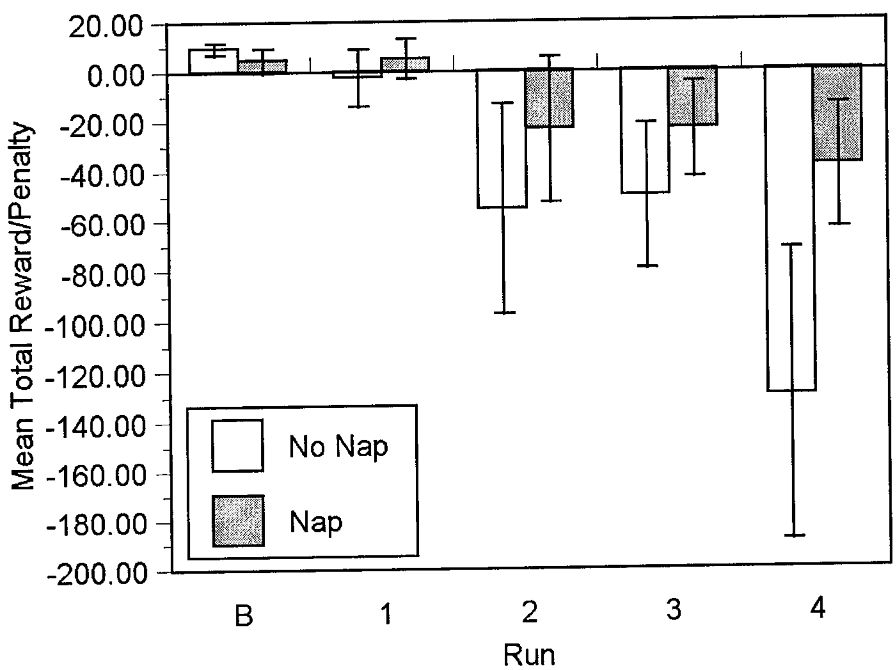


Figure 6. Mean Total Reward / Penalty (LS)

Table 8. Summary of ANOVA results for Total Monetary Reward

	Condition		**Run**		**Cond x Run**	
Dep. Measure	**F**	**p**	**F**	**p**	**F**	**p**
Total Monetary Reward	4.55	.07	3.95	.02	3.38	.04

The main effect of run was statistically significant for this measure, reflecting the significant deterioration in performance between the first and fourth run (Run 1: M = 1.3 SD = 27.5; Run 4: M = -87.5, SD = 133.1). Overall means for the second and third runs were approximately equal (Run 2: M = -38.3 SD = 94.6; Run 4: M = -36.0, SD = 59.9).

5. Mirror Target-Detection

Each 2-hour run contained 36 mirror task trials, including 24 targets and 12 distracters. Performance is measured both by the proportion of targets correctly detected and by the response time for target detection. Figure 7 presents the mean percentages of mirror targets detected by condition and run. Overall, there appeared to be a slight deterioration in performance, as reflected by the approximately 10% decrease in the proportion of targets detected over the course of the nighttime runs. This deterioration was somewhat more pronounced in the No-Nap condition than in the Nap condition.

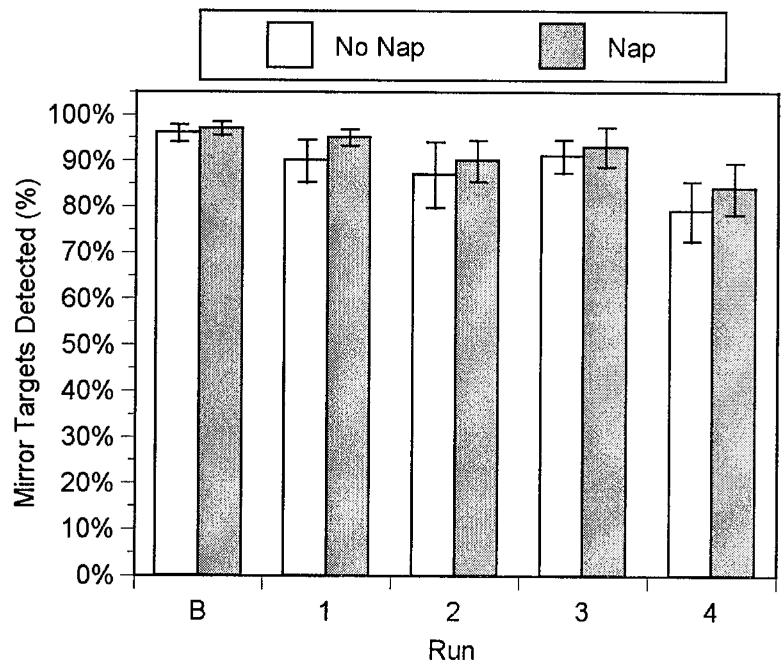


Figure 7. Mean Mirror Target Detection x Run (LS)

Analyses of variance were computed on the proportions of mirror targets detected, the log transformation and the arc sin transformation (Emerson, 1991). Independent variables were condition (two levels: Nap, No-Nap), run (1,2,3,4), and Condition x Run. A summary of the analysis results is presented in Table 9.

The results indicated that the nap had no effect on the proportion of mirror targets detected. The main effect of run was significant across analyses, reflecting the progressive decrease in proportion of mirror targets detected. There was no interaction between condition and run.

Table 9. Summary of ANOVA Results for Mirror Detection Proportions

	Condition		Run		Cond x Run	
Dep. Measure	**F**	**p**	**F**	**p**	**F**	**p**
p (mirror)	1.60	.25	4.03	.02	0.39	.76
Log [p(mirror)]	1.28	.30	4.75	.01	0.62	.61
Arcsin (p[mirror])	2.43	.16	3.28	.04	0.38	.77

Mean response times by condition and run are presented in Figure 8. Mirror target-detection response times increased over the four nighttime runs, in both conditions. The means were consistently faster in the Nap condition than in the No-Nap condition. This difference is also apparent for the baseline run, which is noteworthy because the baseline run occurred before the Nap/No-Nap treatment. The only difference between the conditions at this point was that the subjects knew whether or not they would be having a nap in the afternoon.

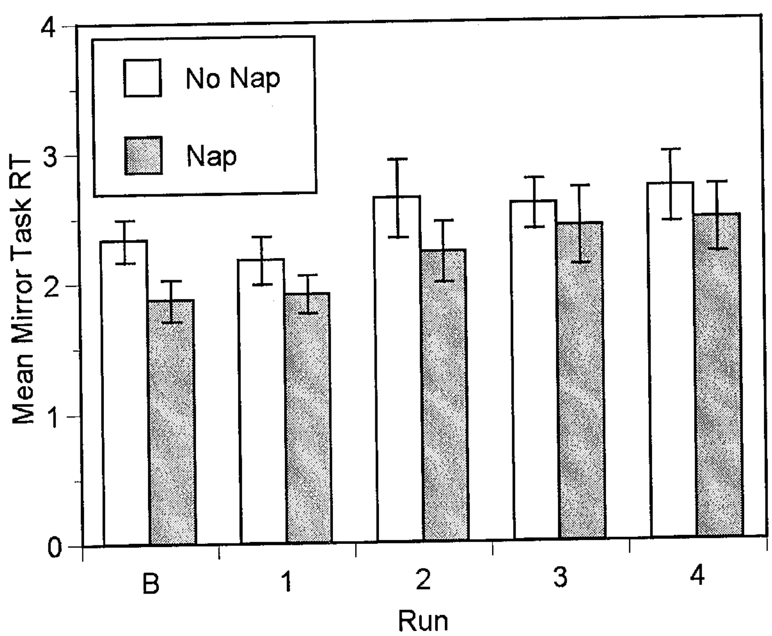


Figure 8. Mean Mirror Task RT x Run x Condition (±SE)

Analyses of variance were computed on the mirror target-detection response times (RTs) and two transformations. Independent variables were condition and run. A summary of the analysis results is presented in Table 10.

The main effect of condition was not statistically significant, indicating no effect of the nap on mirror task response time. The main effect of run was significant for one transformation. The Condition x Run interaction was not statistically significant.

Table 10. Summary of ANOVA Results for Mirror Task Response Time

Dep. Measure	Condition F	Condition P	Run F	Run p	Cond x Run F	Cond x Run p
RT	2.69	.15	2.05	.14	2.21	.12
Log (RT)	3.39	.11	1.37	.28	1.81	.18
Square root (RT)	2.43	.16	3.28	.04	0.38	.77

6. Pedestrian Target-Detection

Each 2-hour run contained 32 pedestrian targets. Performance is reported both by the percentage of targets detected and by the response time for target detection. The pedestrian task actually involved two stages of target detection: stationary target detection and moving target detection. Pedestrians remained stationary for 3 seconds and then began moving. The mean percentages of pedestrians detected are presented in Figure 9.

The bars are composed of three segments. The lower, black, segment represents the proportion of pedestrian targets detected within 3 seconds, while the pedestrian was stationary. The middle, gray, section represents the proportion of pedestrians detected after the pedestrian had started moving. The top, white, segment refers to the percentage of pedestrians that were missed altogether. The results show a general reduction over time in the proportion of stationary pedestrians detected in both the Nap and No-Nap conditions. The percentage detected during the fourth run is slightly less in the No-Nap condition than in the Nap condition. Moreover, the change in percent detected between runs 1 and 4 is greater in the No-Nap condition (30%) than in the Nap condition (18%). Together, these results indicate greater deterioration in performance in the No-Nap condition during the final run.

Although relatively small, the overall proportion of pedestrians that were missed entirely, as shown by the size of the top, white portions of the bar graphs, was slightly greater in the Nap condition, relative to the No-Nap condition. Among alert drivers, pedestrians are almost never missed, since they move very near to the simulated vehicle's path and become quite large before disappearing from the driver's view. Pedestrians that were missed altogether generally indicate a significant level of inattention.

Analyses of variance were computed on the proportions of pedestrian targets detected within 3.5 seconds, and two transformations of these proportions. Independent variables were condition and run. A summary of the analysis results is presented in Table 11.

The results indicated that the nap had no effect on the proportion of pedestrian targets detected. The main effect of run was also not significant, indicating no change in the proportion of pedestrian targets detected across runs. There was no interaction between condition and run.

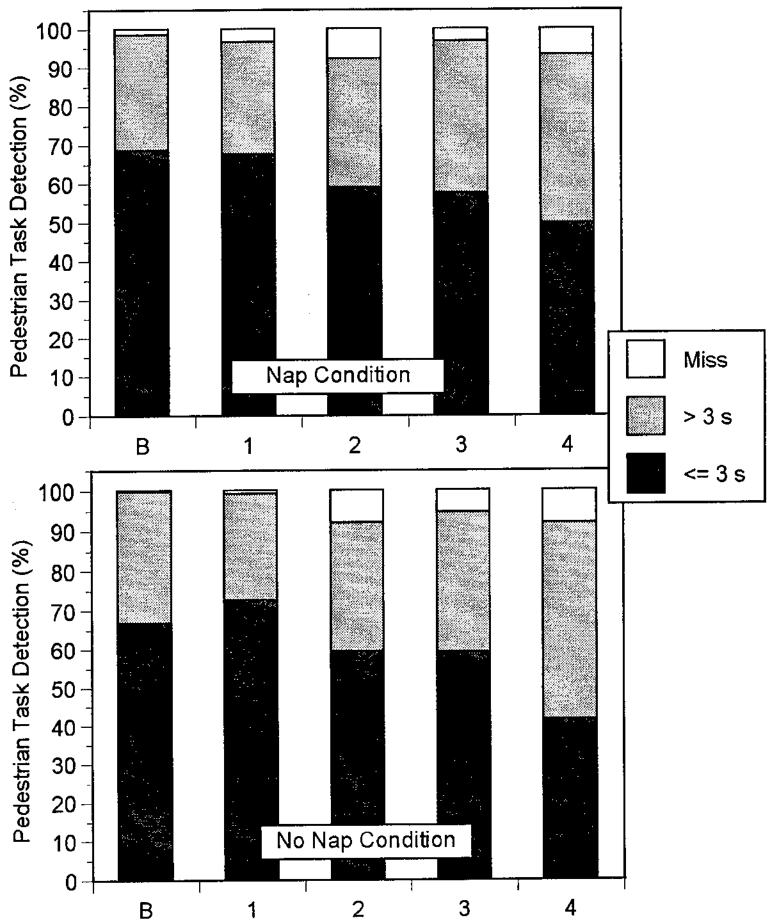


Figure 9. Mean Pedestrian Task Detection x Run x Condition

Table 11. Summary of ANOVA Results for Pedestrian Detection Proportions

	Condition		Run		Cond x Run	
Dep. Measure	**F**	**p**	**F**	**p**	**F**	**p**
p (ped)	0.14	.72	2.44	.09	1.50	.24
Log [p(ped)]	0.88	.38	1.81	.18	1.34	.29
Arcsin (p[ped])	0.00	.98	2.12	.13	1.91	.16

The mean response times for the pedestrian task by condition and run are shown in Figure 9. The mean response times remained constant in the baseline and first numbered run and there were no differences between the Nap and No-Nap conditions. The mean response time increased starting in the second numbered run and continued to increase throughout the remaining runs. Mean response times were greater in the No-Nap condition than in the Nap condition for the last three runs, with the biggest differences occurring in runs two and four. Slower response times in the No-Nap condition are consistent with greater impairment.

Analyses of variance were computed on the Pedestrian target-detection response times (RTs) and two transformations. Independent variables were condition and run. A summary of the analysis results is presented in Table 12.

Table 12. Summary of ANOVA Results for Pedestrian Task Response Time

	Condition		**Run**		**Cond x Run**	
Dep. Measure	**F**	**p**	**F**	**p**	**F**	**p**
Ped RT	2.35	.17	12.04	.00	.85	.49
Log (RT)	1.46	.27	15.13	.00	.49	.69
Square root (RT)	1.91	.21	13.80	.00	0.67	.58

The main effect of condition was not statistically significant, indicating no effect of the nap on pedestrian task response time. The main effect of run was significant for all analyses, reflecting the strong and consistent increase in pedestrian task RT over run. The Condition x Run interaction was not statistically significant.

7. Continuous Driving Performance

Symptoms of inattention are likely to occur either during unexpected increases in task demand or during extended stretches of uneventful driving. Unexpected obstacles, including vehicles parked in the roadway and encroaching oncoming vehicles were included in the experimental design to represent the first scenario. These data were not analyzed for the current presentation.

The experimental design included four 30-minute driving blocks, each divided into three approximately 10-minute segments which varied by the relative level of task demand. Driver workload was highest during the high-demand segments, slightly less in the medium-demand segments and least in the low-demand segments. Four uneventful sub-segments (intervals) were also included in each 30-minute driving block. These included two in the low-demand segment (5,000 feet and 10,000 feet) and one in the medium-demand segment. The following analyses were conducted using data from the shorter sub-segments in the low-demand segments. To ensure that driver behavior was homogeneous, and not influenced by preceding or succeeding events, we removed 500 feet from both the beginning and end so that the resulting length of the sub-segments was 4000 feet. Driving time represented by these segments varied between approximately 39 and 58 seconds, corresponding, respectively, to average travel speeds of approximately 69 and 46 miles per hour.

Means and standard deviations were computed for each subject (8) x condition (2) x run (4) x block (3) combination, for which data were available. Segments with large standard deviations were examined. Segments with standard deviations greater than 12 feet per second were removed, because the associated range of speed approached zero, suggesting a crash either during or immediately preceding the interval. Usable data were obtained from 172 (90%) of 192 possible samples

Analyses of variance were computed for the following summary measures: mean speed (MSPD), standard deviation of speed (SSPD), standard deviation of lane position (SLP), and standard deviation of steering wheel position (SSTR). Independent variables included run (4), condition (2) and block (4). Group means are plotted in Figure 10, and the main effects from these analyses are summarized in Table 13.

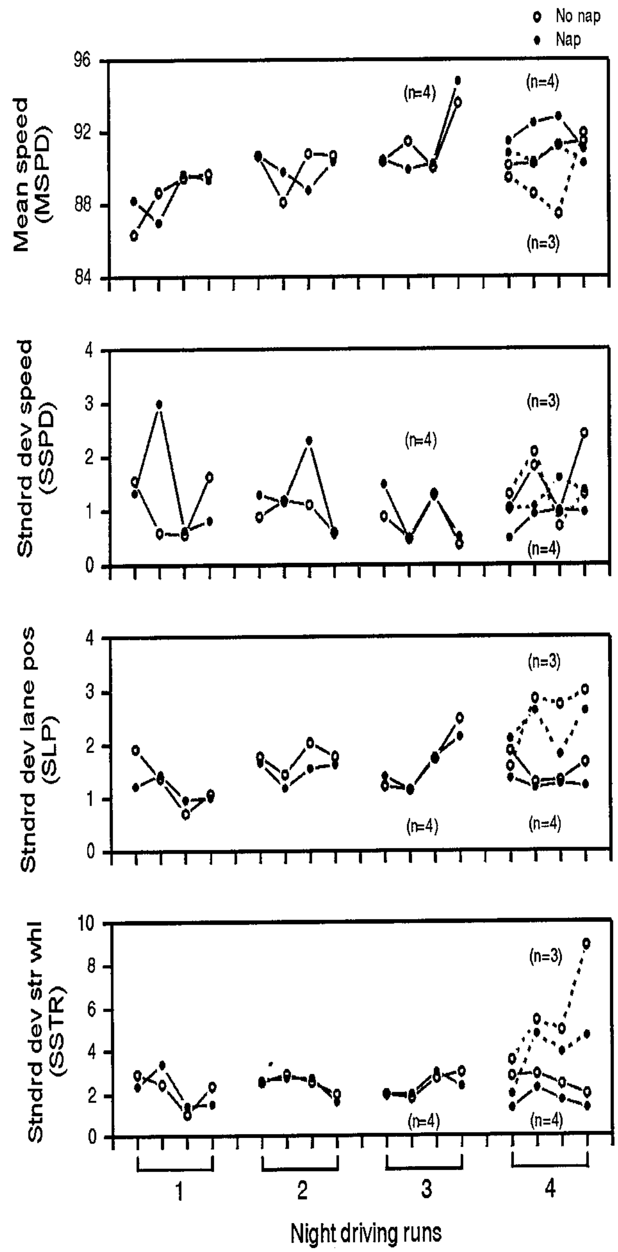


Figure 10. Group Means for Continuous Driving Measures

Table 13. Main effects for vehicle control behavior in uneventful segments

	Run		Condition		Block	
	F	p	F	p	F	p
MSP	2.3	.11	1.47	.27	2.98	.06
SSPD	1.35	.29	0.48	.52	1.52	.24
SLP	5.79	.007	5.89	.05	0.59	.63
SSTR	2.44	.10	4.08	.09	1.18	.35

Differences in vehicle-control performance attributable to the nap are reflected in the main effect of condition. The standard deviation of lane position (SLP) was the only performance measure for which this effect was statistically significant. For this measure, the overall mean for the no-nap condition was greater than the corresponding mean for the nap condition, which is consistent with they hypothesis that the nap improved vehicle control. (Nap: M = 1.54; SD = 0.9; No-nap: M = 1.66; SD = 1.03). None of the interactions with condition were statistically significant.

While the main effect of condition approached statistical significance for the SSTR, the Condition x Run interaction was statistically significant (F [3,12] = 3.71, p = .04). Consideration of the means for this effect revealed a large increase between the third and fourth runs in the No-Nap condition (SSTR No-Nap: Run 3: M = 2.41; Run 4: M = 3.34) that was not apparent in the Nap condition (SSTR Nap: Run 3: M = 2.70; Run 4: M = 2.62).

The main effect of run on the standard deviation of lane position was the strongest effect shown in the table. This effect reflects a progressive increase over runs 1-4.

Main effects of block were included to examine consistent changes across the 2-hour driving run. Of the measures examined, only (MSPD) revealed a main effect that approached statistical significance. Examination of the group means reveals a progressive increase in mean speed over 2-hour runs, across all conditions. None of the interactions with block reached statistical significance.

F. SUBJECTIVE SLEEPINESS AND FATIGUE

VAS values for individual subjects are shown in Appendix C, Figures C1 (Fatigue) and C2 (Sleepiness). The raw values were transformed into z-scores prior to statistical analysis (Appendix C, Figure C3), after eliminating Session 4 data of the three subjects who took unscheduled naps during Session 3. Data from three other sessions were missing in one condition and the corresponding data from the other condition were therefore also excluded. In Appendix C, Figures C1-C3, BL1 refers to data collected at end of the morning baseline driving session, while BL2 refers to data collected at 21:30, just prior to the start of the first night driving session.

The pattern of results was very similar for fatigue and sleepiness. Scores on the two scales increased significantly across the night in both the Nap and No-Nap conditions (repeated measures ANOVA, $p = 0.03$ or better). Except during BL1, scores on both scales were generally higher in the No-Nap than in the Nap condition, indicating greater subjective fatigue and sleepiness. Paired t-test comparisons (one-tailed) showed significantly higher sleepiness scores in the No-Nap condition for all night sessions except Session 3 ($p = 0.02$ or better, indicated by asterisks in Appendix C, Figure C3), and significantly lower sleepiness during BL1 (i.e., before the scheduled afternoon nap). Fatigue scores were significantly higher in the No-Nap condition for BL2 ($p = 0.01$) and Session 2 ($p = 0.03$), and showed a non-significant trend for Sessions 1 ($p = 0.07$) and 4 ($p = 0.08$).

G. PERFORMANCE ASSESSMENT BATTERY (PAB)

Performance measures analyzed for the Two-Letter Memory and Search Test (MAST-2) and the Wilkinson test were Reaction Time (RT) in seconds, Standard Deviation of RT, and Throughput, a measure that combines both speed and accuracy of performance (accuracy, or percent correct, was consistently high, with only occasional errors, and was therefore not analyzed separately). For the Time Wall test, the measures were Reaction Time (which in this case represents the subject's estimate of a standard 10-second interval), and the Coefficient of Variation of RT.

Appendix D, Figures D1-D3 show mean values on the three tests, plotted separately for the subjects who took unscheduled naps during Session 3 (N = 3, right panels) and for those who did not (N = 5, left panels).

In general, the MAST-2 and Wilkinson data show a deterioration in performance at night, especially in the No-Nap condition, as indicated by longer RTs, larger standard deviations, and lower throughputs. This deterioration occurred earlier in the night for the three subjects who took unscheduled naps than for the five subjects who did not: in the former case, the deterioration was already visible during the 21:30 and 24:00 test sessions, while in the latter this didn't occur until the 05:00 and 07:30 sessions.

The Time Wall RT data showed no systematic trends across the night, and no clear differences between the Nap and No-Nap conditions. However, the Coefficient of Variation increased across the night, especially in the No-Nap condition, with the largest increase occurring earlier in the night for the three subjects who took unscheduled naps than for the other five subjects.

For statistical analysis, data from Session 4 for the three subjects who took unscheduled naps were excluded from the analysis. There were also a few additional missing data points due to computer malfunction. Group means for the three tests are shown in Appendix D, Figures D4-D6.

Changes across sessions were analyzed using repeated-measures ANOVA. Significant effects were found for Standard Deviation in the MAST-2 and for the Coefficient of Variation in the Time Wall. Differences between the Nap and No-Nap conditions were tested using paired t-tests. Significant differences were found for MAST-2 RT at 21:30, while several other measures approached statistical significance ($0.05 < p < 0.1$), including MAST-2 Standard Deviation and Throughput at 07:30, Wilkinson RT and Throughput at 07:30, and Time Wall Coefficient of Variation at 05:00.

H. EEG DATA ANALYSIS

1. Rational

There is ample evidence that EEG activity in the theta (4-8 Hz) and alpha (8-12 Hz) frequency bands is sensitive to the effects of sleep loss and performance fatigue and to variations in task demands. Field studies of shift workers, train drivers and truck drivers using continuous EEG recordings report increased alpha and/or theta activity during the night relative to baseline or daytime levels, as well as increases during the night shift, associated with higher subjective sleepiness and performance decrements (Akerstedt et al., 1991; Torsvall and Akerstedt, 1985, 1987; Cabon et al., 1993; Kecklund and Akerstedt, 1993).

2. Data Analysis

EEG during driving was recorded referentially using the following derivations: O1-A2, O2-A1, P3-A2, P4-A1, C3-A2, C4-A1, F3-A2, and F4-A1. During subsequent, off-line analysis, pairs of referential channels were combined to yield the following bipolar derivations: O1-P3, O2-P4, C3-P3, C4-P4, F3-C3, and F4-C4.

Records for individual 10-minute driving segments (high-, medium-, and low-demand segments) were first identified and edited for artifacts, as described below. Data from all channels (referential and bipolar) were then subjected to spectral analysis (Fast Fourier Transform on 4-second epochs), yielding total amplitude and power in the following frequency bands: delta (0.75-3.75 Hz), theta (4-7.75 Hz), alpha (8-11.75 Hz), sigma (12-14.75 Hz), and beta (15-24.75 Hz). An additional frequency band comprising fast theta and slow alpha activity was also analyzed (6-9.75 Hz).

EEG data showing large artifacts in the O1-A2 or P3-A2 referential channels, usually associated with high EMG activity and large eye movements (especially large horizontal saccades), were eliminated from the records. Smaller artifacts in the referential channels, including artifacts associated with eye blinks, usually disappeared when the data were converted to a bipolar derivation and were therefore not edited out. This meant that the lowest and highest frequency bands could, in some cases, have been contaminated by low frequency noise (e.g., from blinks and eye movements) and high frequency noise (e.g., from muscle activity), respectively. Therefore, only data from the middle frequency bands, alpha, theta, and 6-10 Hz, are reported, and only for the bipolar O1-P3 derivation. In addition, 10-minute driving segments with fewer than 30 epochs (2 minutes) of edited

EEG data were excluded from further analysis, because the data were not considered representative of the entire segment. Data from Subject 6 in the No-Nap condition did not generally meet this criterion, and were therefore completely excluded from the analysis.

3. Driving EEG

Spectral amplitudes in the alpha, theta, and 6-10 Hz frequency bands for individual subjects are shown in Appendix E. Since all electrodes were reapplied before the start of the baseline driving run and before the start of the first night run in each condition, a normalization procedure was used to ensure that differences within or between conditions were not due to slight differences in electrode locations.

For each condition, absolute spectral amplitude during the four night driving runs was expressed as a percentage of the amplitude for that frequency band obtained during the resting EEG test (eyes open) that was performed at 21:30, just prior to the start of Run 1. Since the baseline driving runs were not preceded by resting EEG testing, the data from these sessions were expressed as a percentage of the amplitude obtained during the first 3 minutes of driving.

Figure 11 illustrates average spectral amplitudes in the alpha, theta and 6-10 Hz frequency bands for the twelve driving segments comprising each driving session in the two conditions. Data for Session 4 is plotted separately for the three subjects who took an unscheduled nap.

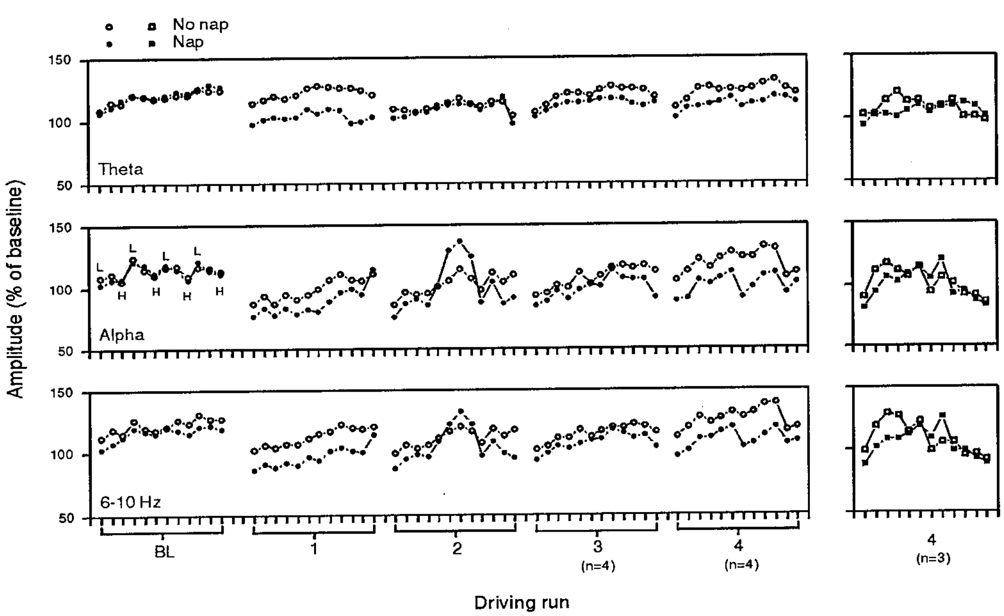


Figure 11. Mean driving EEG data

In general, all three frequency bands showed increases in spectral amplitude across sessions (TOD effect) as well as within sessions (TOT effect) . Amplitudes were also generally higher in the No-Nap than in the Nap condition, but only during the night driving sessions.

A gradual increase across the night sessions in the spectral amplitude of the alpha, theta and 6-10 Hz frequency bands was observed in Subjects 1-5 (Appendix E, Figures E1-E5). Sudden changes in amplitude levels from one session to the next characterized the results for Subject 7 (Session 1 to 2 in both Nap and No-Nap; Appendix E, Figure E6) and Subject 9 (Session 2 to 3 and 3 to 4, both conditions; Appendix E, Figure E7). The sudden decrease in amplitude observed in these cases is probably related to the reapplication of the electrodes from the O1-P3 derivation. Since all electrodes were applied both before the morning baseline recording and before the evening and night sessions, slight differences in electrode montage may also explain the generally higher amplitude levels observed during baseline relative to Session 1.

An increase in amplitude in all three frequency bands was also generally observed in most subjects within each driving session, across successive 10-minute driving segments.

Higher amplitudes in the No-Nap compared to the Nap condition during the night sessions were observed in Subject 1, especially in the alpha band, in Subject 3 in the second half of Sessions 3 and 4, and in Subject 9, especially in Session 4. Subject 2 showed slightly higher amplitude levels in the Nap than in the No-Nap condition. In this case the lack of an effect of the afternoon nap might be partially explained by the fact that in the Nap condition this subject obtained only 10 more minutes of total sleep time (TST) than in the No-Nap condition. Subject 5 showed a steep rise in alpha amplitude in the middle of Session 2 only in the Nap condition. The results for the three subjects who took an unscheduled nap are difficult to interpret on account of missing data for Session 3, while the results for Session 4 were affected by the unscheduled nap and appeared to differ between frequency bands in different subjects.

When individual data were averaged across the four night sessions and compared between the two conditions, spectral amplitudes in the theta and 6-10 Hz frequency bands were significantly higher in the No-Nap condition (1-tailed paired t-tests: $p=0.02$).

4. Effects of Driving Demand Level

The driving scenarios comprised separate high-, medium-, and low-demand segments, each segment lasting approximately 10 minutes. The expectation was that a high-demand segment would be more stimulating and, therefore, more alerting than a low-demand segment, and that this would be reflected in some of the physiological measures of alertness that were recorded in this study.

The effects of demand level on the spectral amplitudes in the alpha frequency band are illustrated in Figure 11 for the Baseline Session (high- and low-demand segments indicated by H and L, respectively).

Alpha amplitude was significantly higher during low-demand driving segments than during medium- and high-demand driving segments at Baseline and in Sessions 1 and 2 of the No-Nap condition, as well as in Sessions 2 and 4 of the Nap condition (ANOVA, $p<0.05$).

5. EEG Changes Preceding Crashes

EOG records during driving showed many episodes of eye closure which lasted for several seconds and often preceded off-road accidents and collisions. As expected, examination of the EEG records showed an increase in alpha activity during such episodes.

This is illustrated in Appendix F, Figure F1a, which shows, from top to bottom, the simulator event channel, followed by vertical EOG, horizontal EOG, two EEG, EMG, EKG, and two additional EEG channels. The 8-second record shows that the subject had his eyes closed for about 6 seconds (upward deflection of the vertical EOG trace) before driving off the road (off-road accident indicated on the top trace). An increase in alpha activity during the eyes closed episode is clearly visible in the EEG channels, especially the occipital derivation (O1-A2). Appendix F, Figure F1b shows amplitude spectra for successive 4-second epochs encompassing that accident (Off-road Accident 1) as well as a second accident (Off-road Accident 2) which occurred approximately 30 seconds later (note that time starts at the bottom of the chart). A peak around 11-12 Hz, in the alpha frequency range, is clearly visible in most of the spectra, but the size of that peak increases just before the first off-road accident, when the subject closed his eyes.

Appendix F, Figure F2 shows similar records from another subject who also had two off-road accidents in rapid succession (Off-road Accidents 2 and 3). The EOG records show that the subjects' eyes were open at the time of off-road Accident 2 (Appendix F, Figure F2a) and closed at the time of off-road Accident 3 (Appendix F, Figure F2b), but spectral analysis of the EEG shows that both accidents were preceded by a large increase in alpha activity. These results indicate that EEG alpha activity is a reliable indicator of fatigue and alertness levels, even in subjects whose eyes remain open, and that increases in alpha often precede a deterioration in driving performance.

6. Resting EEG

EEG recordings obtained at rest with eyes open and eyes closed provide an additional index of fatigue. As described above, these recordings were obtained immediately after the end of each driving session and about 20 minutes before the start of Session 1.

In well-rested subjects, the EEG shows little activity in the alpha frequency band when the eyes are open and high activity when the eyes are closed. An increase in fatigue level is usually accompanied by an increase in alpha with eyes open and a decrease with eyes closed.

These effects are illustrated in Appendix G, Figures G1 and G2, which show EEG tracings and corresponding amplitude spectra during the transition from eyes open to eyes closed for Subject 5. The data in Appendix G, Figure G1a,b were recorded at 21:30, when the subject was still well-rested; they show low alpha activity during eyes open, and sustained high activity during eyes closed (sharp peaks around 12 Hz in Appendix G, Figure G1b). In contrast, at 02:30, when the subject was very tired (this subject asked to take an unscheduled nap shortly thereafter), alpha activity during eyes open was higher than it was at 21:30, and showed only a transient increase when the subject closed his eyes (Appendix G, Figures G2a,b).

The ratio of alpha amplitude with eyes open over alpha amplitude with eyes closed was calculated for each subject. This method represents a form of standardization, since the eyes open/eyes closed ratio is relatively immune to small differences in electrode placement and to inter-individual differences in alpha amplitude. As explained above, that ratio is expected to increase with fatigue.

Average data for the five resting EEG recording sessions are shown in Figure 12. Group means for the five test sessions are illustrated separately for the five subjects who adhered to the experimental protocol (upper panel) and for the three subjects who needed to take an unscheduled nap (lower panel). In both groups the alpha ratio was lowest at 21:30, indicating maximum alertness, and was higher during the night than during the day. The alpha ratio showed an increase at 0:00 and then remained fairly stable throughout the night in the group of subjects who completed all driving sessions. In contrast, the alpha ratio for the subjects who required to take an unscheduled nap around 5:00 showed a much sharper increase at 0:00 and continued to rise in the next testing session.

Statistical analyses were performed on data from all subjects for the first four resting EEG tests, while data from the three subjects who took an unscheduled nap were eliminated from Sessions 3 and 4. A significant TOD effect was found in both the Nap and No-Nap conditions (One-way ANOVA: F=6.703, p=0.004 and F=7.094, p=0.005, respectively). In the Nap condition the eyes open/eyes closed alpha ratio was significantly lower at 21:30 compared to 12:00, 0:00 and 2:30. In the No-Nap condition the eyes open/eyes closed alpha ratio was significantly lower at 12:00 and 21:30 compared to 0:00 and 2:30 (t-tests, p<0.05).

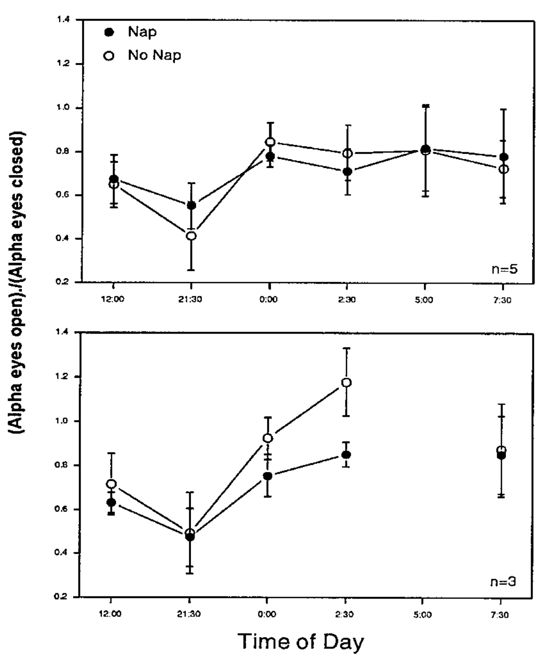


Figure 12. Resting EEG: ratio of alpha amplitude with eye open over alpha amplitude with eyes closed.

7. Discussion

As expected, alertness levels, as reflected in EEG activity, were significantly lower at night than during the day. The ratio of eyes open over eyes closed alpha, which is inversely proportional to the level of alertness, was lower at 21:30, before the start of the night driving runs, than at 0:00 and 2:30, regardless of experimental condition. This peak in alertness in the early evening corresponds to the nadir in sleepiness between 20:00 and 22:00 which has been termed a "forbidden zone" for sleep (Levie, 1986). Alertness levels were also higher at 21:30 than at 12:00, although the difference was statistically significant in the Nap condition only. Torsvall and Akerstedt (1983, 1987) obtained continuous EEG recordings in train drivers and reported a gradual, steady increase in alpha power during night driving, as well as increases of 50% and 30% in alpha and theta power, respectively, at night relative to daytime levels. The magnitude of the increase is comparable to the difference in alpha ratio observed in our study between the 12:00 session and the two night sessions.

The alpha frequency band was also sensitive to changes in task demand level during driving, with significant increases in amplitude, indicating lower alertness, during the low demand 10-minute driving segments. In a field study of car driving, increases in alpha activity have been shown to be inversely correlated to the degree of stimulation offered by the driving conditions, with city driving showing less alpha than rural driving (Lecret and Pottier, 1971). Using an EEG-derived activation index (theta + alpha/beta), deWaard and Brookhuis (1991) reported a gradual decrease in activation during a 165-minute highway driving test, followed by an increase when the drivers were asked to perform a 15- minute "car following" test, indicating that under certain conditions, even a mild increase in task demand level can counteract the effects of driving fatigue.

The effects of the afternoon nap on alertness levels during the night driving runs were significant in the theta and 6-10 Hz frequency bands, with lower amplitudes in the Nap relative to the No-Nap condition. The presence of theta activity is one of the signs of sleep onset, and increases in theta activity are associated with periods of mild to severe sleepiness during vigilance tasks (Torsvall and Akerstedt, 1988), night driving (Mackie and Miller, 1978), and sleep deprivation (Lorenzo et.al., 1995). Thus, it can be concluded that the 2-hour afternoon nap was effective in reducing overall sleepiness levels during the following night.

I. EOG DATA ANALYSIS

1. Rationale

The inclusion of EOG recordings in this study served two purposes: (1) to evaluate the utility of various EOG-derived ocular parameters in predicting loss of alertness and consequent impairment in driving performance, and (2) as a means of assessing the validity and reliability of ocular parameters derived from concurrently recorded eye tracker data.

Ocular parameters useful in predicting driver fatigue would be expected to show chronic changes as a function of TOT and of TOD, as well as phasic changes immediately preceding traffic accidents and other events indicative of impaired performance. Two types of analysis were therefore performed, one aimed at assessing TOD and TOT effects, and the other looking for possible changes in EOG-derived measures associated with traffic accidents.

In the present context, TOT effects refer to changes within each 2-hour driving run, while TOD effects refer to changes across successive runs. Note that any TOD effects obtained would be attributable to sleep deprivation (i.e., an increase in time spent awake), to diurnal variations in the parameters that were examined, or to a combination of these two factors.

2. Data Analysis: TOT and TOD Effects

Vertical and horizontal EOG were continuously recorded at a rate of 128 Hz, along with stimulus and response event stamps from the driving simulator and with one channel of analog data from the eye tracker, representing horizontal point-of-gaze. The data were first converted to a sampling rate of 100 Hz using MATLAB, then converted to PCWUPDRS format for analysis.

Six minutes of data were sampled out of every 10-minute driving segment. Data from every three consecutive segments were then combined, providing 18 minutes of sampled data for every 30-minute driving block. There were some problems in identifying driving segments since the output from the simulator occasionally missed or mislabeled the event signaling the transition from one segment to the next.

Available algorithms were used for blink and saccade identification in PCWUPDRS. An operator reviewed computer-identified events and edited the data as necessary. The following information was abstracted from the vertical EOG channel and manually edited:

<u>Blink amplitude</u> (in arbitrary A/D units).

<u>Blink closing duration</u> - the time from onset of lid closure to maximum closure.

<u>50% window</u> - this measure identifies the point in time where the lid is half closed, then finds the point in time during the reopening phase where the lid passes through the same voltage level. The time difference between these two points identifies the 50% duration window.

<u>Blink frequency</u> (blinks per minute).

The data were not corrected for periods of lid closures, nor were long closure duration blinks (i.e., blinks with a 50% window greater than 500 milliseconds) included in this analysis.

3. Blink Measures: Effects of TOD and TOT

Individual data from seven subjects, plotted in 30-minute blocks, are shown in Appendix H, Figures H1-H7. Data for some blocks, runs, or entire conditions were either missing or could not be analyzed due to missing event stamps or to poor quality of the EOG signals.

Mean blink frequencies for the seven subjects are shown in Figure 13. The group means were obtained after first calculating averages for individual subjects across the two conditions. Note that the means for Run 3 are based on a smaller number of subjects, since three subjects took unscheduled naps at that time.

Blink Closing Duration and 50% Window

These two measures were similar in several respects. Both showed clear TOD effects, with increases across successive night driving runs. This was observed in both the Nap and No-Nap conditions. The two measures also showed TOT effects in most subjects. These effects consisted of a sustained increase across successive 30-minute blocks, or an initial increase followed by a decrease. The latter pattern occurred more frequently in the later runs. TOT effects were less consistent during the baseline run.

Effects of the scheduled nap could not be determined with any certainty because only four subjects provided data from both conditions. Three of these subjects (Subjects 1, 4 and 7) showed longer blink closing duration and 50% windows in the No-Nap condition, especially during the later runs. The fourth subject (Subject 5) showed little or no difference between the two conditions in Runs 1 and 2. In Run 4, which was completed after that subject took an unscheduled nap, the two blink measures were longer in the Nap condition.

Blink Frequency

Figure 13 shows that, on average, blink frequency also showed a TOD effect, with increases from Run 1 to Run 2, and from Run 2 to Run 4 (the data for Run 3 did not fit that pattern, but they were derived from a smaller number of subjects and are therefore not directly comparable). As seen in Appendix H, Figures H1-H7, however, interindividual variability was considerably higher than it was for the blink duration measures. Indeed, increases across successive runs were only shown by Subject 1 (both conditions), Subject 4 (Nap condition only) and Subject 9 (Nap condition).

As with the blink duration measures, changes in blink frequency with TOT consisted of an initial increase followed by a decrease. However, the decrease started earlier (after the second or third 30-minute block), and interindividual variability was higher. Subject 3 was exceptional in that his blink frequency was highest in the first 30-minute block and decreased thereafter.

Blink Amplitude

No consistent trends in blink amplitude as a function of either TOT or TOD were apparent in the data.

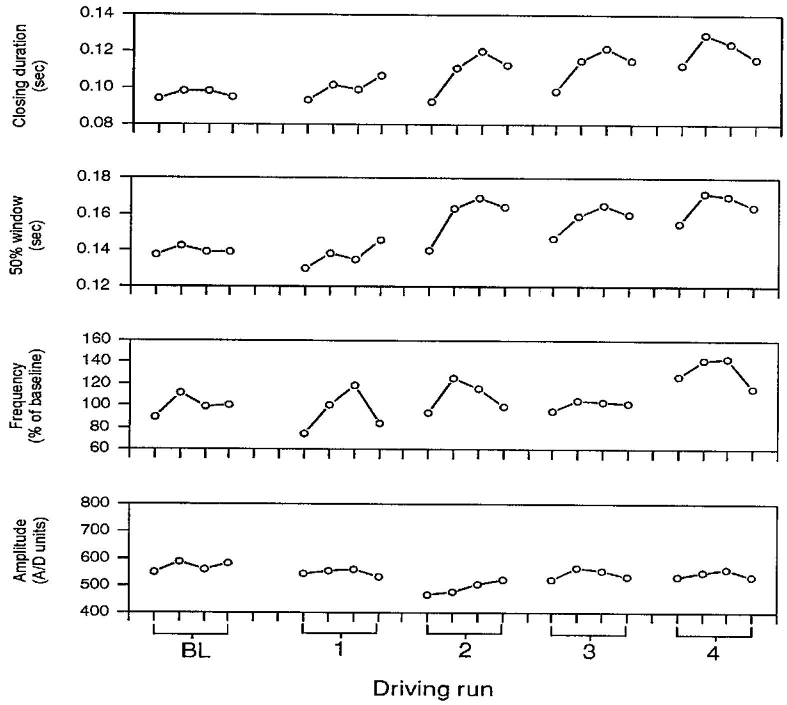


Figure 13. EOG Blink Measures

4. Data Analysis: Off-Road Accidents

To further evaluate the utility of oculometric parameters in predicting driver fatigue, changes in EOG measures associated with off-road accidents were analyzed. Off-road accidents were selected because there was a reasonable number of such accidents while other types of accidents (such as collision with another vehicle or hitting a pedestrian) occurred less frequently.

Only unprovoked off-road accidents were selected, that is, accidents not immediately preceded (within 20 seconds) by events such as wind gusts or an oncoming car encroaching on the driver's lane. In addition, the time between successive accidents had to be at least 90 seconds. For subjects who took unscheduled naps instead of completing Run 3 (Subjects 4, 5 and 7), accidents were selected from Runs 1 and 2, while for the remaining subjects, accidents were selected from Runs 3 and 4. In some cases, the accidents were taken from both the Nap and No-Nap conditions, while in others, they were all taken from the same condition.

Results for two drivers could not be evaluated. One of these (Subject 8) participated in one condition only and felt sick during the night driving runs, while the other (Subject 6) had only one off-road accident in both conditions. Adequate data were available for the remaining seven drivers. For each subject, the earliest eight accidents meeting the above criteria were selected. Thus, the analyses were performed on a total of 56 unprovoked off-road accidents.

Nine consecutive 10-second epochs were evaluated, starting 60 seconds before and ending 30 seconds after each accident. The following information was abstracted for each 10-second epoch, and evaluated statistically using the Friedman two-way analysis of variance by ranks.

Blink frequency (blinks per minute).

Blink closing duration - the median closing duration was calculated for each driver. The algorithm for determining closing duration is described below.

Eyelid closures - any lid closure with a duration longer than 0.5 second, measured from blink initiation to the point where the eyelid started to reopen.

Saccade frequency - Eye movements, identified by visual inspection of the EOG records, were considered saccades if they met minimal slope and amplitude criteria. The measure analyzed was the number of such saccades in each 10-second epoch. Because of muscle activity, EEG "noise" and other artifacts, this was the least reliable of the measures obtained.

Blink closing duration was determined by visual inspection of the data and application of the following algorithm. Blinks were first identified using gross morphological characteristics. The lowest point in the blink (lowest A/D value, representing full closure) was determined and sampled backward (in time) until a criterion of "minimal change" was met, which consisted of an increase of at least three A/D units between successive samples. In cases where the data were too noisy (due to EEG, muscle, or other artifacts), the "minimal change" criterion was applied to alternate data points. Closing duration was defined as the time difference between the lowest point in the blink and the point meeting the "minimal change" criterion.

5. EOG Changes Associated With Off-Road Accidents

Blink Frequency

Figure 14 (top panel) depicts mean blink frequency (median blinks/minute per driver) for each of the nine 10-second epochs. The arrow identifies the time of occurrence of the accident.

The chi squared value for this data set was 11.26, which, with 8 d.f., is not statistically significant. Of the seven drivers whose data were analyzed, five showed a reduction in blink rate immediately preceding the accident while two showed an increase.

Blink Closing Duration

Mean blink closing duration is shown in Figure 14 (second panel). An ANOVA produced a chi squared value of 25.13, which is significant at the $p<.01$ level. Closing duration started to decrease in the 10-second epoch immediately preceding the accident, and was shortest in the three epochs following the accident. It should be noted that the effect, though reliable, reflects relatively small differences in closing duration. The

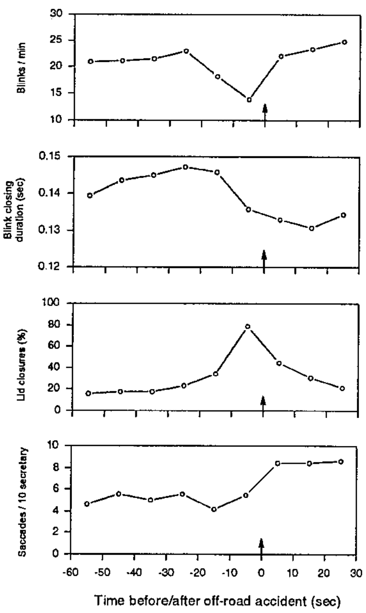


Figure 14. EOG analysis of off-road accidents

difference between the largest and the smallest mean closing duration was 16 milliseconds.

To determine whether this closing duration was, in general, longer than "normal", we evaluated data from twenty five consecutive blinks sampled at two time points: during the calibration period prior to initiation of a driving run, and early in a run, within the first 10 minutes of driving. These data were sampled from either the baseline run or the first night run.

Average closing duration for both sets of data was 120 milliseconds, which is about 10 milliseconds less than the lowest point in Figure 14. For four of the drivers, the average duration for each of these two periods was smaller than for any of the nine epochs associated with the off-road accidents. For one subject they were smaller than eight and for another subject they were smaller than seven of the nine epochs evaluated. One subject was exceptional in that his "normal" values were equal to the largest values associated with the off-road accident. Thus, the difference between the "normal" closing duration and those obtained during epochs surrounding the accidents is, for most drivers, quite robust.

Eyelid Closures

An ANOVA was performed on the number of times a driver showed one or more lid closures in each of the nine 10-second epochs (since eight accidents were selected per subject, the maximum number of lid closures per epoch was eight). The ANOVA achieved a chi squared value of 23.23, which is significant at the $p<0.01$ level.

Figure 14 (third panel) shows the total number of lid closure occurrences per epoch shown by all subjects, converted into a percentage (with seven subjects and eight accidents per subject, the maximum total number of lid closures was fifty six). It is readily apparent that the epoch immediately preceding the accident had the largest number of lid closures. The frequency of lid closures started to increase 20-30 seconds before the accident, and showed a gradual decrease during the 30-second period following the accident.

It should be noted that even in the earliest epoch sampled (starting 60 seconds before the accident) there were nine occurrences of lid closures, representing 16% of the maximum number possible. Although lid closures were not systematically sampled at other time points, inspection of the EOG records and videotapes suggests that such closures were rare at the start of driving in the early runs (Baseline or Run 1).

The duration of these eyelid closures ranged from 0.5 second to a maximum of 22 seconds, with median closure duration per driver ranging from 0.7 to 2.4 seconds (because of occasional long duration closures the median is a more representative measure than the mean.) Of the forty four eyelid closures identified in the epoch immediately preceding an accident, three overlapped the preceding epoch and eleven overlapped the epoch immediately following the accident.

Saccade frequency
These results are presented in Figure 14 (bottom panel). The ANOVA chi squared value was 23.228, which is significant at the $p<0.01$ level. This effect consisted mainly of a marked increase in saccade frequency during the 30 seconds following the accident.

6. Discussion

TOD and TOT Effects
The TOD and TOT results described above are in agreement with data reported in the literature in some respects, but differ in others. In the case of blink duration and frequency, one notable difference is that, during the night runs, these parameters often showed an initial increase followed by a decrease, rather than a steady increase as reported in most studies (reviewed by Stern et al. 1994a, see also Goldstein et al. 1985, Bauer et al. 1985, 1987, Stern et al. 1994 a, b). In addition, interindividual variability in blink frequency was higher and the effect of TOT less consistent than was expected.

Most previous studies involved single test sessions lasting between 0.5 and 4 hours and scheduled during the daytime. In contrast, subjects in the present study were required to complete four 2-hour driving runs at night, starting at 22:00, with only 30-minute breaks between runs. Thus, our subjects were undoubtedly more tired than those of most other studies that have examined TOT effects on blink parameters.

Higher fatigue levels would be expected to lead to an increase in long duration blinks as well as in slow eye closures. For example, Morris (1984) studied subjects operating a flight simulator for 4 hours following a night of sleep deprivation, and found that the frequency of long closure duration blinks was one of the parameters that correlated best with flight performance. Pfaff et al. (1976) reported that, over the course of 3 hours of real driving, mean blink duration increased from 180-200 milliseconds to about 500-600 milliseconds, with individual blink duration ranging from 150 milliseconds to 2.4 seconds. Although Pfaff et al. did not distinguish between short and long duration blinks, the latter were clearly included in the analysis and must have contributed to the TOT effect described in that study.

Based on these observations, it is likely that had long duration blinks been included in the present analysis, they would have contributed to the TOD and TOT effects, and may have at least partially offset the decline in blink frequency and duration observed in the later portions of the night driving runs.

The results of the off-road accident analysis showed a dramatic increase in eyelid closures immediately preceding the accidents, accompanied by a reduction in blink closing duration and in blink frequency (the latter effect, though not statistically significant, was nevertheless observed in five of the seven drivers). If lid closures increase with fatigue, as we expect, then this reciprocal relation between lid closures and blink frequency and closing duration would also help account for the decrease in the two blink measures observed in the later portions of the night driving runs.

It should be noted that in the off-road accident analysis, lid closures were defined as any closures lasting one half second or longer and, therefore, included what in the TOD and TOT analysis were considered long duration blinks (i.e., blinks with 50% window of one half-one second). Whether the decrease in blink frequency and closing duration preceding off-road accidents would have been offset by the inclusion of one half-one second closures in the blink category remains to be determined.

Blink Amplitude

Changes in blink amplitude as a function of TOD and TOT have received much less attention than those in blink frequency and duration. In one of the studies described above (Morris 1984), a significant decrease in blink amplitude was observed over the 4 hours of task performance. More recently, however, subjects performing a simulated air traffic control task for 2 hours showed a significant increase in blink amplitude over time (Stern et al. 1994 b,c).

Other factors being equal, blinks initiated when the eyelids are partially closed will have smaller amplitudes than blinks initiated when the eyes are fully open. This would account for the decrease in blink amplitude observed by Morris, since his subjects were sleep-deprived and, therefore, more likely to show partial eye closures than the subjects studied by Stern et al. (evidence that partial eye closures increase with fatigue is described in Section K below). In addition, task duration was twice as long in the Morris study than in the Stern et al. study.

There are, however, other considerations that complicate the interpretation of EOG-based measures of blink amplitude. For example, lid position is not independent of eye and head position. Thus, if the head is held straight, the eyelids will be farther apart and blinks will have larger amplitudes than when the person is looking downward. Similarly, when the eyes are fixating an object directly in front, the eyelids will be more open if the head is drooping forward than if it is leaning backward. In addition, rotation of the eyeball in an upward direction produces voltage changes in the same direction as the lid moving downward. This also adds to the difficulty of interpreting EOG-derived blink amplitude measures.

These considerations, and because blink amplitude showed no consistent changes with either TOD or TOT in the present study, suggest that this measure may not be a useful predictor of driver fatigue. It should also be noted that blink amplitude is not a parameter that can be readily measured with the eye tracker used in this study, since the image of the pupil is lost before full closure of the eyelids is reached.

Eyelid Closures
Eyelid closures have long been considered one of the primary signs of drowsiness. Indeed, Skipper and Wierwille (1986) define drowsiness as "that moment when persistent eyelid closure occurred, persistent meaning that slow ramp closures were present." (p. 529). These authors also found a close relationship between eyelid closures and driving performance, especially variability in lane position and in low steering velocity movements.

The results depicted in Figure 14 show a similar relationship between eyelid closures and unprovoked off-road accidents, i.e., accidents that are likely to be attributable to fatigue effects alone. Thus, over 78% of the accidents included in this analysis were immediately preceded by one or more lid closures. However, while the increase in lid closures in the 10-second period preceding the accidents is certainly dramatic, it is probably not very useful as a warning signal since it doesn't allow enough time for implementation of any countermeasures.

The results also suggest a relatively high frequency of lid closures throughout the entire 60-second period preceding the off-road accidents, and the 30-second period afterwards. Thus, it will be important to determine when, in relation to the accidents, lid closures first start to occur, and whether this initial rise in lid closure frequency is useful in predicting impaired driving performance.

Saccade Frequency
A high frequency of saccadic eye movements reflects active scanning of the visual environment. The finding that saccade frequency was significantly lower in the 60 seconds preceding the off-road accidents than in the following 30 seconds is consistent with the view that subjects are less likely to engage in visual scanning behavior during periods of drowsiness. Although very limited, these observations may warrant a more complete assessment of the relation between fatigue and saccade frequency.

J. EYE TRACKER DATA

1. Rationale

The analyses performed on eye tracker data were aimed at (a) identifying a number of blink and eyelid closure parameters, (b) following their progression across TOD and TOT, and (c) examining possible associations with off-road accidents. Limited comparisons were made with corresponding EOG data.

Analysis of blinks using eye tracker data was based primarily on vertical and horizontal pupil diameter measures. We assumed that changes in pupil diameter in the vertical plane would reflect the closing and reopening portions of the blink, while full lid closure would result in the loss of point-of-gaze information. During such data drop-outs, the vertical and horizontal pupil diameter values became zero, while the vertical and horizontal point-of-gaze measures assumed a preset default value.

Eye tracker data drop-outs were also expected whenever the driver turned his/her head to look at the rear view mirrors. Head (and body) movements proved to be much more frequent than was anticipated, and they appeared to increase with the level of sleepiness that the subject was experiencing. The drivers tended to move and change position more during the later portions of a driving run and during the later runs, apparently in an attempt to remain awake and counter the effects of fatigue. As a result, data drop-outs also increased as a function of TOT and TOD.

The high rate of data drop-outs, especially in the later runs and the later portions of runs, precluded any reliable estimates of blink frequency and eye closures based on eye tracker data. Thus, the blink analysis described below focused primarily on blink duration measures, while the eye closure analysis involved partial eye closures only.

2. Data Analysis: Blinks

A window of acceptance was defined for identifying blinks. That window was set so that drop-outs ranging between 50 and 400 milliseconds, with 100 milliseconds of acceptable data preceding and following the drop-out, were flagged as possible blinks. If pupil diameter differences between the onset and termination of a blink (defined below) were less than 1 mm, the event was accepted as a blink.

Blinks were also broken down into three components, corresponding to the closing portion, the portion during which point-of-gaze information was lost, and the reopening portion. The closing portion of each blink (which the EOG data base suggests as being in the neighborhood of 100 milliseconds) was measured as follows. The first vertical pupil diameter value preceding lid closure (data loss) was identified and the data sampled backward to a point where three successive samples were either stable, or the earliest one (in time) was lower than the next value. The middle sample was identified as the time of initiation of lid closure. A similar strategy was used to define the reopening portion of the blink.

These procedures thus provided the following four blink duration measures, which were assessed in both the horizontal and vertical planes of the pupil.

Blink duration - time from onset of pupil diameter decrease to recovery of stable pupil diameter.

Pre-blink closing time (PRBCT) - time from onset of pupil diameter decrease to time of data loss.

Data loss (DATLOS) - the period during which eye position information was lost.

Post-blink opening time (POBOT) - time from end of data loss to recovery of stable pupil diameter.

Analysis of eye tracker data based on 10-minute driving segments (with different task demand levels) could not be performed as planned, because stimulus and response event codes generated by the simulator were not always transmitted correctly to the eye tracker (the problem was corrected later in the study and the interface functioned properly for the final two subjects). To examine TOT effects, each driving run was therefore divided into four equal blocks, each lasting approximately 30 minutes.

The maximum number of blinks identified in any one 30-minute block was in the neighborhood of 600. In two cases, the number of blinks identified in a block was very low, suggesting that the blink identification procedure did not function reliably in these subjects.

3. Blink Measures: Effects of TOD and TOT

Blink duration, data loss, and blink closing and opening times from eight subjects are presented in Appendix I, Figures I1-I8. Group means, calculated after first averaging the data from the Nap and No-Nap conditions, are plotted in Figure 15.

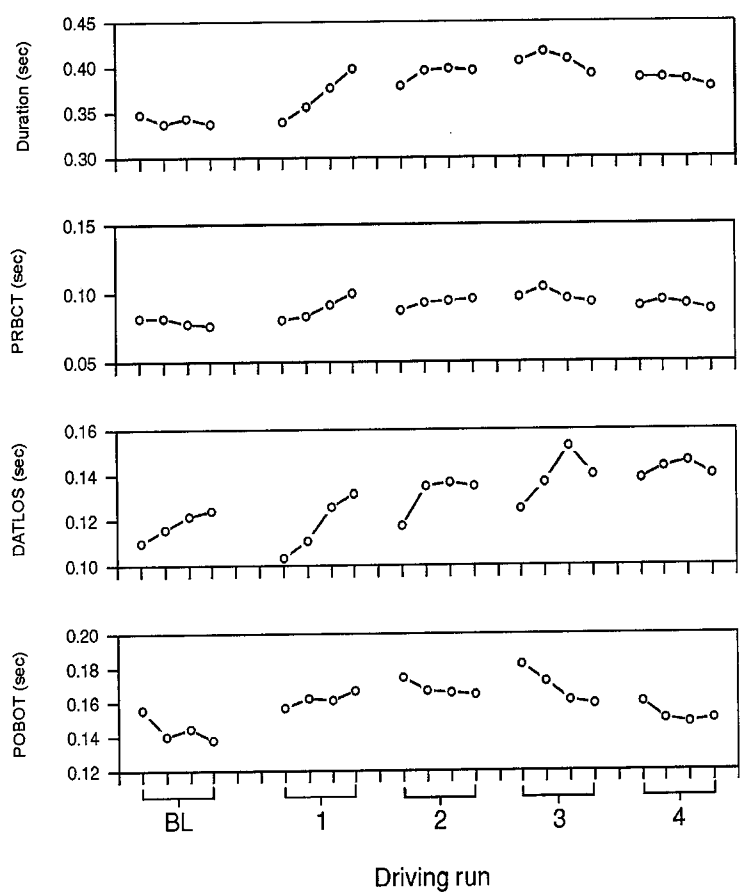


Figure 15. Eye tracker blink measures. The data are group means, calculated after first averaging the data from Nap and No-Nap conditions.

Six of the subjects showed appreciable increases in all four blink measures as a function of TOD, at least up to the second or third night driving run. This was true for both the Nap and No-Nap conditions. Data from one other subject (Subject 9) were too limited for assessing TOD effects. The remaining subject (Subject 6) showed no consistent changes across driving runs. Examination of Figure 15 shows that, on average, data loss increased steadily across all five driving runs while blink duration increased across the first four runs then decreased somewhat in the last.

Most subjects showed increases in blink duration and data loss across successive 30-minute blocks in Run 1, and about half the subjects showed such effects in Runs BL and 2. In later runs, these measures generally showed either a steady decrease, or an initial increase followed by a decrease. Based on the mean values for all eight subjects, the data loss measure appears to be more sensitive to TOT effects than blink duration. Pre-blink closing and post-blink opening times showed no consistent TOT effects.

Other Measures

A limited amount of work was done on horizontal pupil diameter immediately preceding blink occurrence (data not shown). The analysis was based on the assumption that fatigue would be reflected in pupillary constriction, and we therefore concentrated primarily on the 5th percentile measure, which represents pupil diameter at the constricted end of the continuum.

Examination of data from five subjects showed no consistent TOD effects. Evaluation of TOT effects showed that, when there was a consistent pattern, it consisted primarily of a reduction in pupil diameter across successive thirty-minute blocks. Two subjects showed this pattern in both the Nap and No-Nap conditions, while one subject showed it in the No-Nap condition and in two of the three nighttime runs of the Nap condition (this subject took an unscheduled nap in place of Run 3). The remaining two subjects showed variable patterns.

4. Data Analysis: Fixations

The analysis described in this section was performed with the EYENAL FIX software, which was developed for use with ASL eye tracking systems. The FIX program reduces raw eye position data to a series of fixations based on a set of adjustable criteria, and calculates several variables relating to these fixations.

Essentially, a fixation is identified whenever point-of-gaze remains within a 1 degree by 1 degree area for at least 100 milliseconds. Pupil losses of 200 milliseconds or less, such as would occur during blinks, are ignored, while those longer than 200 milliseconds cause the fixation to end at the last acceptable data sample.

Fixation measures provided by EYENAL include fixation duration, inter-fixation interval, inter-fixation distance, point-of-gaze, as well as mean vertical and horizontal pupil diameters. We used the latter two variables to derive a measure of eye closure consisting of the ratio of vertical pupil diameter over horizontal pupil diameter, or V/H.

As the eyelid closes, the pupil gets occluded, and both vertical pupil diameter and V/H decrease. However, using V/H rather than vertical pupil diameter alone has the advantage that it is not affected by pupillary constriction and dilation, or by the subject moving his/her head toward or away from the eye tracker camera (which would alter apparent pupil diameter).

By definition, eyelid closures during fixations are partial closures, since complete or near-complete closures result in data drop-out and cannot, therefore, be part of a fixation.

5. Fixation Data

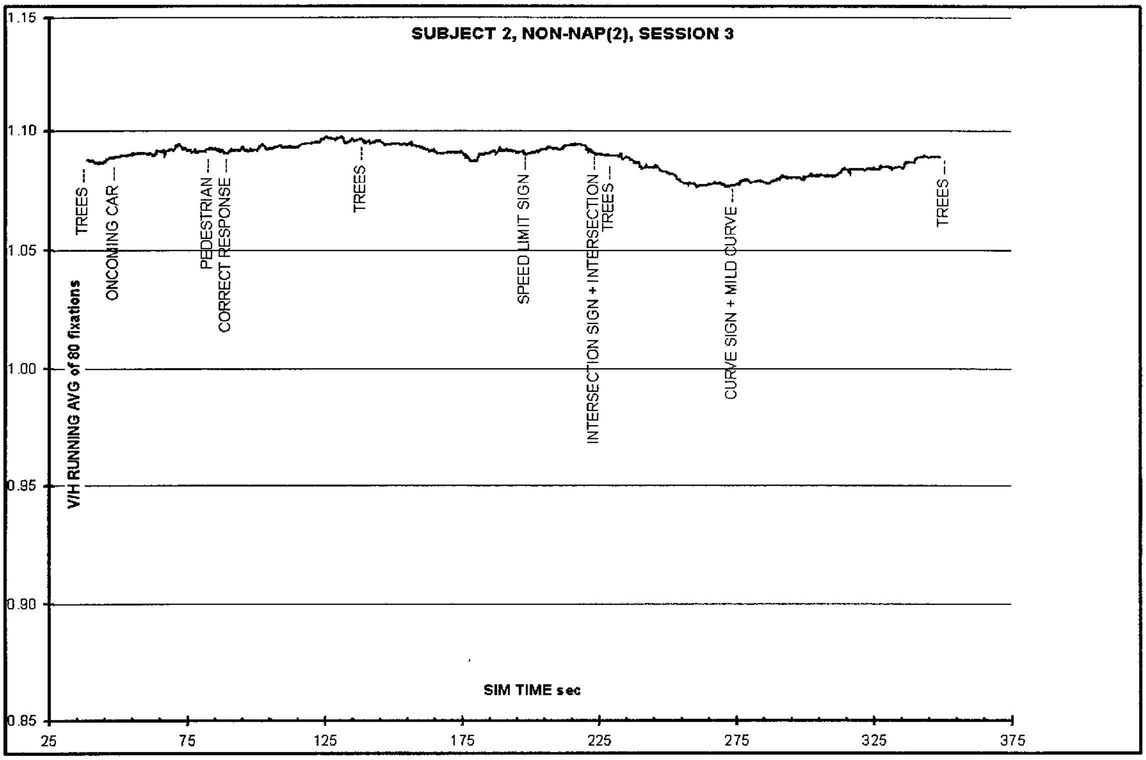
V/H During Fixations

Fixation analysis was performed on seven complete driving runs taken from three subjects. The number of fixations per run ranged from about 11,000 to 15,000. Data representing V/H for each individual fixation in each of the seven runs are shown in Appendix J, Figures J1 - J4. One of these runs (Run 3 in Appendix J, Figure J1) was also plotted as a running average across eighty consecutive fixations and shown in Appendix J, Figure J5. In addition, running averages (across eighty consecutive fixations) were obtained for 5-minute data segments taken at the beginning of run, and 10-minute segments starting 5 minutes before and ending 5 minutes after selected off-road accidents. An example of this running average for an early 5-minute segment and a later 10-minute segment(including and after an off-road accident) is shown in Figure 16. Other samples are shown in Appendix J, Figures J6 and J7.

The figures representing complete runs also show all off-road accidents that occurred during the runs, while the figures representing short driving segments show all stimulus and response events. Off-road accidents were divided into two categories: those preceded in the last 30 seconds by wind gusts, by an oncoming vehicle encroaching on the driver's lane, or by another off-road accident are referred to as "provoked" accidents, while all others are considered "unprovoked". The reason for this distinction is that while most unprovoked accidents are attributable primarily to fatigue, provoked accidents may be accounted for by other factors. For example, avoiding an encroaching vehicle requires moving slightly to the side of the roadway. How far drivers can move to the side without crashing is something that they can only learn through trial and error, and this learning process was often incomplete at the start of the night driving runs.

Two aspects of the V/H data require clarification. First, the highest concentration of data points, especially early in a run, is around a value of V/H slightly greater than 1.0, which implies that the vertical pupil diameter is slightly greater than the horizontal. This effect appears to be due to a slight inaccuracy in the calibration of the eye tracker at the start of the run.

The second aspect is that V/H occasionally reaches values of 1.4 or more, which are too high to be due to calibration inaccuracies. Rather, they appear to represent fixations with the eyeball rotated to the left or right (as when subjects are fixating the rear-view mirrors). Because of the curvature of the cornea, the horizontal diameter recorded by the eye tracker would, in such cases, underestimate real pupil diameter. Close examination of EOG data recorded concurrently with eye tracker data showing high V/H values confirmed this interpretation.



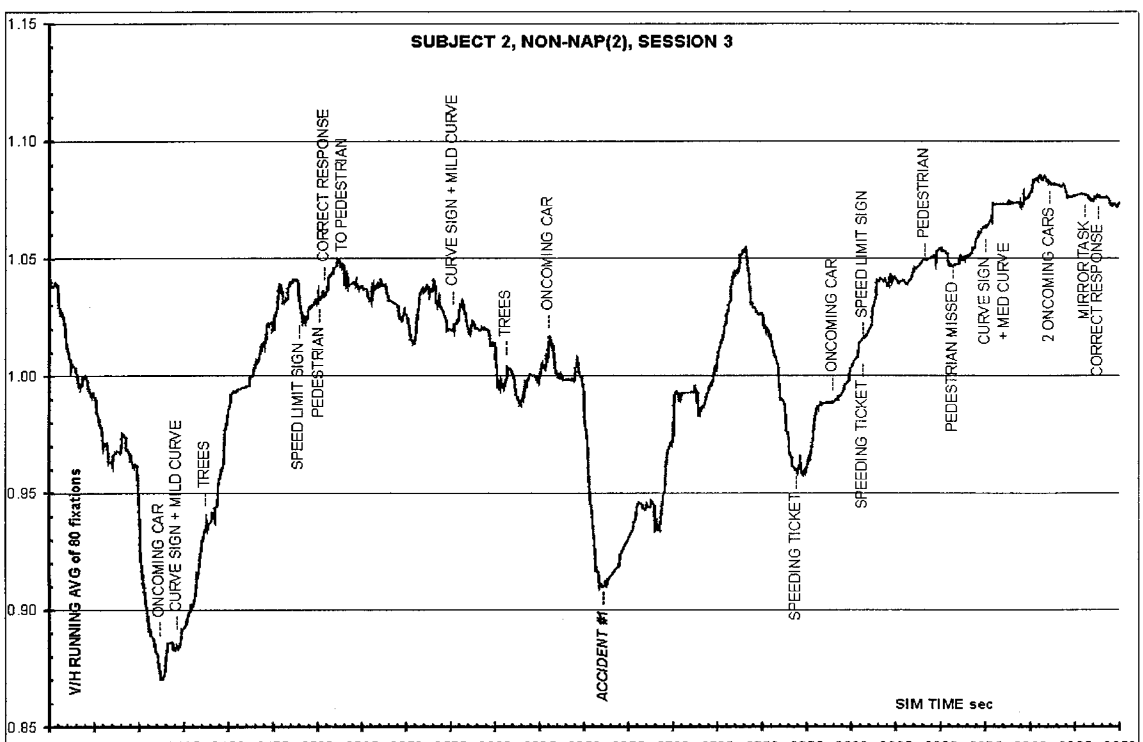


Figure 16. Pupil occlusion measure plotted for a single subject early in a driving session (top plot) and later in the same driving session (bottom plot) during a period in which an accident occurred. The data are average vertical pupil diameter during fixations divided by average horizontal pupil diameter during fixations (V/H), computed as a running average across eighty consecutive fixations. Lower values correspond to more pupil occlusion (less of the pupil exposed).

Finally, there were also instances where the eye tracker misinterpreted certain reflections off the subject's face (or glasses), treating them as pupillary images and calculating vertical and horizontal "pupil diameter". Such instances, however, rarely satisfied the criteria for fixations, and are therefore unlikely to have affected the data in any significant way.

The data from the seven driving runs illustrate the following points:

(1) At the start of a driving run, V/H values are tightly clustered near a value of 1.0 (or slightly higher, as explained above), indicating that the drivers' eyes were fully open during most fixations.

(2) Later in the run, the drivers start to show isolated episodes with low V/H values, indicating partial eye closures. As the run progresses, these episodes increase in frequency and/or duration, eventually merging into one long stretch with many partial closures.

(3) The initial portion of the driving run, when V/H values are clustered near 1.0, is considerably shorter in the later than in the earlier runs (see Appendix J, Figures J1, J3 and J4).

(4) Unprovoked off-road accidents generally occur during episodes of low V/H values. This is best illustrated in Figures 16, and Appendix J, Figures J6, and-J7 showing V/H (running average across eighty consecutive fixations) during the 10-minute periods surrounding such accidents.

(5) Unprovoked off-road accidents, even the earliest ones in a run, are also consistently preceded by drops in V/H occurring over a period of several minutes. The earlier accidents are often followed by a recovery period during which the eyes are fully open again (e.g., first accident of Run 3 in Appendix J, Figure J1, first accident of Run 2 in Appendix J, Figure J3). This can be more clearly seen in the eighty fixation running averages as shown in Figure 16. In this case an oncoming car event and a traffic ticket (indicated by a police siren) also occur at relatively low V/H values and are immediately followed a recovery period in which the eyes are open.

(6) The earliest periods with low V/H values to occur in a driving run are not always followed by an off-road accident (e.g., Run 3 in Appendix J, Figure J1, Run 2 in Appendix J, Figure J2, Run 2 in Appendix J, Figure J4).

The V/H plots also show brief periods with few or no fixations, especially in the later runs and in the later portions of the runs. This is consistent with our observation that the drivers move more as they get tired, resulting in more data drop-outs.

Fixation Duration

The duration of all fixations recorded during the same seven driving runs is plotted in Appendix K, Figures K1-K4. Unlike the corresponding V/H data, fixation duration shows no discernible pattern, either within a run or across successive runs. Furthermore, there is no apparent relation between fixation duration and off-road accidents.

6. Discussion

Blink-Related Measures

TOD effects were observed in all four blink duration measures, but they were most pronounced for data loss and overall blink duration. These two measures also showed the clearest TOT effects, which consisted either of a sustained increase across successive 30-minute blocks, or an initial increase followed by a decrease. These results are thus very similar to those observed with the two EOG-derived blink duration measures, even though the latter measures were different from those abstracted from the eye tracker data, and the EOG and eye tracker data sets selected for analysis did not completely overlap.

Examination of pupil diameter at the start of blinks showed no consistent TOD or TOT effects.

Fixation Data

There is considerable evidence that during sleep deprivation and/or long work hours, alertness levels and performance ability do not show a steady, monotonic decline. Rather, a person will first experience brief, intermittent episodes of drowsiness and performance impairment, followed by a return to normal performance levels. These episodes, termed blocks (Bills 1931) or lapses (Williams et al. 1959), increase in frequency and duration over time

The V/H data described above show exactly that pattern, which suggests that partial eye closures are a sensitive index of drowsiness. In addition, all unprovoked off-road accidents were preceded by frequent sharp drops in V/H occurring over periods of a few minutes or more. In the data analyzed, the shortest such period lasted about 3 minutes (first accident, Run 2 in Appendix J, Figure J3). Thus, drops in V/H occurred early enough before an accident to be useful as a component of a feedback loop aimed at alerting the driver and implementing countermeasures.

The data also show that periods with frequent drops in V/H are not invariably followed by off-road accidents. This, of course, is not unexpected, since in real life, drivers do not have a traffic accident every time they feel drowsy. It is possible, however, that some of the periods with low V/H values are associated with other, more subtle signs of driving impairment, e.g., slower reaction times, or increased variability in lane position.

IV. ALGORITHM DEVELOPMENT

A primary aim of this pilot study was to develop and propose a preliminary version of an algorithm for detecting changes in alertness level and predicting impaired driving performance using eye tracker data. The preliminary algorithm described below is based on data obtained in this 6-month pilot study. Background information is based on a critical review of pertinent literature.

A. BACKGROUND

Wierwille and his colleagues have proposed several algorithms for detecting drowsiness based, at least in part, on a measure of slow eye closure (SEC) known as PERCLOS, which is defined as the proportion of time the eyes are 80% to 100% closed. This is usually measured from videotape recordings of the subject's face, by having an experimenter manually track the subject's eyelids using a linear potentiometer.

Initial laboratory experiments with sleep deprived subjects driving a car simulator showed good correlation between PERCLOS and several indices of driving performance (lane-, steering-, and accelerometer-related measures), as well as performance on a secondary task and subjective evaluations of drowsiness. Based on this correlation, estimates of PERCLOS, termed ePERCLOS, were derived using different combinations of driving and secondary task performance measures. Algorithms for drowsiness detection were then generated based on PERCLOS, ePERCLOS, various lane-related measures, and/or secondary task performance.

An experiment designed to validate the drowsiness algorithms was performed on subjects driving a moving-base simulator from 12:15 a.m. to 3:00 a.m. (Wierwille et al., 1996, part II). Although detection of drowsiness was obtained, all detection was based entirely on driver performance, and none on eye closure.

A subsequent study (Wierwille et al., 1996, part III) was performed after correcting some procedural problems in the previous experiment and developing revised algorithms. However, the results showed only low correlation between drowsiness measures based on eye closure and performance-related measures. The authors concluded that "to avoid run-off-road accidents in a drowsy driver detection system, an auxiliary system that can measure lane position directly and nearly instantaneously must be used" (p. 39).

These results indicate that, while increases in PERCLOS are a clear sign of drowsiness, driving performance (at least in a simulator) may start to deteriorate before any observable changes in that measure. This suggests that algorithms designed to predict drowsiness and impaired driving performance cannot be based on complete (or near-complete) SECs alone.

The algorithm proposed for this study is based on an index of eye closure, weighted by the degree of closure. In principle, the eyes are closed during blinks and during SECs. Many studies have looked at either blinks or SECs as they relate to performance (including driving performance), to TOD and TOT, and to other behavioral and physiological indices of drowsiness. The proposed algorithm includes eye closure due to both blinks and SECs.

B. RATIONALE

The two major categories of eye closure, blinks and SECs, can each be subdivided into two components: the former into "normal" blinks and long duration blinks, and the latter into partial and complete (or near-complete) SECs.

The preliminary algorithm proposed here encompasses all four components. There is evidence from this study that demonstrates that each of these components is sensitive to fatigue effects, showing systematic changes in frequency and/or duration associated with TOD, TOT, and/or performance impairment. Other representative studies supporting this are listed below:

1. "Normal" blinks:	This study	Stern et al. (1994 b; 1996)
2. Long duration blinks:		Morris (1984)
3. Partial SECs:	This study	
4. Complete (>80%) SECs:	This study	Wierwille et al. (1996)

The algorithm proposed here has the following advantages:

1. It is based on a single, but all encompassing eye closure variable. This variable is the proportion of time the eyes are closed (partially or completely), and therefore can be implemented with an eye tracker system.

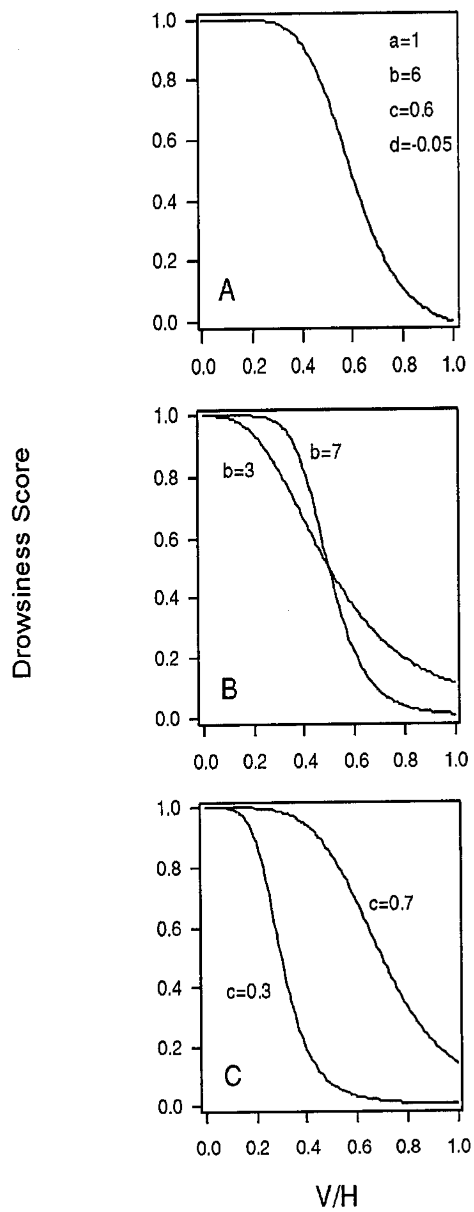
2. That variable combines all eye closures, regardless of whether they are due to blinks (including long duration blinks) or to SECs (including partial SECs). Thus, it reflects the increase in blink frequency and duration which starts to occur at early stages of task performance degradation, the increase in partial eye closures that precedes off-road accidents (as shown in this study), and the increase in SECs which occur in the later stages of degraded alertness when drowsiness levels are dangerously high.

3. The algorithm does not require the added complexity of distinguishing between "normal" blinks, long duration blinks, and SECs. Such distinctions are often arbitrary, especially in very tired subjects.

C. PRELIMINARY ALGORITHM

The algorithm is based entirely on a measure of eye closure, or, more precisely, of pupillary occlusion, consisting of the ratio of vertical pupil diameter over horizontal pupil diameter, or V/H. The reasons for using V/H rather than vertical diameter alone were summarized in Section III.J.4.

The following 4-parameter logistic function is used to derive a drowsiness score weighted by the degree of eye closure:



Drowsiness score = (a-d) / [1 +$(x/c)^b$] + d

where: x = V/H

a = asymptotic maximum score of V/H

b = slope factor (the slope is negative when b>0)

c = inflection point

d = asymptotic minimum score

The function is plotted in Figure 17, which shows V/H going from 0 (eyes fully closed) to 1 (eyes fully open) on the x-axis, and the corresponding drowsiness score, which also ranges from 0 to 1, on the y-axis. Note that there is an inverse relationship between drowsiness score and V/H: the smaller V/H the higher the drowsiness score.

The curve in Figure 17a was generated using the following parameters:

a = 1.0

b = 5

c = 0.5

d = 0.0

Figure 17. Drowsiness Algorithm

The advantage of a logistic function over other functions (e.g., linear or exponential) is that the upper and lower ends of the curve can be differentially manipulated to give more or less weight to one extreme of the V/H distribution or the other. Figure 17b shows the effects of making the slope steeper (b=7) or shallower (b=3), while Figure 17c shows the effects of changing the inflection point, thereby lengthening the flat portion at the upper (c=0.7) or lower (c=0.3) end of the curve.

Parameter values would be assigned in such a way as to maximize the correlation between drowsiness scores and driving performance measures. For example, increasing the slope at the high end of the V/H distribution would give more weight to small, partial, eye closures, which we found to be a good predictor of driving impairment. Conversely, by decreasing the slope at the low V/H end, near-maximum drowsiness scores would be reached before full eye closure. This would be analogous to the PERCLOS measure, which considers any closure above 80% as a full closure.

D. IMPLEMENTATION

Applying the algorithm required distinguishing between a data drop-out resulting from a blink and all other data drop-outs (driver looking in side view mirrors, excessive head or body movement to fight fatigue, etc.). Manual editing of the data was used to distinguish between these two types of drop-outs using the following criteria. A series of data points where either the vertical (V) or horizontal (H) pupil diameter was zero was considered an eye closure, and assigned a V/H value of zero, if it was immediately preceded by three or more data points with decreasing V values and immediately followed by three or more data points with increasing V values. If these criteria were not met, the data were considered of the second type and no value was assigned to V/H.

Since the data were sampled at 60 Hz, editing was laborious. To meet the scheduling commitments of the study, there was only time to apply the algorithm to a limited number of short driving segments. An example is shown in Figure 18. The data represent three consecutive driving segments, each lasting 5 minutes (18,000 data points), with the first segment starting 5 minutes after the beginning of the driving run. Figure 18 shows a running average of the drowsiness scores calculated over 10 seconds (600 data points). Algorithm parameters used were: a = 1.0, b = 6.0, c = 0.7, d =-0.05.

The first 5-minute segment, appearing to be an alert driver segment, shows generally low average drowsiness scores, with occasional small peaks, none of which exceeded 0.4 (Figure 18a). An isolated peak in drowsiness, reaching a score of about 0.6, occurred approximately 640 seconds after the start of the driving run (Figure 18b). A series of larger peaks then followed, beginning 860 seconds after the start of the run and culminating in an unprovoked off-road accident 140 seconds later (Figure 18c). During the remainder of the last segment, drowsiness scores were generally low, except for a single large peak approximately 90 seconds after the accident.

Examination of the EOG records for the same driving segments showed that many eye closures (blinks and SECs) were correctly identified in the eye tracker data. However, there were also several instances where real eye closures were considered artifact, as well as instances where data drop-outs caused by head movements were considered eye closures.

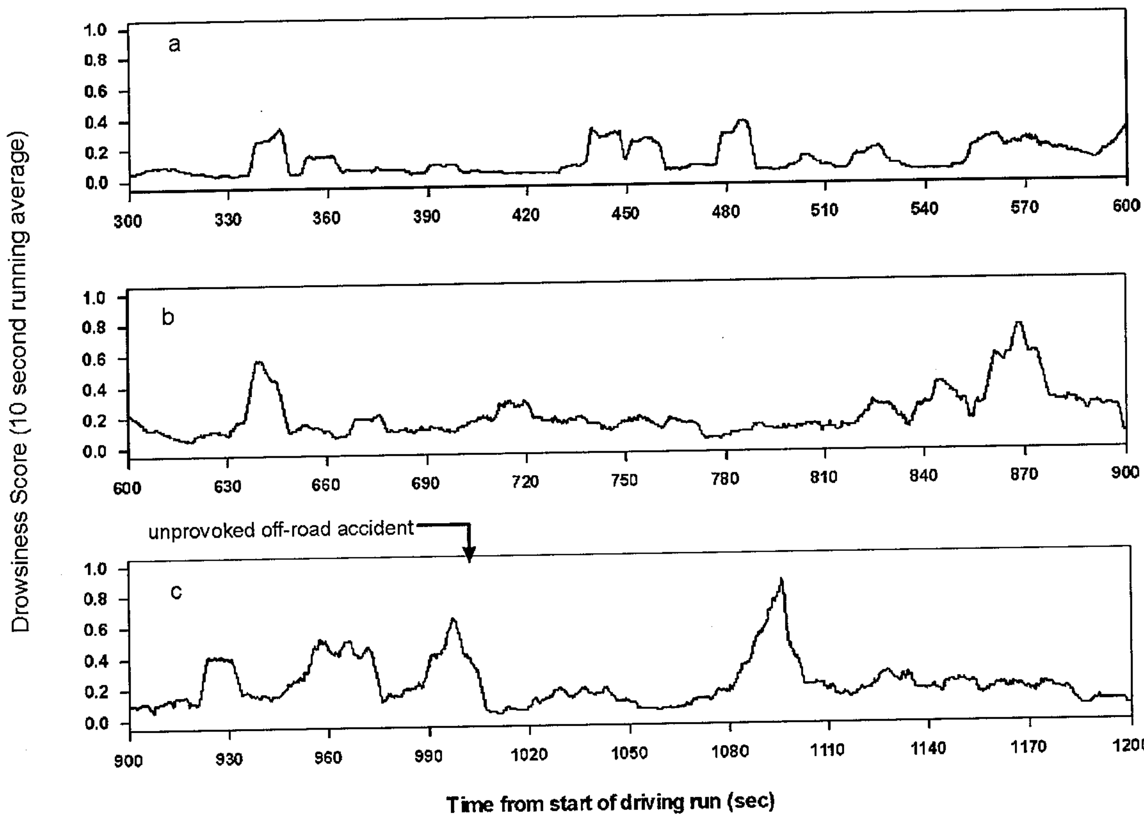


Figure 18. 10 second running average of drowsiness scores during the first 20 minutes of a driving session. Algorithm parameters were: a=1.0, b=6.0, c=0.7, d=-0.05.

In summary, the algorithm described above appears to hold promise for detecting drowsiness and predicting driving impairment. It is clear however, that the algorithm could not be sufficiently evaluated in the available time frame designated to the study. A follow on study is needed to more effectively evaluate the algorithm.

V. CONCLUSIONS AND RECOMMENDATIONS

A. OCULAR DYNAMICS AS A PREDICTOR OF REDUCED ALERTNESS

A major objective of this pilot study was to evaluate the potential of an eye tracking system for detecting drowsiness in professional long-haul drivers, with the aim of developing and testing a preliminary algorithm for predicting fatigue-induced driving impairment. Concurrent EOG recordings collected during this study served to corroborate the utility of the eye tracker data as well as identify other potentially useful ocular parameters that could be employed in an eye tracker specifically designed for this application.

Criteria used to determine if identified ocular parameters are potentially good candidates for detecting drowsiness are as follows:

1. If parameters showed systematic changes across successive driving runs (TOD effect) or across successive segments within a driving run (TOT effect), or if they showed changes associated with off-road accidents,

2. If parameters could effectively be measured using an existing eye tracker or EOG equipment such that an eye tracker specifically designed for this application could potentially be developed.

3. If parameters could be used in an algorithm to provide early and reliable indications of reduced alertness.

4. If parameters could be measured unobtrusively.

The 6-month pilot project was successful in identifying six specific parameters that meet the defined criteria as stated above. These parameters are summarized as follows:

1. Blink duration measures derived from eye tracker as well as EOG data showed unquestionable evidence of TOD and TOT effects. The latter consisted of a sustained blink duration increase, or an initial increase followed by a blink duration decrease, especially in the later driving runs. Blink closing duration, effectively derived from the EOG, was also longer in the minute preceding off-road accidents than immediately after, and considerably longer than just before or at the start of the early driving runs.

2. Blink frequency, as derived from EOG data, showed favorable TOD and TOT effects which were similar to those of blink duration. However, the results with blink frequency were less consistent than with blink duration.

3. Partial eye closures during fixations, measured from eye tracker data as the ratio of vertical to horizontal pupil diameter (V/H), showed clear TOD and TOT effects. This measure showed consistent and obvious V/H ratio decreases, especially when

compared to alert state measures, which led up to an off-road accident. The results show that this parameter is particularly impressive as an indicator of degraded alertness by at least 2 to 3 minutes before an accident and probably by as much as 10 to 12 minutes before an accident. Furthermore these fluctuations in the parameter seem to show a logical relationship with events that can be assumed to induce brief periods of increased alertness (e.g., oncoming car, police siren, etc.).

4. Eye closures occurred with a relatively high frequency in the minute preceding off-road accidents and showed a further, dramatic increase starting 20-30 seconds before the accidents.

5. Saccade frequency, measured from EOG data, was markedly higher in the 30 seconds following as compared to the 60 seconds preceding off-road accidents. Although the study did not conclusively identify this parameter as an effective measure of reduced driver alertness, it did show enough change around unprovoked off-road crashes to deserve further study.

6. Large head and body movements, although not specifically measured in the study, were observed as being obvious indicators of fatigue. These movements should be explicitly measured in a follow-on study.

Measures of blink amplitude, horizontal pupil diameter immediately preceding blinks, and fixation duration were also investigated. While there was weak evidence to show that changes in one or more of these parameters may have any correlation with unprovoked off-road accidents, time constraints of the study did not allow for a thorough analysis. These results suggest that further analysis of these parameters is warranted in a follow-on study

In summary, the ocular parameters showing great promise in this application are partial eye closures during fixations, SECs, blink duration, and blink frequency. An even stronger case can be made for the latter two measures when one considers that the results from this study were based on "normal" blinks only. By extending the analysis to include long duration blinks the association between blink frequency and amplitude on one hand, and TOD, TOT, and driving performance on the other is expected to be even more compelling. Data from other studies suggest that long duration blinks are highly correlated with fatigue-induced performance impairment.

Based on these results a preliminary algorithm was proposed which uses V/H as a measure of eye occlusion. The advantage of using this measure in the algorithm is that it takes into account all eye closure measures identified in this study as good indicators of degraded alertness. These include blink frequency, "normal" and long duration blinks, partial SEC, and complete SEC measures. A 4-parameter logistic function is used to assign a drowsiness score to each V/H value, weighted by the degree of eye closure.

Initial efforts, although limited, to apply the algorithm to short data segments in the study showed promising results. Drowsiness scores early in the run, presumably while the driver is alert, showed generally low scores. This was followed by larger scores later in the run and then larger peaks leading up to an unprovoked off road accident. However, time limitations of the study did not allow for extensive application of the algorithm to longer data segments. Thus, follow-on efforts will be required to further test, validate, and perhaps expand on the algorithm.

B. PROPHYLACTIC NAPPING AS A FATIGUE COUNTERMEASURE

The second major objective of this pilot study was to evaluate the effects of a scheduled 3-hour afternoon nap on nighttime driving behavior and on several other measures of alertness and performance in professional long-haul drivers.

Several of the measures analyzed suggest that nighttime alertness levels were higher and performance was better in the Nap than in the No-Nap condition. Evidence of such beneficial effects of the scheduled nap include:

1. Lower subjective fatigue and sleepiness scores.
2. Faster, more accurate, and/or less variable performance on the Performance Assessment Battery tests.
3. Improved performance on the driving simulator, including larger rewards and/or smaller penalties, fewer crashes, less variability in lane position, and better identification performance on the two secondary tasks.
4. Lower spectral EEG amplitudes in the theta and 6-10 Hz frequency bands.

Although the scheduled nap was intended as a prophylactic nap, because the subjects' sleep on the preceding night was restricted to 5 hours implies that the nap was partly recuperative, rather than being strictly prophylactic. It is worth noting, however, that partial sleep restriction is not unusual during long-haul (over-the-road) driving, even when the drivers comply with hours of service regulations. For example, the results of the recently completed "Commercial Motor Vehicle Driver Fatigue and Alertness Study", which involved eighty drivers in the U.S. and Canada, showed that during duty tours, average time in bed was just over 5 hours per day, or about 2 hours less than what the drivers considered to be their ideal sleep time (Mitler et al., 1997). In retrospect, however, the sleep restriction was perhaps unnecessary, given how difficult it was for the subjects to complete the nighttime driving sessions on the simulator.

As stated earlier, three of the subjects were unable to start or complete the third night driving run and were allowed to take an unscheduled nap instead, both in the Nap and No-Nap conditions. Interestingly, the unscheduled naps, which lasted 1-2 hours, had few beneficial effects. For example, subjective sleepiness and fatigue levels remained almost as high as they were at the end of the second night driving run. More importantly, driving performance was worse during the fourth driving run, after the unscheduled nap, than it was during the second, as indicated by larger nominal penalties and more crashes.

Even after the unscheduled naps, driving performance during Run 4 remained better and spectral EEG amplitudes remained generally lower in the Nap than in the No-Nap condition. Finally, while the subjects' performance on the PAB tests improved after the unscheduled nap, this only happened in the No-Nap but not in the Nap condition.

These latter results suggest that prophylactic naps may be more beneficial than recuperative naps during all-night driving situations. Other studies have shown that naps taken early within a long period of sustained wakefulness have more beneficial effects than ones taken at later times (e.g., Dinges et al., 1987). This question could be tested experimentally in future studies of scheduled napping as a fatigue countermeasure in long-haul driving operations.

C. RECOMMENDATIONS FOR FOLLOW ON WORK

The pilot study has provided guidance both for further research on alertness measurement algorithms, and for eventual development of an eye tracker that will be appropriate for operational use in alertness detection.

The test protocol used in the pilot study produced extreme sleepiness in all subjects well before the last driving session. In follow on studies it would be reasonable to eliminate sleep restriction the night before driving sessions (the pilot study limited sleep to 5 hours on the night before driving sessions), and it would also be reasonable to decrease the length and number of driving sessions for each subject. This will make more efficient use of simulator and staff time, allowing the use of more subjects. Very long periods of driving beyond the point at which drivers becomes desperately sleepy are of limited value.

During a follow on study, provision should be made to measure PERCLOS along with eye tracker data during at least some of the driving sessions. In the short term, this can probably be best accomplished using the same semi-automated video tape analysis technique used by Weirwille.

PERCLOS and performance measures should be used as comparative alertness measures during the driving task rather than EEG measures. The current pilot study has shown performance indicators to be preferable to EEG for this purpose, and if EEG is not measured it will be possible to use a magnetic head tracker as described below.

The extreme head and trunk motions observed after subjects became drowsy may prove useful as an alertness detection parameter, but was also the primary cause of eye tracker data loss and difficulties in data analysis. These problems can be minimized during further studies that rely on existing eye tracker equipment by using magnetic head tracking in conjunction with a current remote eye tracker, and by collecting at least some data with a head mounted eye tracker.

The technology does exist to improve the ability of a current remote eye tracker to follow vigorous head motion by providing head position information from an independent head tracker. Such a head tracker was not used for the current pilot study because it interfered with EEG measurements; but it has been recommended above that follow on efforts not record EEG data and interference with EEG will therefore not be a factor.

Eye tracking with head mounted optics does decrease "face validity" by requiring drivers to wear head gear that they normally would not wear, but it will not prevent drowsiness, does not restrict the field of view, and will not be significantly affected by head motion. Of course the eventual eye tracker to be developed for operational use in vehicles must handle the vigorous head motion without any head mounted components. In the meantime, some head mounted eye tracker data would be extremely valuable.

Follow on efforts should use a head tracker not only to assist the eye tracker, but also to explicitly record head movement so that the head motion effect observed during the current pilot study can be quantified and analyzed for inclusion in an alertness detection algorithm. The same existing technology that allows integration of an eye tracker with a magnetic head tracker also provides a means to record the head position data.

An eventual eye tracker system (for operational use in vehicles) can probably operate at 60 samples/second or slightly less. None of the alertness parameters that so far appear the most promising require high temporal bandwidth.

The eventual eye tracker system (for operational use in vehicles) must recognize extreme head motions that seem to be characteristic of extreme drowsiness. It may not be essential to make accurate gaze and lid behavior measures during these violent motions, since the presence of such motion may in itself prove sufficiently indicative of drowsiness, but the system must at least detect these motions. This translates to a requirement that the system detect at least the presence of the eye or some other facial feature over a range that is probably somewhat greater than the standard "head box". Quantitative measurement of head motions, as proposed above for follow on research, will more precisely determine the required range.

VI. REFERENCES

Akerstedt T, Kecklund G, Knutsson A. Manifest sleepiness and the spectral content of the EEG during shift work. *Sleep* 14(3):221-225, 1991.

Akerstedt T, Torsvall L, Gillberg M. Sleepiness in shiftwork. A review with emphasis on continuous monitoring of EEG and EOG. *Chronobiol Intern* 4(2):129-140, 1987.

Bills AG. Blocking: a new principle of mental fatigue. *Am J Psychol* 43:230-245, 1931.

Cabon Ph, Coblentz A, Mollard R, Fouillot JP. Human vigilance in railway and long-haul flight operation. *Ergonomics* 36(9):1019-1033, 1993.

Cajochen C, Brunner DP, Kräuchi K, Graw P, Wirz-Justice A. Power density in the theta/alpha frequencies of the waking EEG progressively increases during sustained wakefulness. *Sleep* 18(10):890-894, 1995.

de Waard D, Brookhuis KA. Assessing driver status: a demonstration experiment on the road. *Accid Anal & Prev* 23(4):297-307, 1991.

Douglass A, Bornstein R, Nino-Murcia G, Keenan S. Creation of the "ASDC Sleep Disorders Questionnaire". *Sleep Res* 15:17, 1986.

Douglass A, Bornstein R, Nino-Murcia G, Keenan S, Miles L, Zarcone VP Jr, Guilleminault C, Dement WC. The Sleep Disorders Questionnaire I: Creation and multivariate structure of SDG. *Sleep* 17:60-167, 1994.

Emerson JD. Introduction to transformation. In Hoaglin DC, Mosteller F, Tukey JW, eds. *Fundamentals of Exploratory Analysis of Variance*. New York: John Wiley & Sons, 1991.

Folkard S, Spelten E, Totterdell P, Barton J, Smith L. The use of survey measures to assess circadian variations in alertness. *Sleep* 18:355-361, 1995.

Gortelmeyer R. On the development of a standardized sleep inventory for the assessment of sleep. In Kubicki S, Herrmann WM, eds. *Methods of Sleep Research.* Stuttgart: Gustav Fischer Verlag, pp S93-S98, 1985.

Horne JA, Östberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Internat J Chronobiol* 4:97-110, 1976.

Iman RL, Hora SC, Conover WJ. Comparison of asymptotically distribution free procedures for the analysis of complete blocks. *J Am Stat Assoc* 79:674-685, 1984.

Kecklund G, Akerstedt T. Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics* 36(9):1007-1017, 1993.

Kirk RE. Experimental Design: Procedures for the Behavioral Sciences, 2nd ed. Belmont, CA: Brooks/Cole, 1982.

Knipling, RR and Wang, JS. Crashes and fatalities related to driver drowsiness/fatigue, *NHTSA Research Note*, November, 1994.

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

Lorenzo I, Ramos J, Arce C, Guevara MA, Corsi-Cabrera M. Effect of total sleep deprivation on reaction time and waking EEG activity in man. *Sleep* 18(5):346-354, 1995.

Mackie RR, Miller JC. *Effects of Hours of Service, Regularity of Schedules, and Cargo Loading on Truck and Bus Driver Fatigue*. HFR Technical Report 1765-F. Goleta, CA: Factors Research Inc., 1978.

Myers JL, Well AD. *Research Design and Statistical Analysis*. New York: Harper Collins Publishers Inc., 1991.

Pfaff V, Fruhstorfer H, Peter JH. Changes in eye-blink duration and frequency during car driving. *Pfluegers Archiv* 362:R21, 1976.

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

Skipper JH, Wierwille WW. Drowsy driver detection using discriminant analysis. *Human Factors* 28:527-540, 1986.

Stern JA, Boyer D, Schroeder D. Blink rate: a possible measure of fatigue. *Human Factors* 36:285-297, 1994a.

Stern JA, Boyer D, Schroeder D, Stoliarov N. *Blinks, Saccades, and Fixation Pauses During Vigilance Task Performance: I. Time on task*. Washington, DC: Federal aviation administration, Office of Aviation Medicine Report No. DOT/FAA/AM-94/26, 1994b.

Stern JA, Boyer D, Schroeder D, Stoliarov N. *Blinks, Saccades, and Fixation Pauses During Vigilance Task Performance: II. Gender and Time of Day*. Washington, DC: Federal aviation administration, Office of Aviation Medicine Report No. DOT/FAA/AM-96/9, 1996.

Torsvall 1, Akerstedt T. Sleepiness during day and night work; a field study of train drivers. *Sleep Res* 12:376, 1983.

Torsvall L, Akerstedt T. Eye closure, sleepiness and EEG spectra. In Koella WP, Ruther E, Schulz H eds. *Sleep '84*. Stuttgart: Gustav Fischer Verlag, pp 464-466, 1985.

Torsvall L, Akerstedt T. Sleepiness on the job: continuously measured EEG changes in train drivers. *Electroenc Clin Neurophysiol* 66:502-511, 1987.

Torsvall L, Akerstedt T. Extreme sleepiness: quantification of EOG and spectral EEG parameters. *Intern J Neuroscience* 38:435-441, 1988.

Wierwille WW, Lewin MG, Fairbanks RJ III. *Research on Vehicle-Based Driver Status/Performance Monitoring, Parts I, II, III.* Washington, DC: National Highway Traffic Safety Administration, Report Nos. DOT HS-808638, HS-808639, HS-808640, 1996.

VII. APPENDICES

APPENDIX A

Table A 1. Interactive events in driving scenarios

Event	Parameters
Pedestrian	Position (4), brightness (2), tree presence (+/-)
Mirror Target	Mirror (2), target location distractor/target
Curves	Curvature (mild, medium, sharp)
Oncoming Traffic	Number of cars (1-4)
Encroaching Oncoming Vehicle	1st/2nd of 2 cars
Signalized Intersection	Traffic signal phase (all green, yellow-stop, yellow-go)
Non-Signalized	Stop Sign or no traffic control
Intersection	
Stopped Vehicle Ahead	Oncoming traffic (+/-)
Workzone	Barrel layout (straight or jogged)
Cross Winds	Strength (2)
Building	Two examples
Plane	Direction (L-R, R-L0
Speed Limit Signs	
Trees	Random arrangement

Table A 2. Event frequencies for the three levels of task demand

Event	Segment Event Frequencies		
	Low Demand	**Med. Demand**	**High Demand**
Pedestrian	2	3	3
Mirror Target	0	3	6
Curves	3	4	5
Oncoming Traffic (group)	0	1	2
Encroaching Oncoming Vehicle	3	3	6
Intersection*	1	1	1
Stopped Vehicle Ahead	0	1	2
Work zone	0	0	2
Cross Winds	0	0	1
Building/Plane	1	1	1
Tree-lined roadway sections	8	8	8
Speed Limit Signs	2	2	2
Uneventful section**	2	1	0

*type of intersection (non-signalized, stop sign, signalized) and stopping decision varied between segments

**for low-demand segments there was one 5,000 ft. section of roadway and one 10,000 ft. section of roadway with no events; for medium-demand segments there was one 10,000 ft. section only.

APPENDIX B

Intentionally blank –(Figures inserted in main body of report)

APPENDIX C

Figures C1-C3. Visual analog scale results.

Figure C1. Visual Analog Scale: FATIGUE

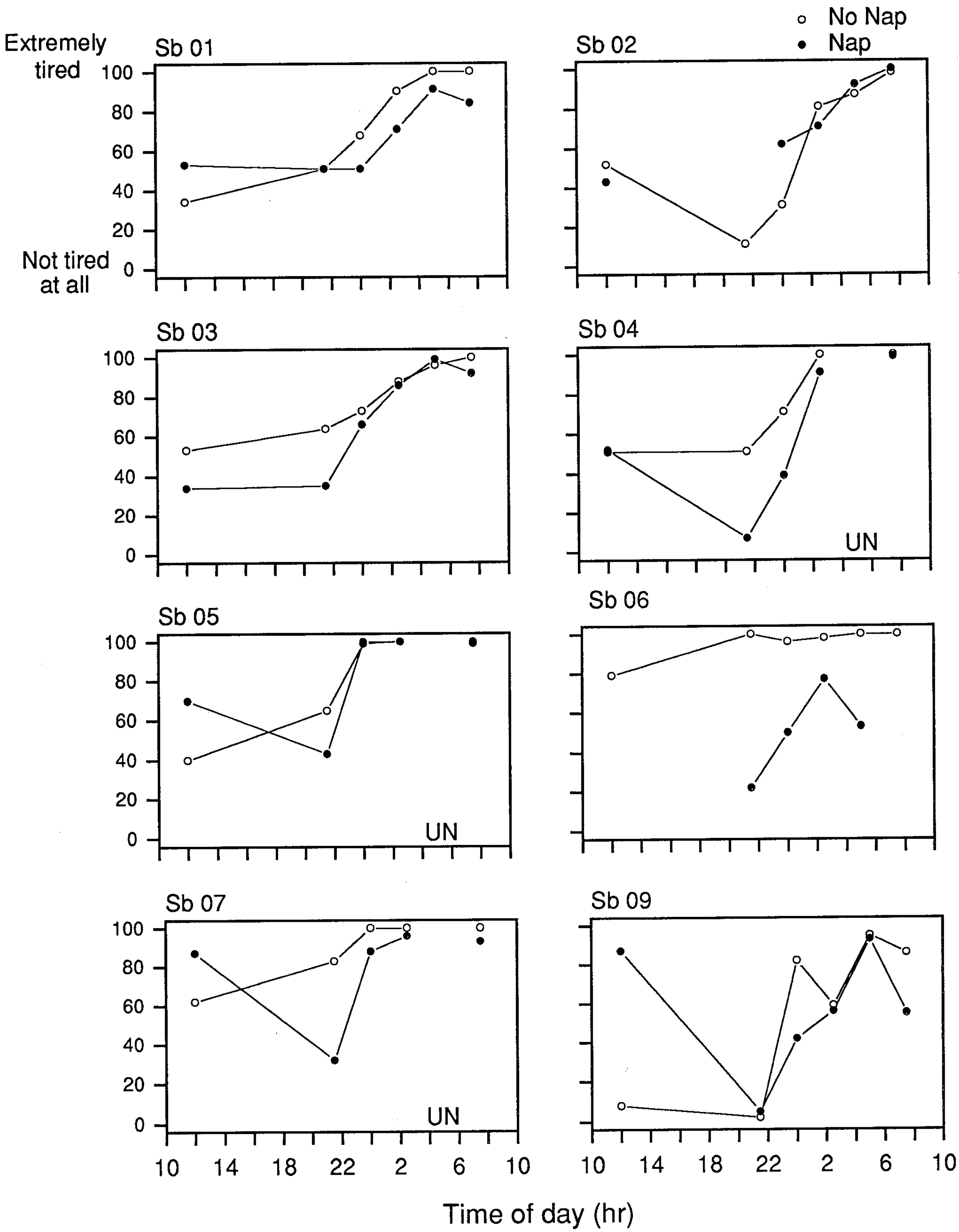


Figure C2. Visual Analog Scale: SLEEPINESS

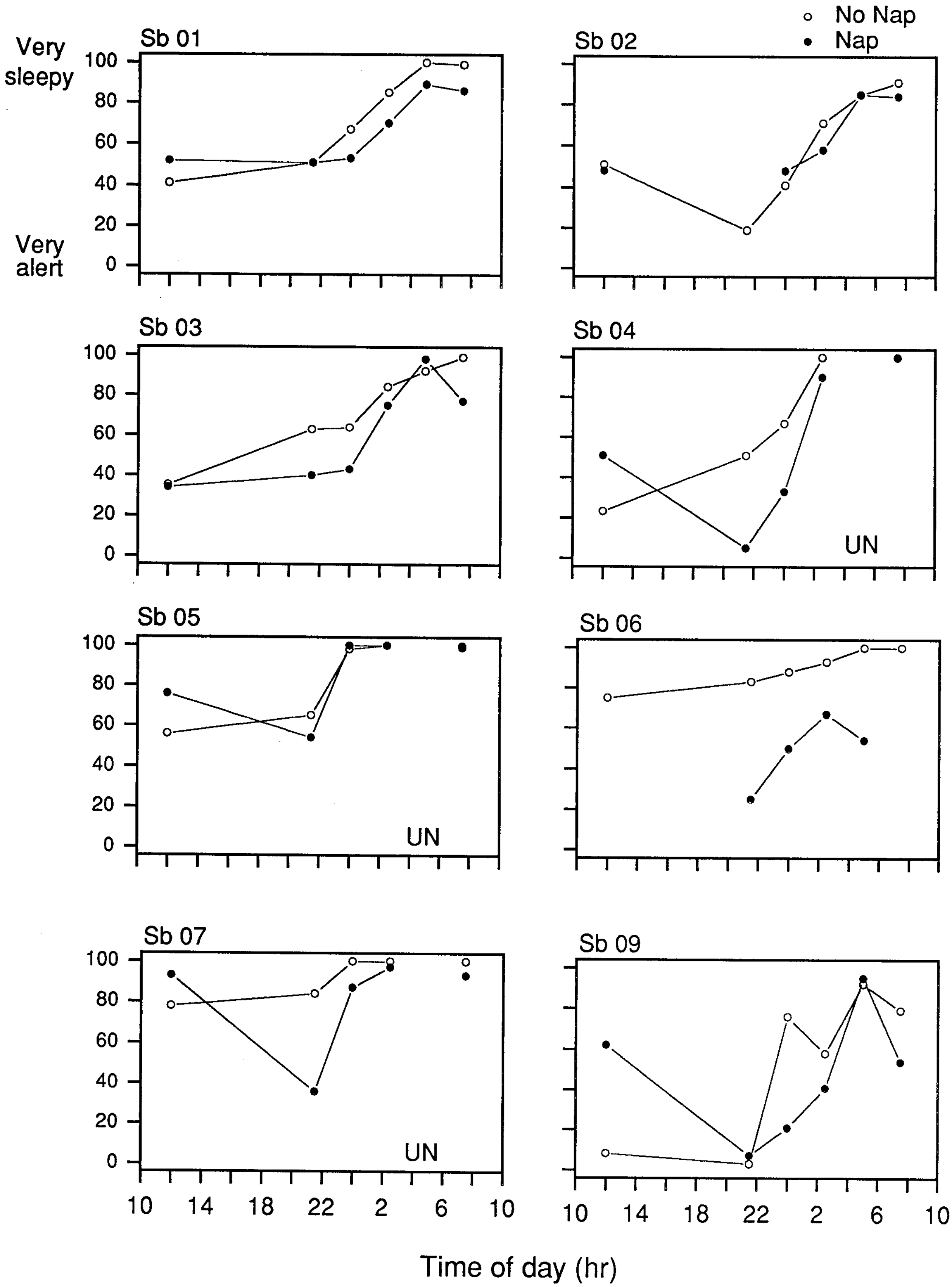
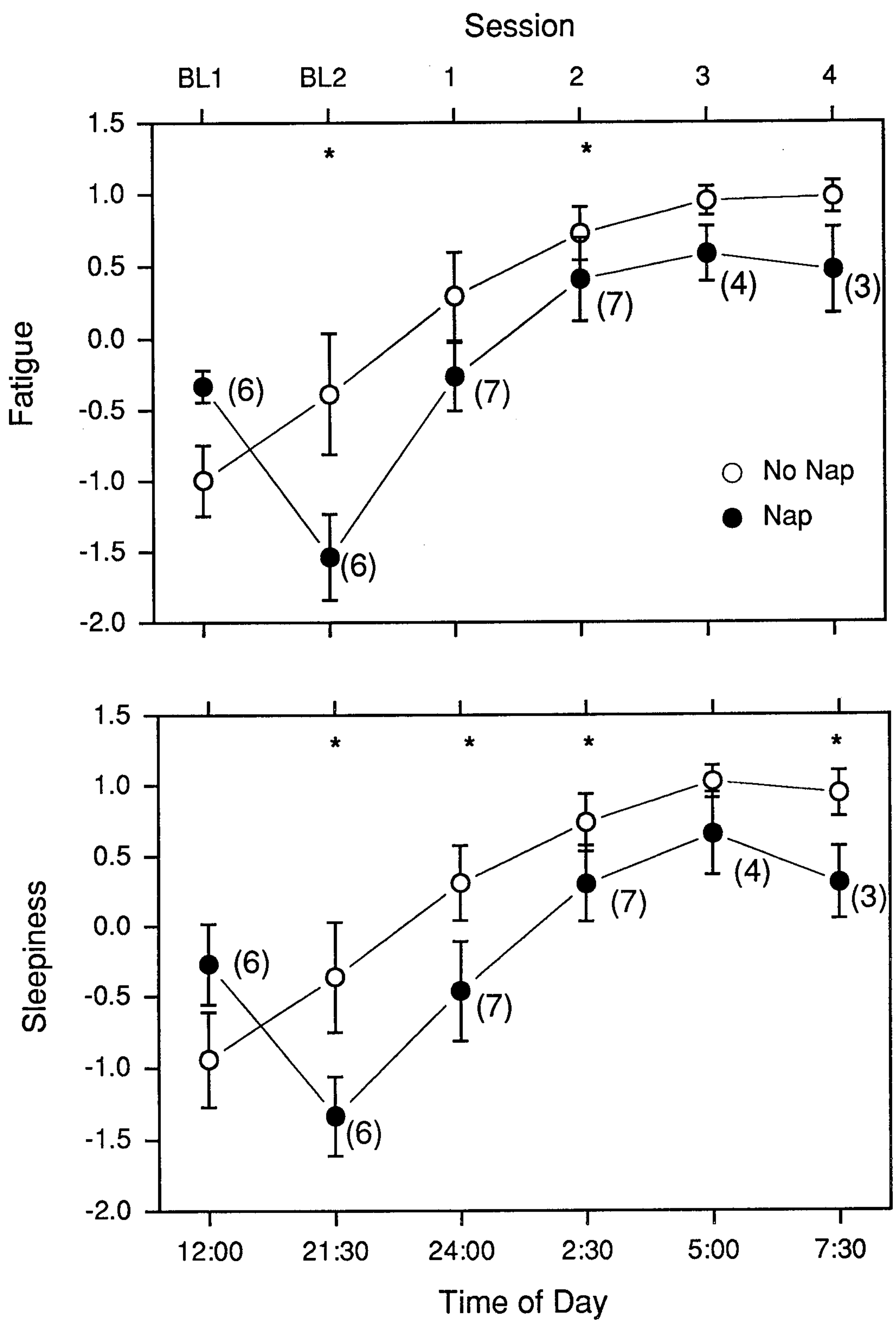


Figure C3. Visual Analog Scales (mean standard score**s** ± **s.s.)**



APPENDIX D

Figures D1-D6. Performance Assessment Battery results.

Figure D1. PAB: Two-letter Memory and Search Task

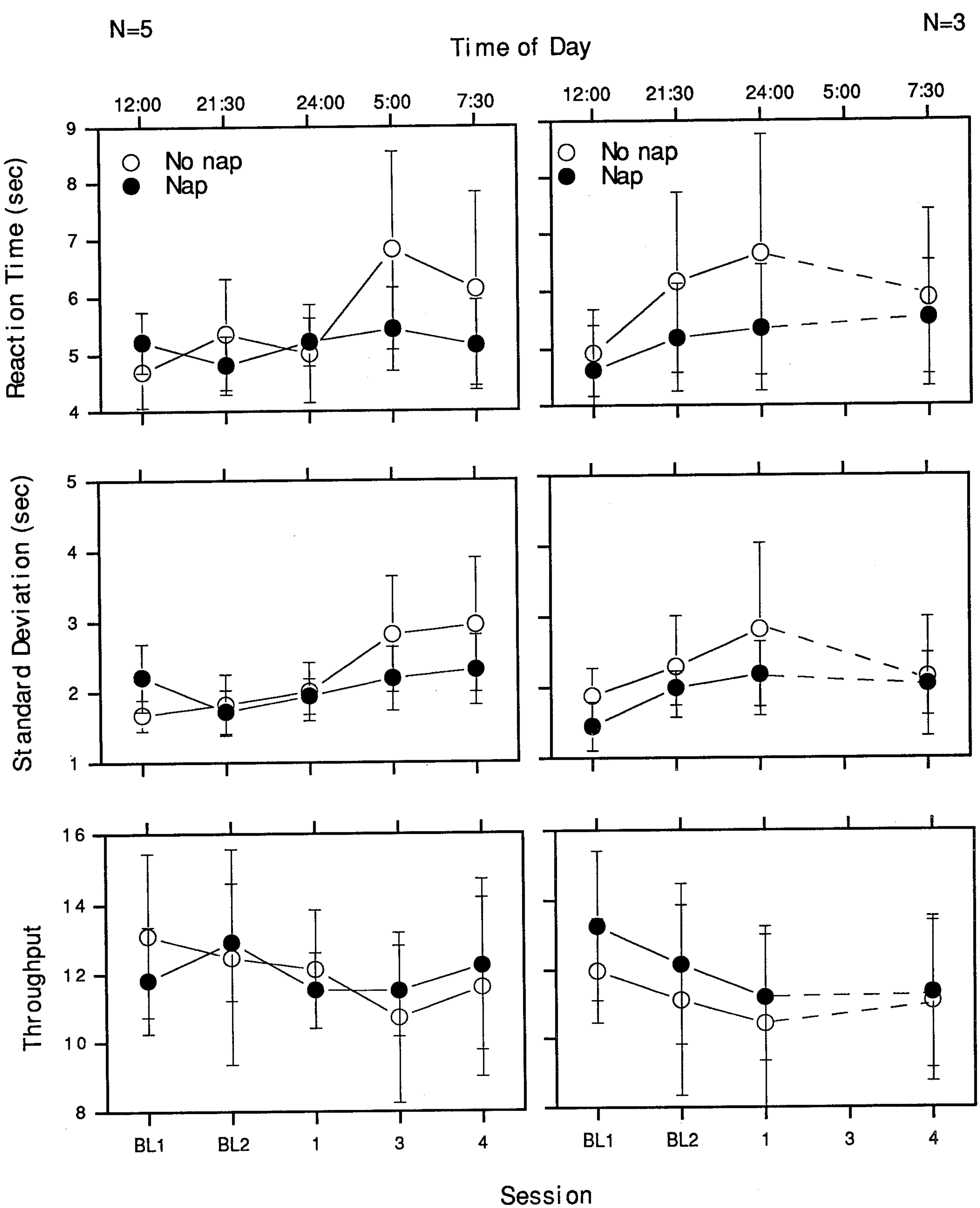


Figure D2. PAB: Wilkinson Four-choice Serial RT

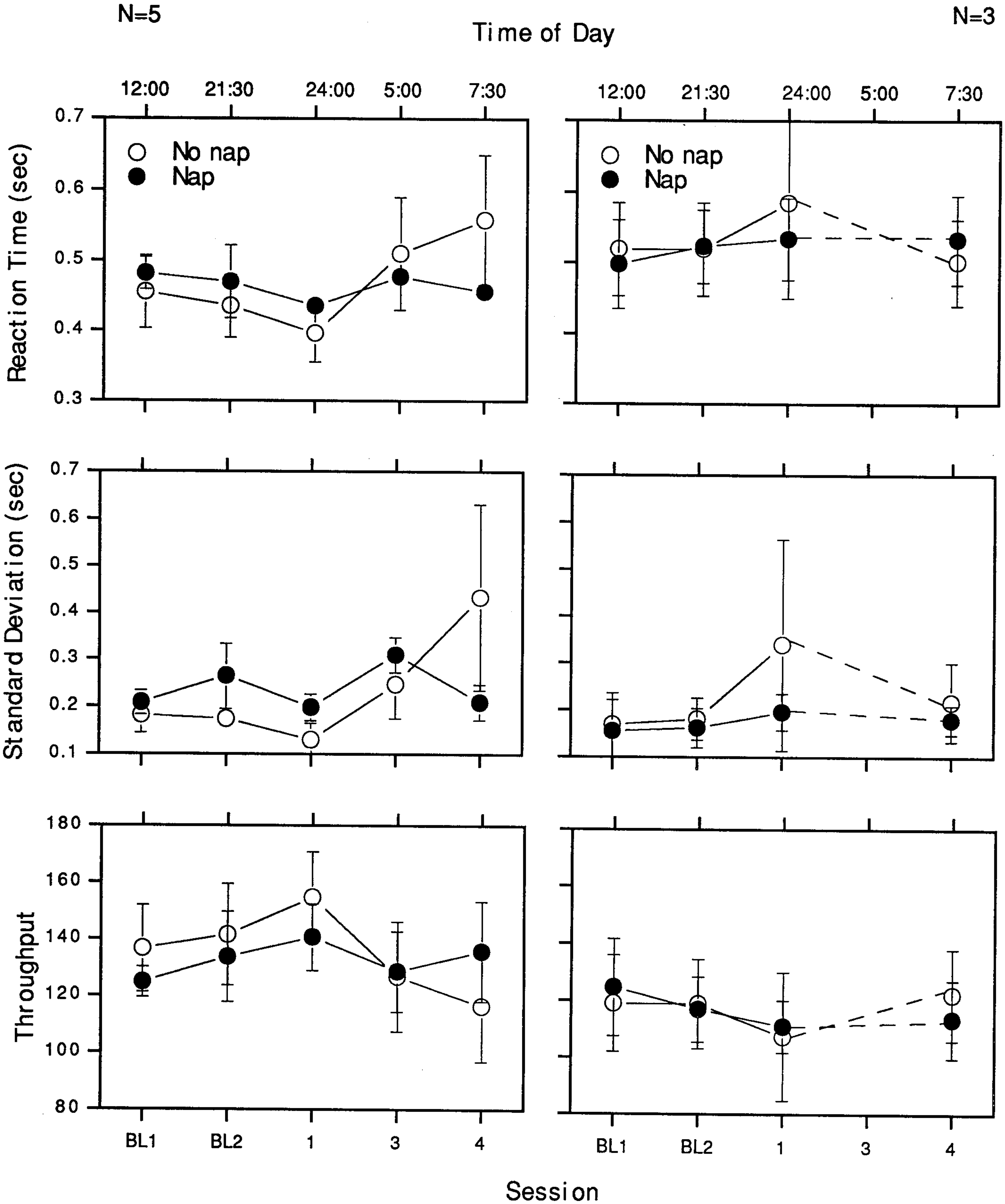


Figure D3. PAB: Time Wall

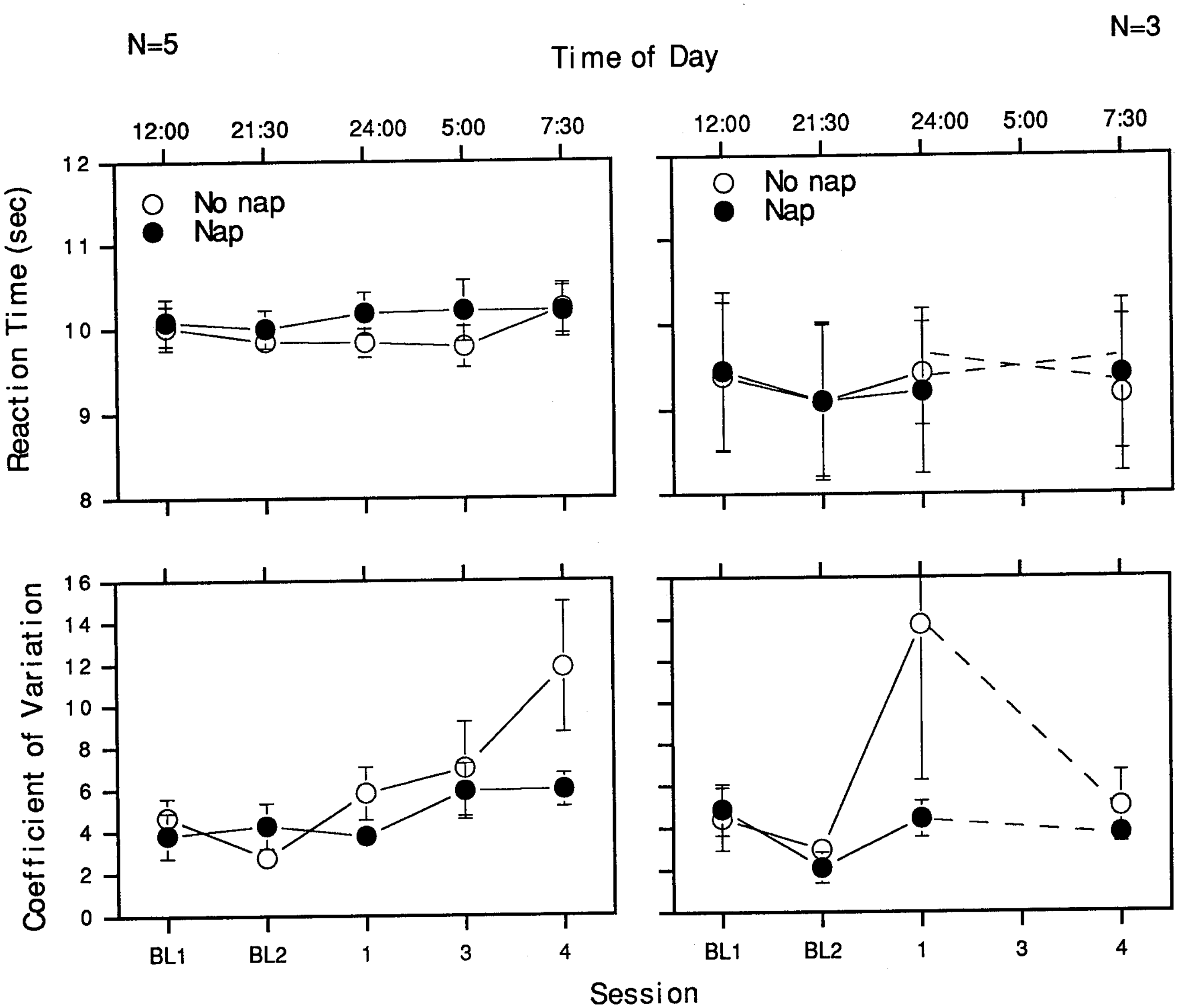


Figure D4. PAB: Two-letter Memory and Search Task

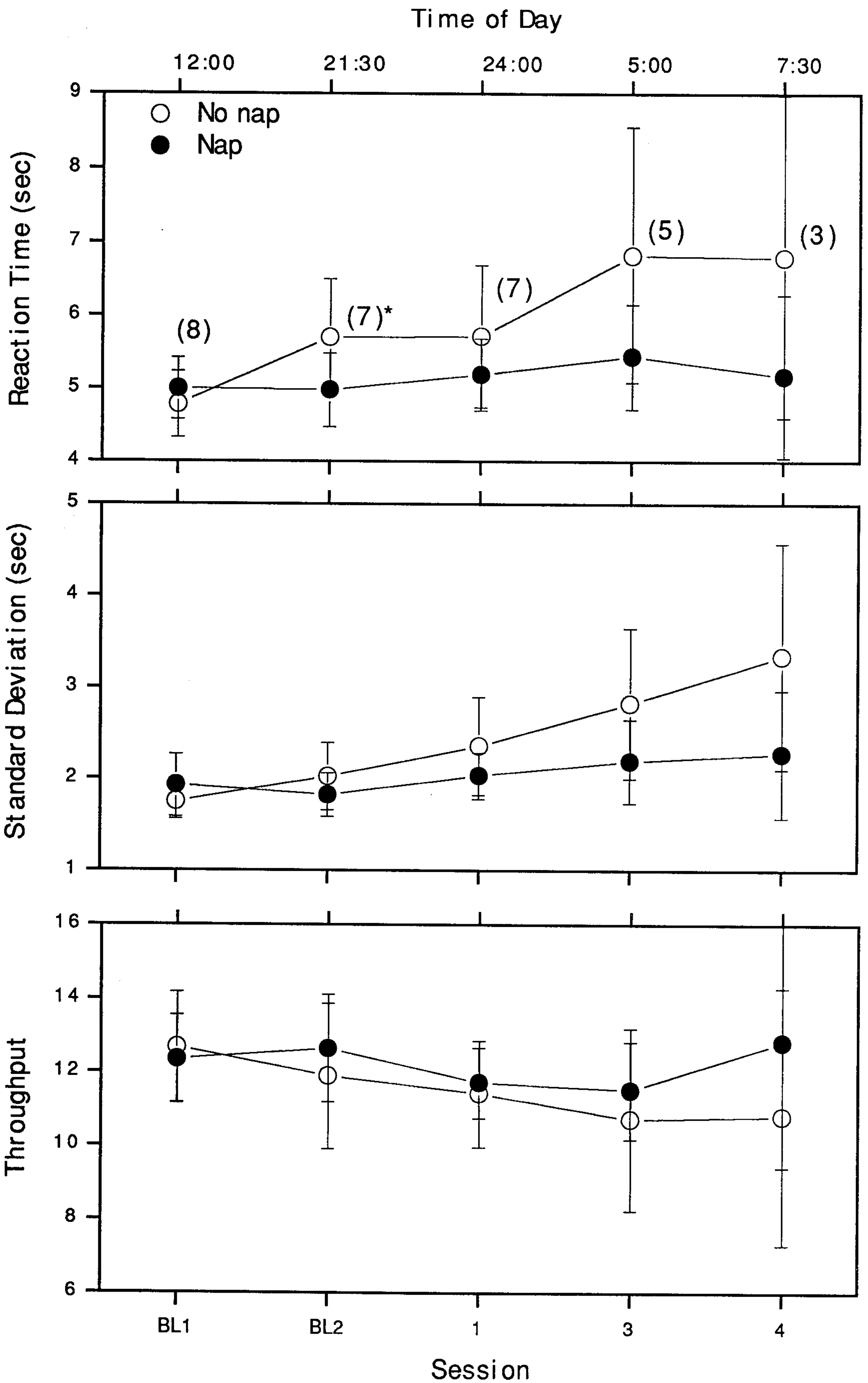


Figure D5. PAB: Wilkinson Four-choice Serial RT

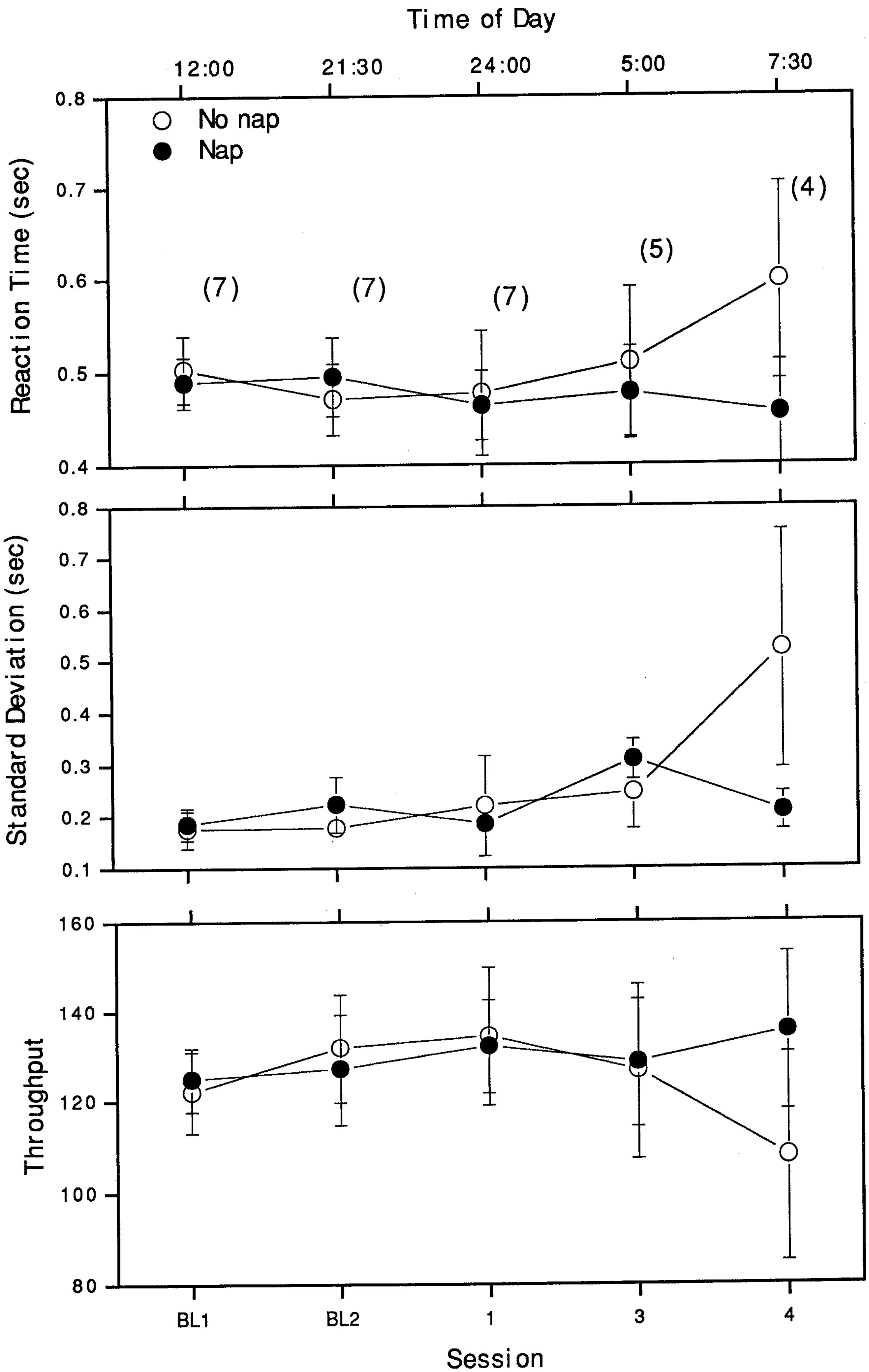
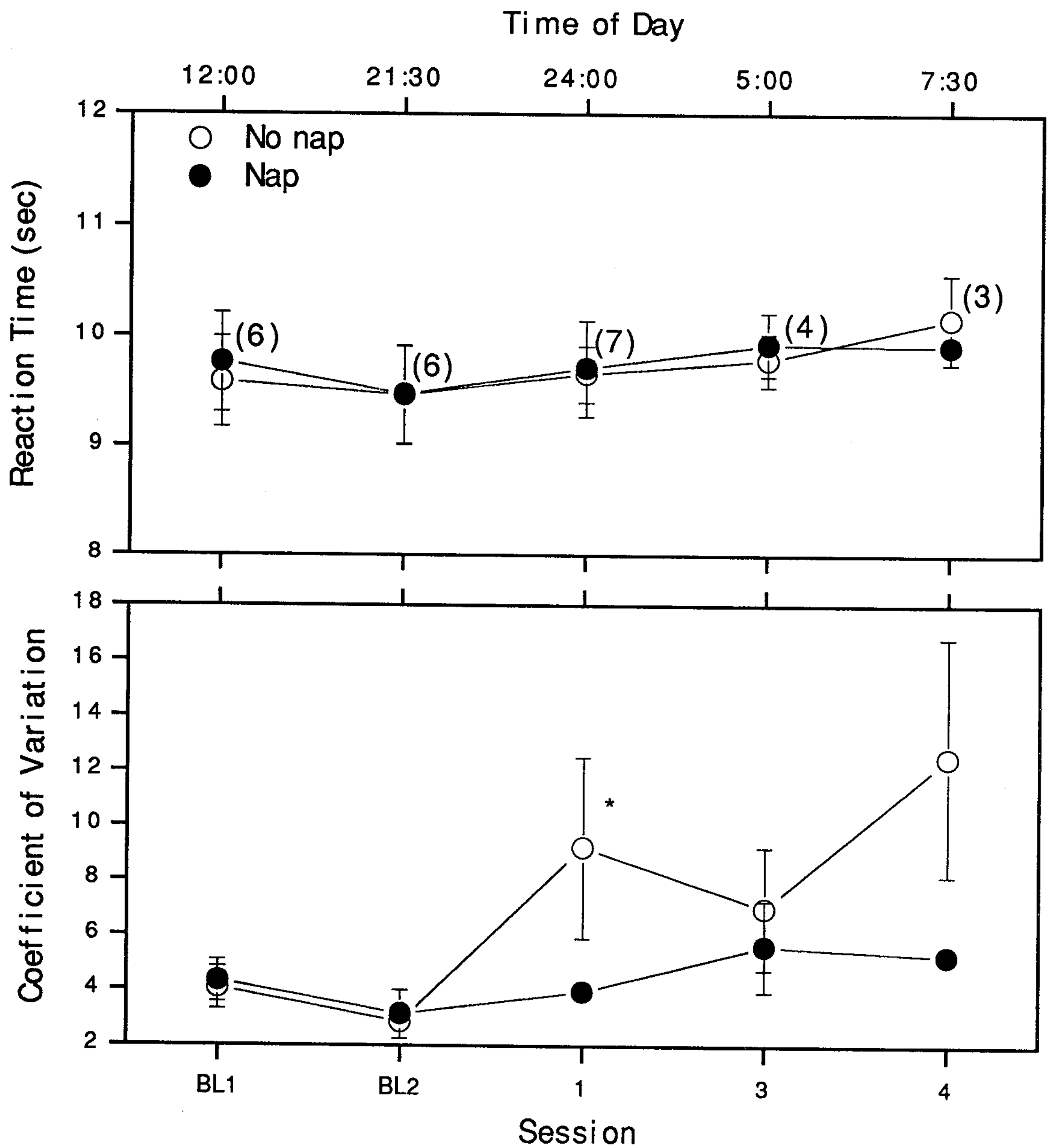


Figure D6. PAB: Time Wall



APPENDIX E

Figures E1-E7. Driving EEG results for individual subjects.

Figure E1. Driving EEG - Subject 1

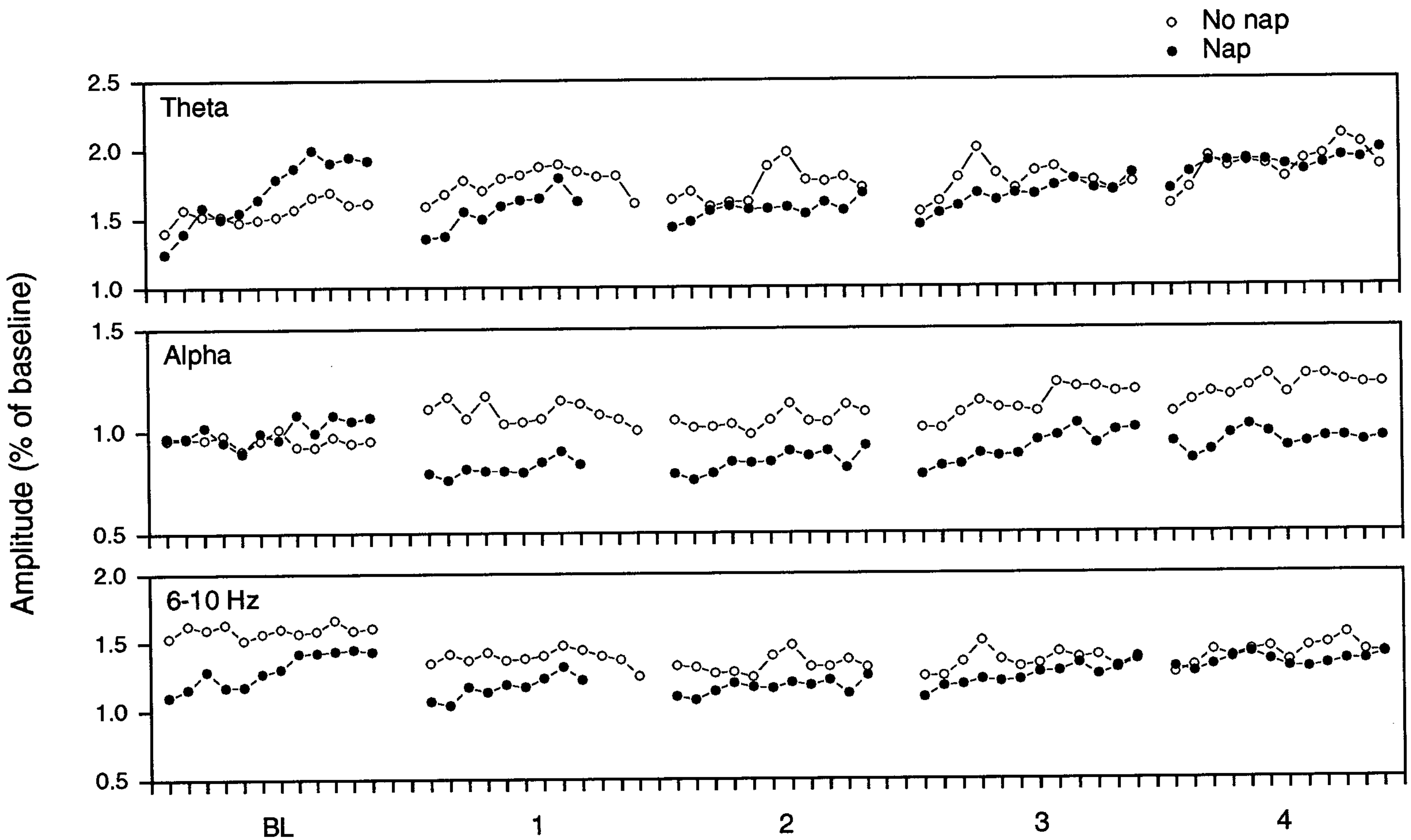


Figure E2. Driving EEG - Subject 2

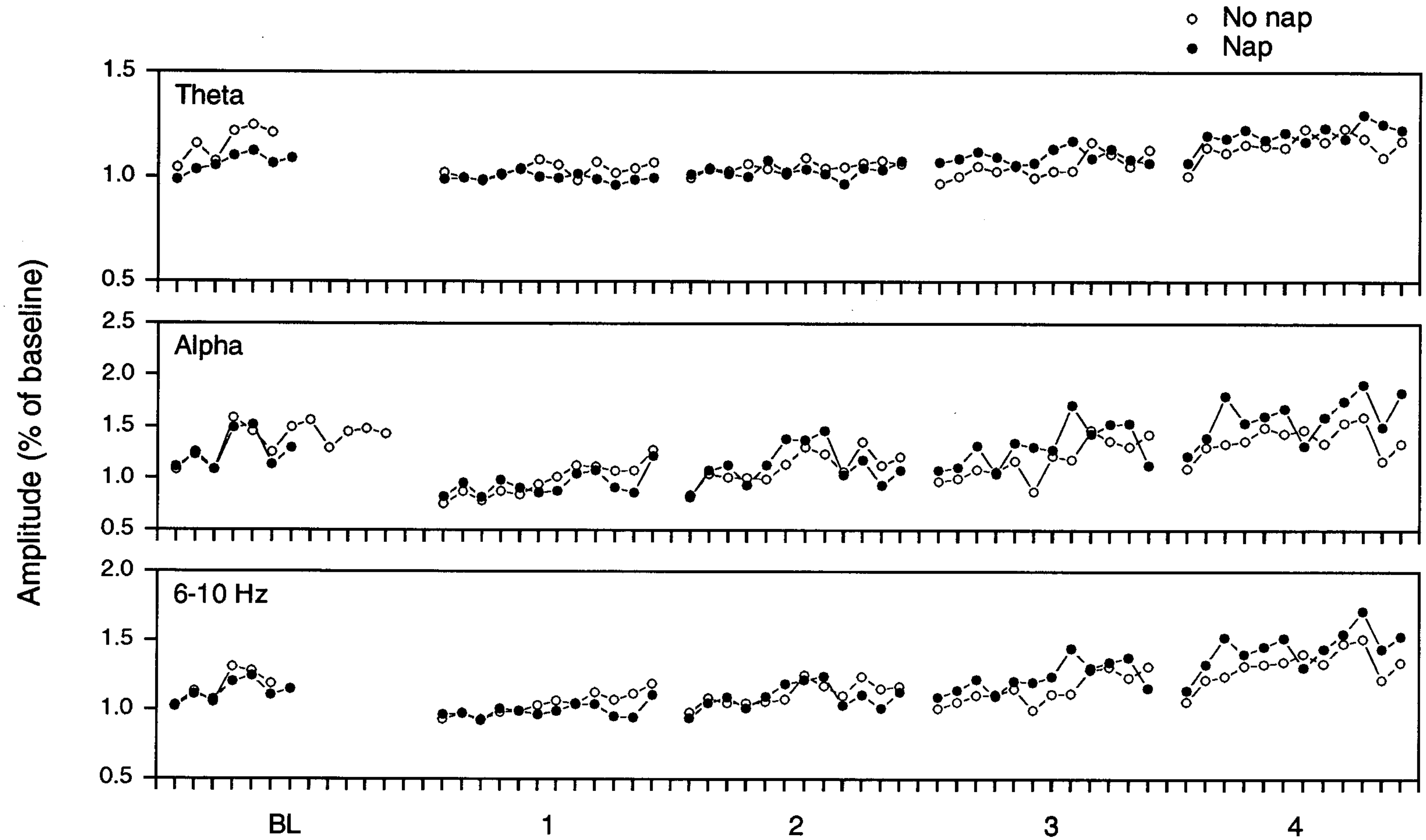


Figure E3. Driving EEG - Subject 3

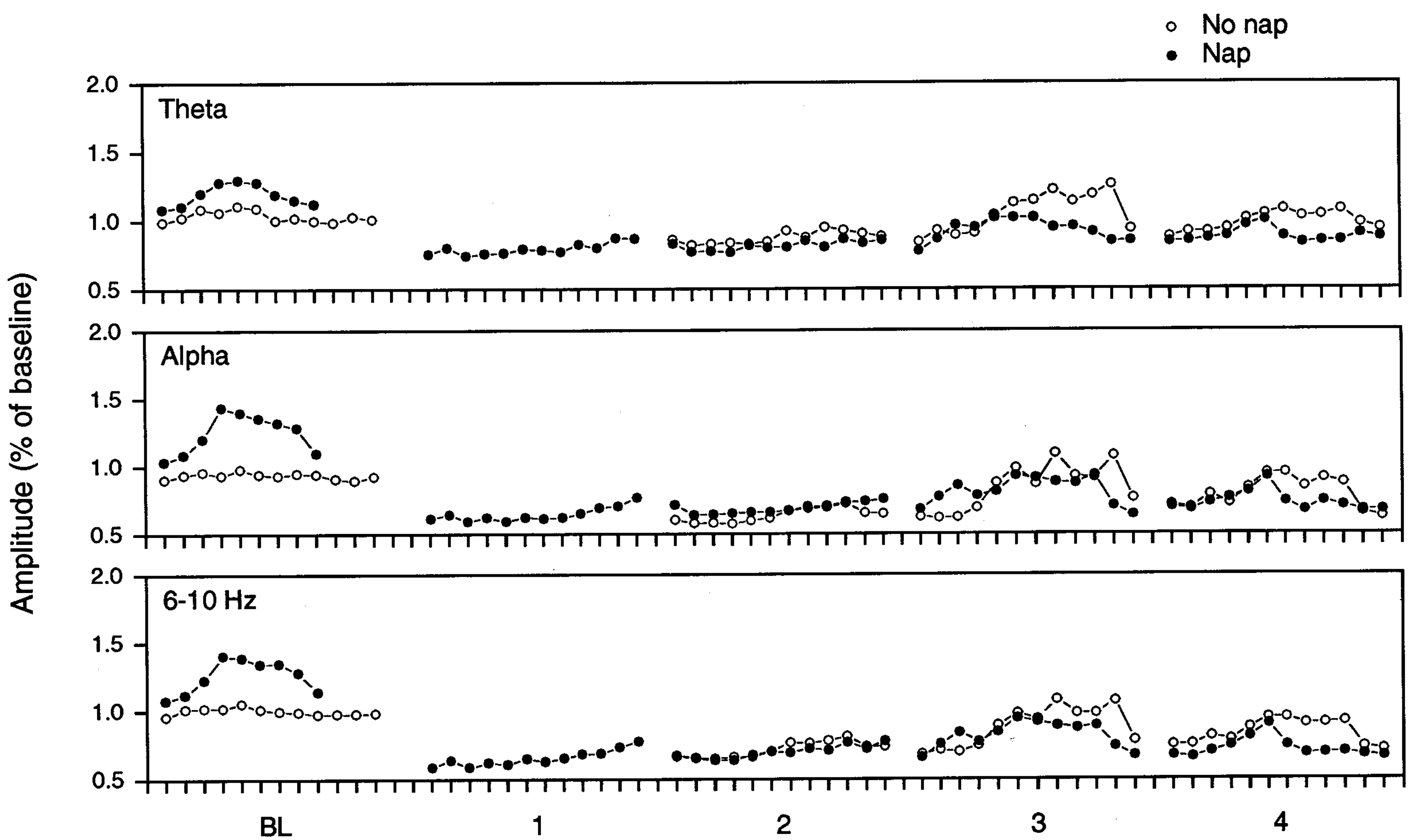


Figure E4. Driving EEG - Subject 4

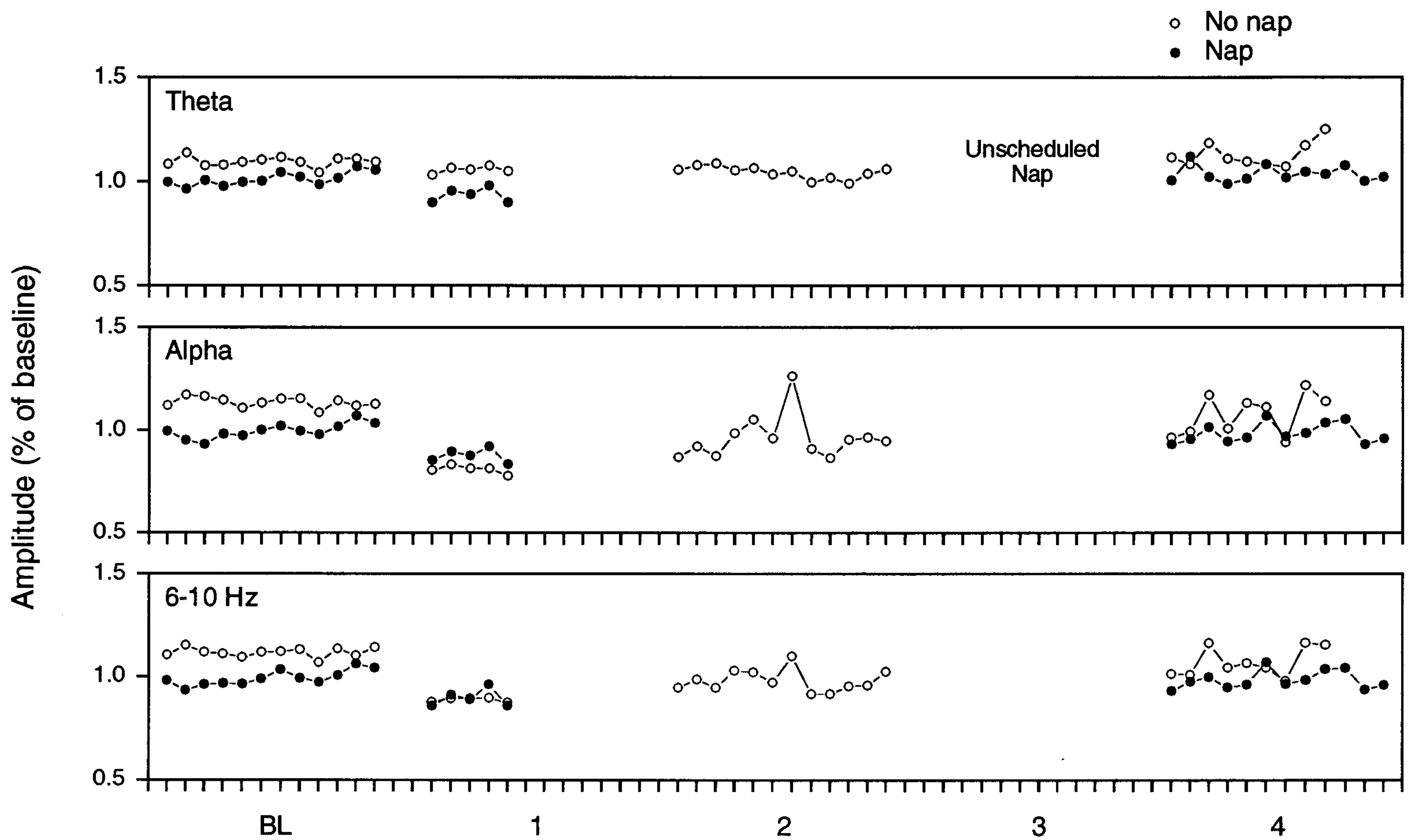


Figure E5. Driving EEG - Subject 5

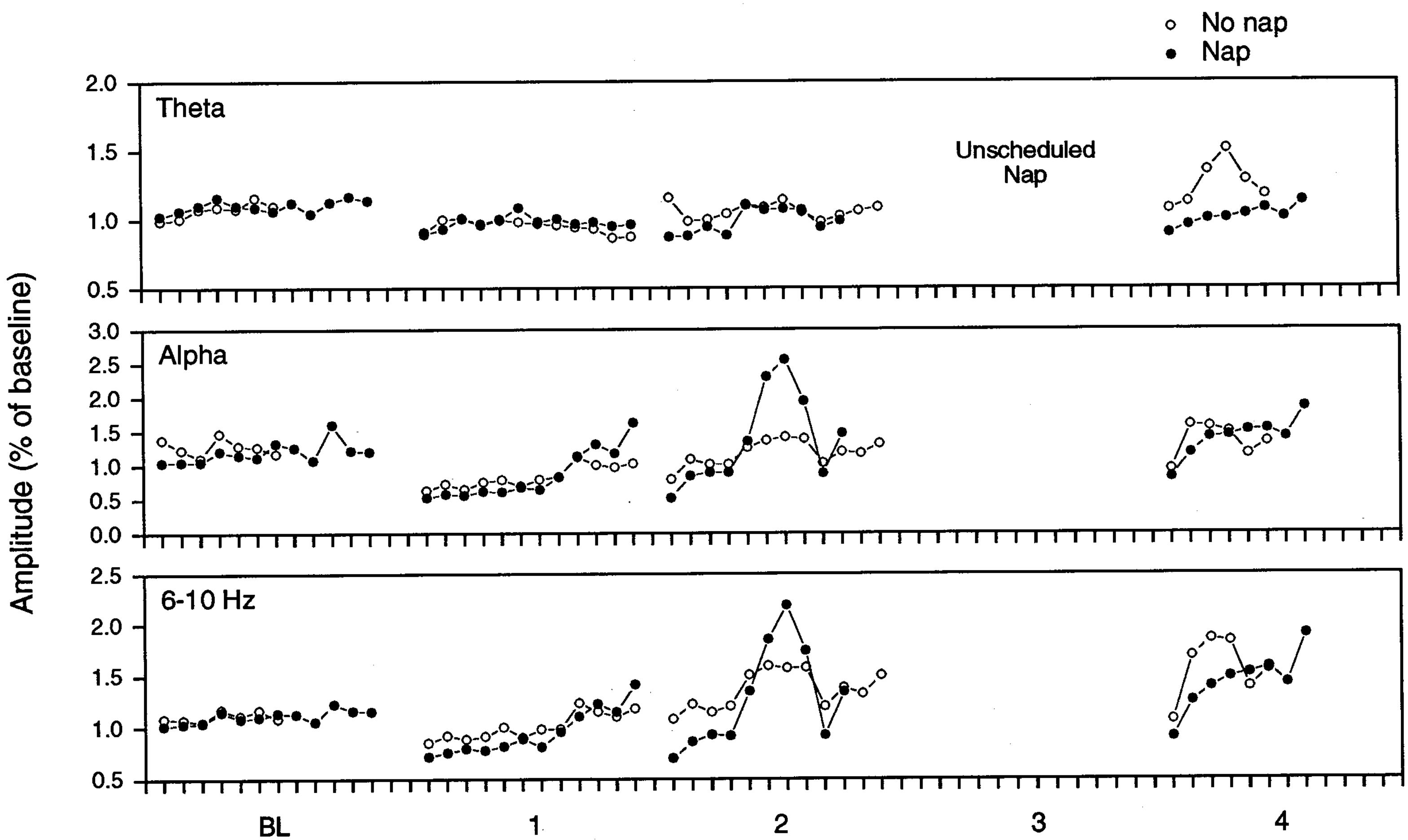


Figure E6. Driving EEG - Subject 7

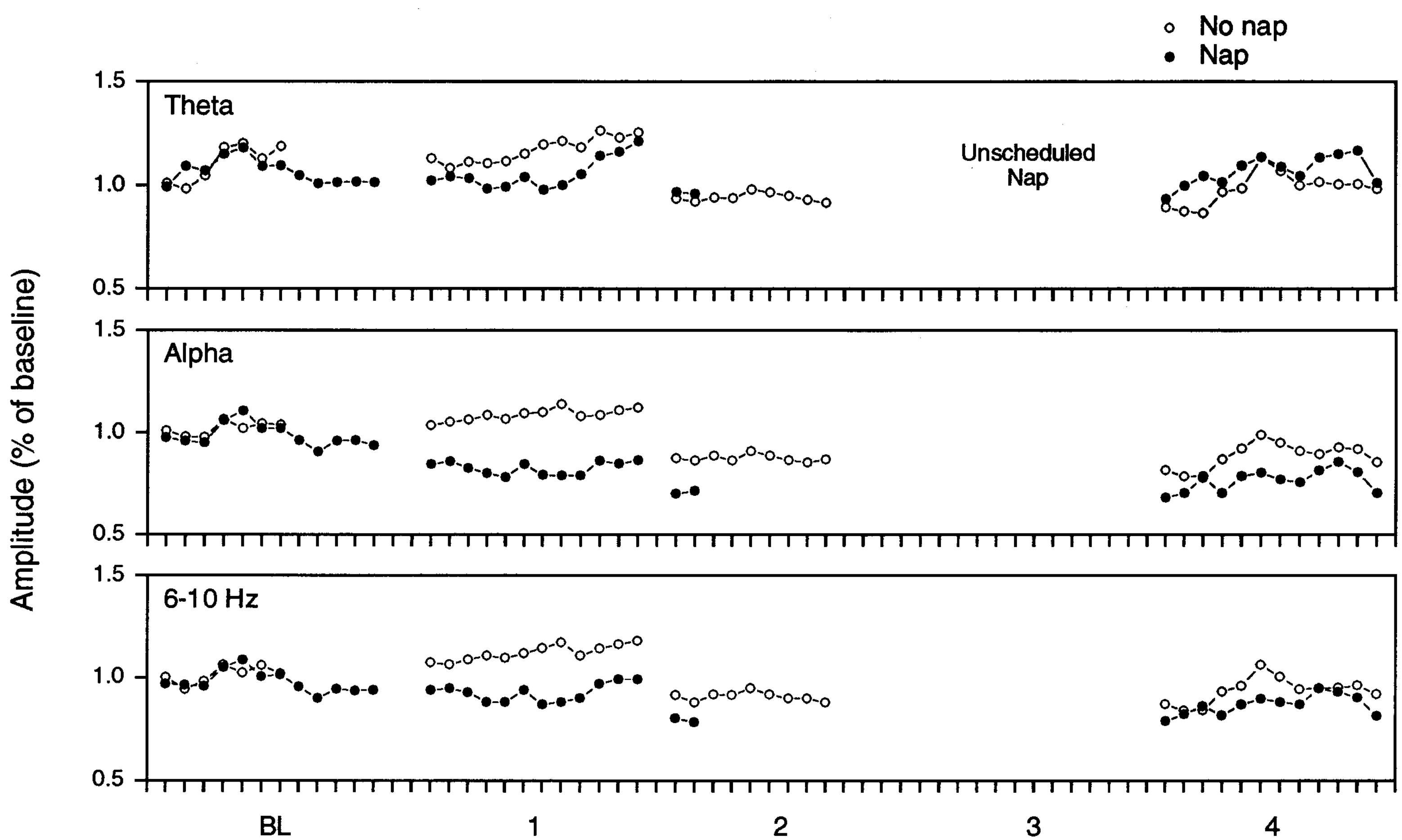
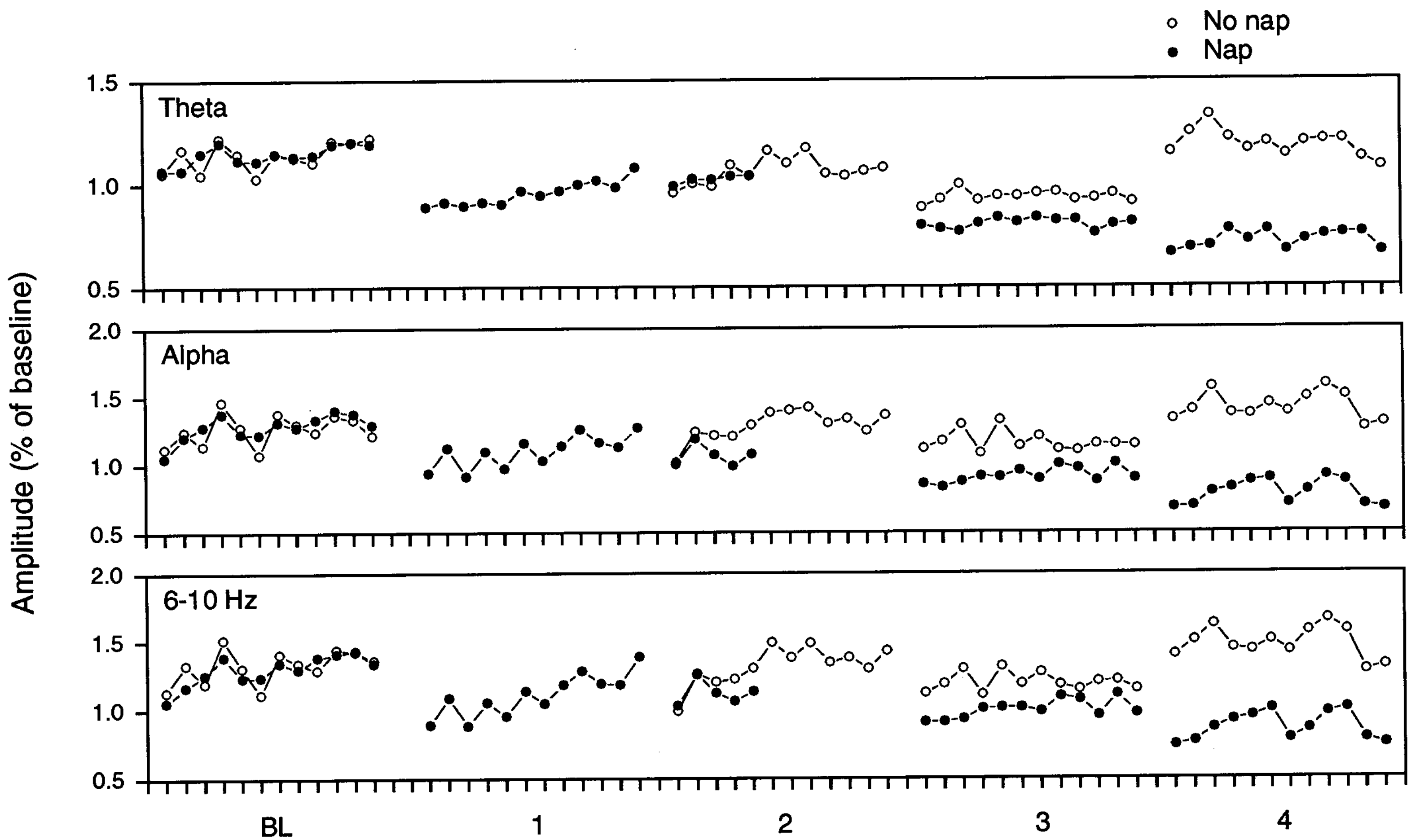


Figure E7. Driving EEG - Subject 9



APPENDIX F

Figures F1-F2. Resting EEG records and corresponding amplitude spectra for Subject 5 at two times of day.

Figure F1a. Driving EEG record (Subject 2)

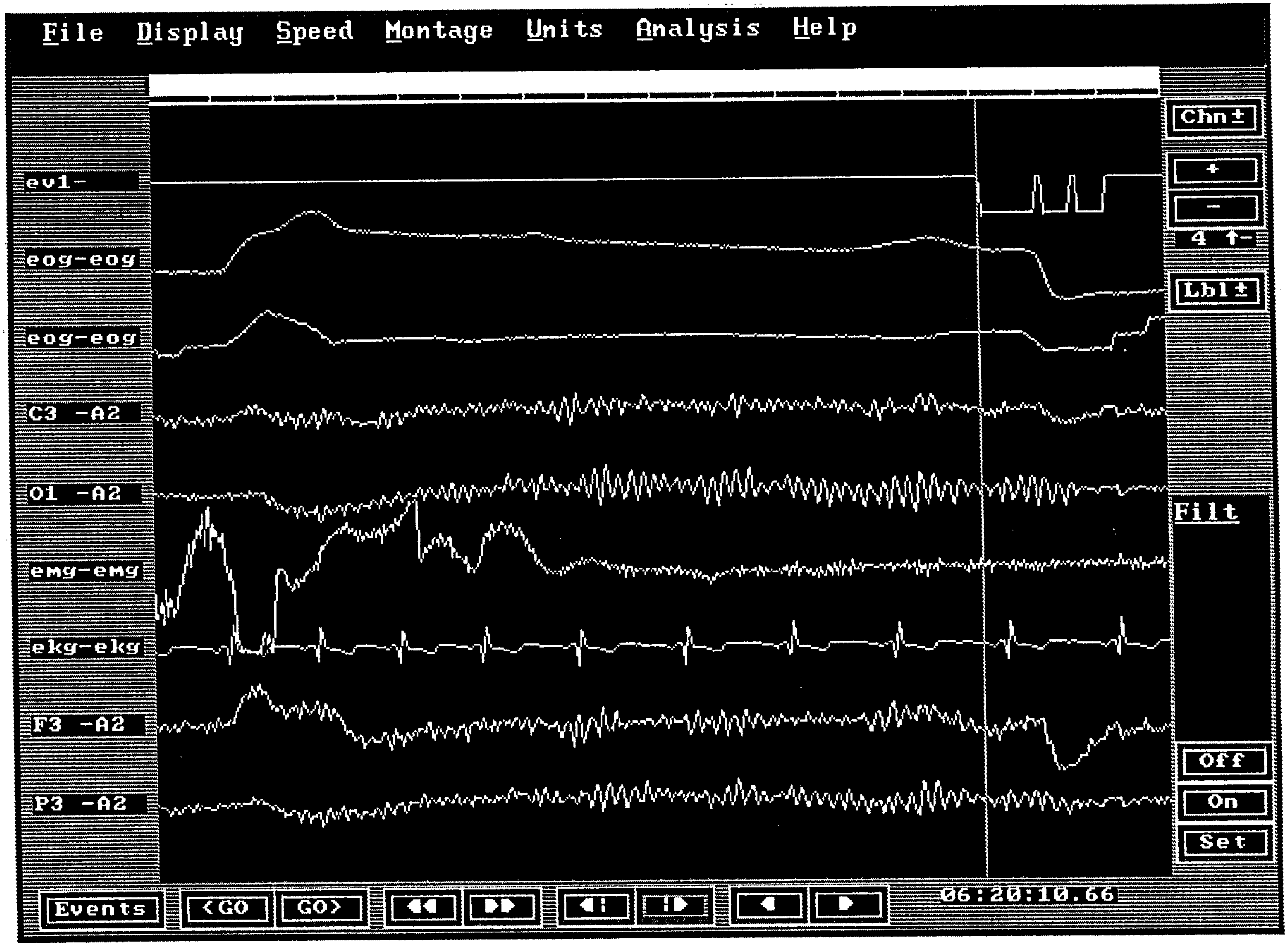


Figure F1b. Amplitude spectra (Subject 2, Off-road Accidents 1 & 2)

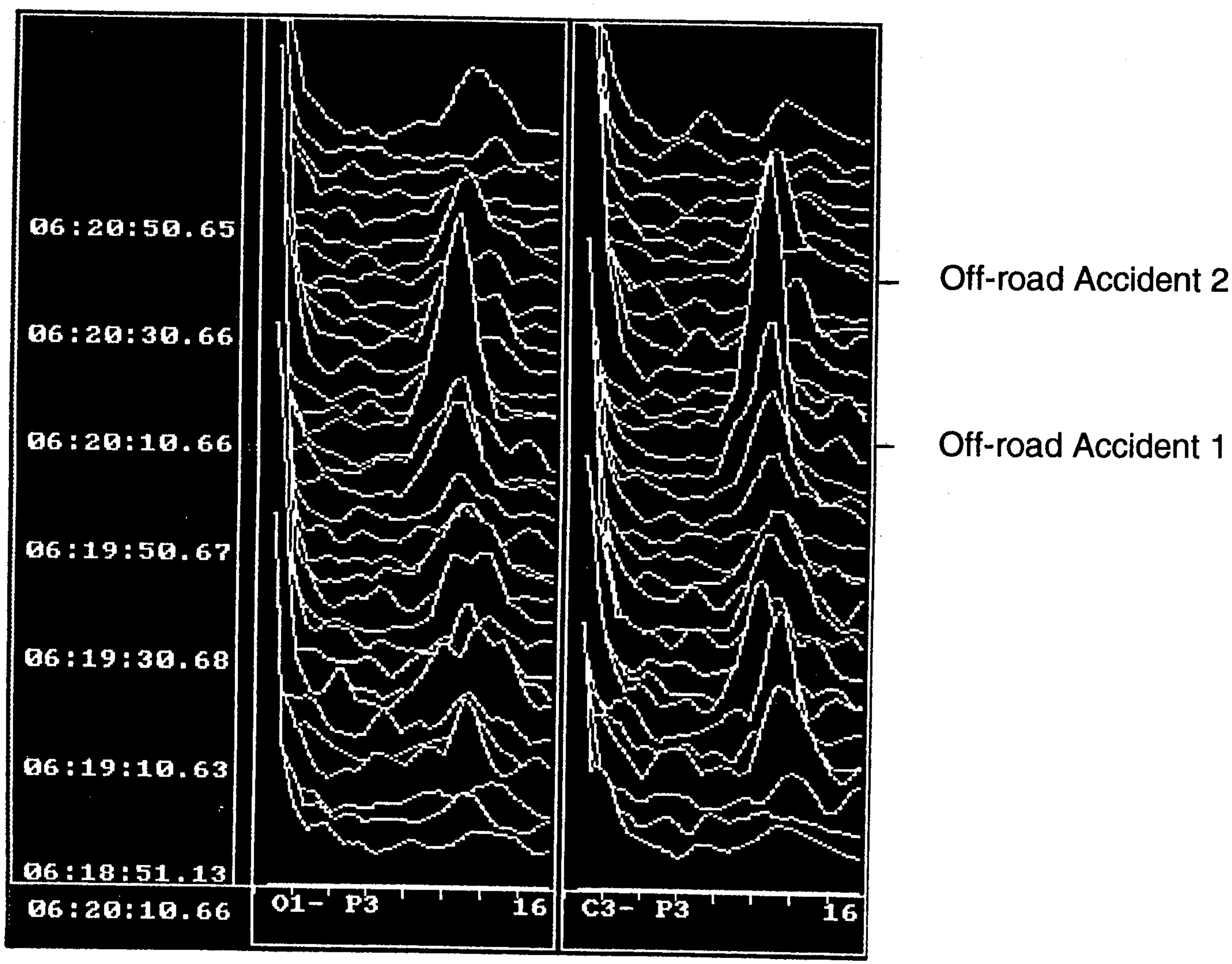


Figure F2a. Driving EEG record (Subject 5, Off-road Accident 2)

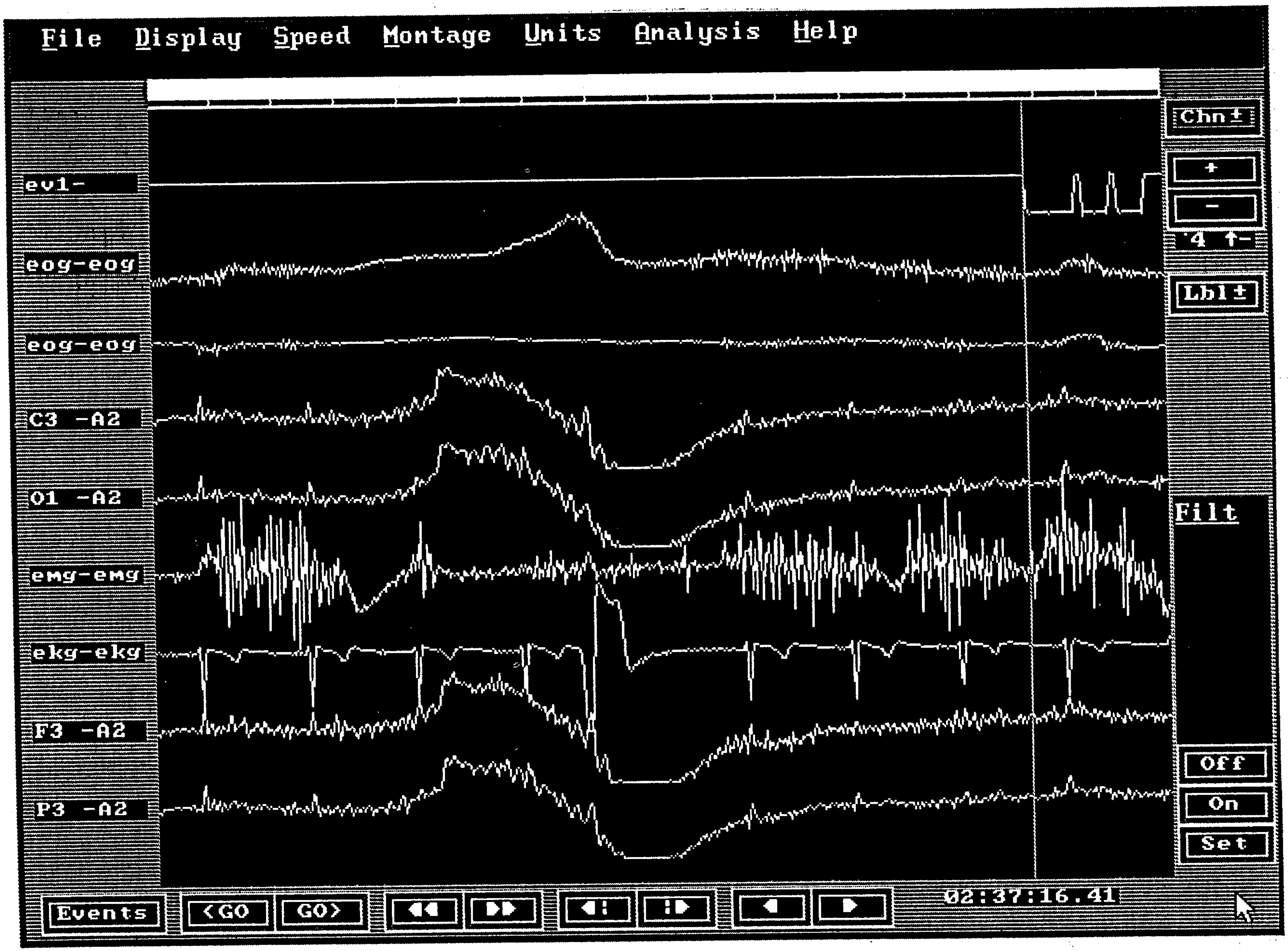


Figure F2b. Driving EEG record (Subject 5, Off-road Accident 3)

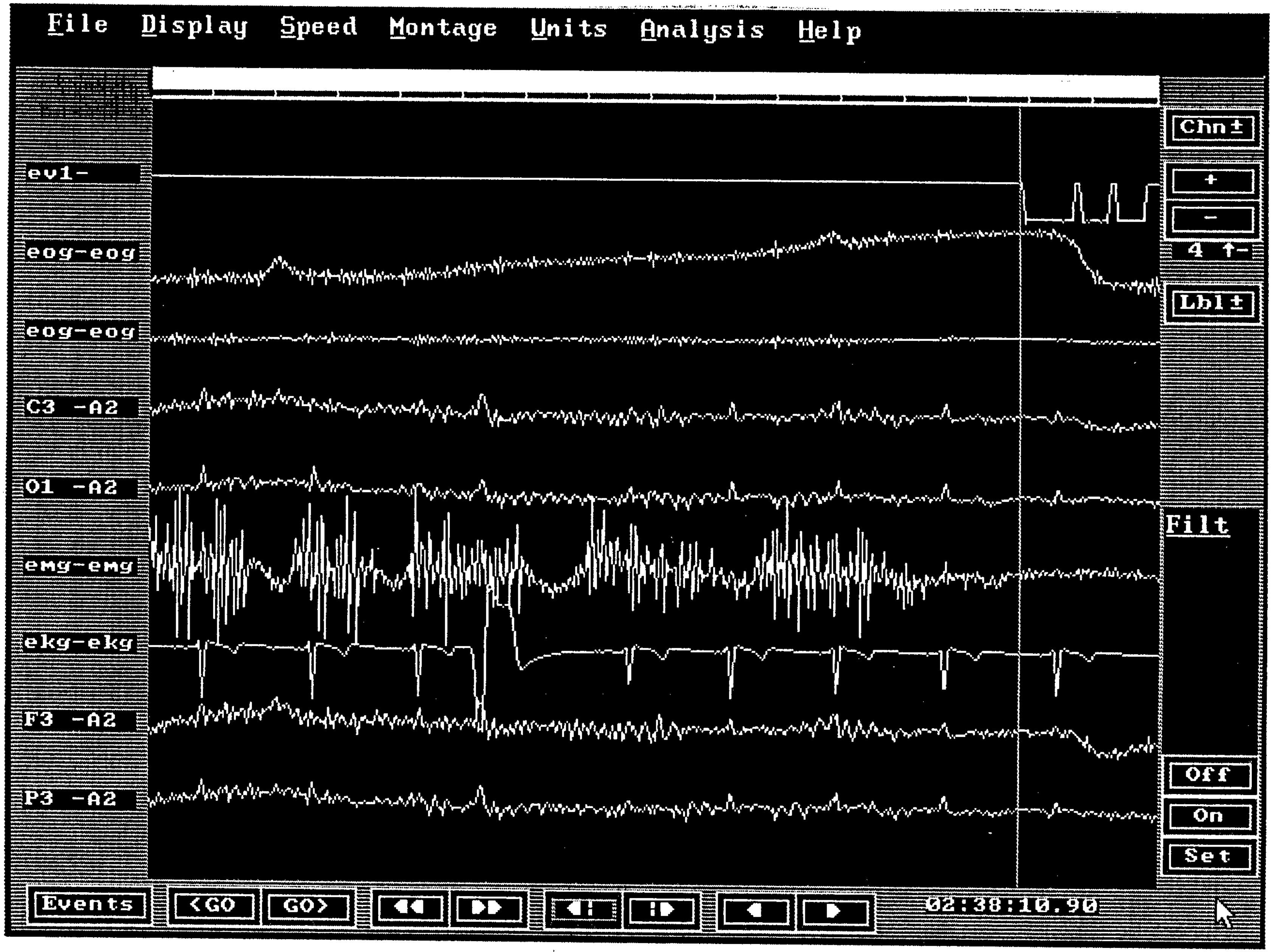
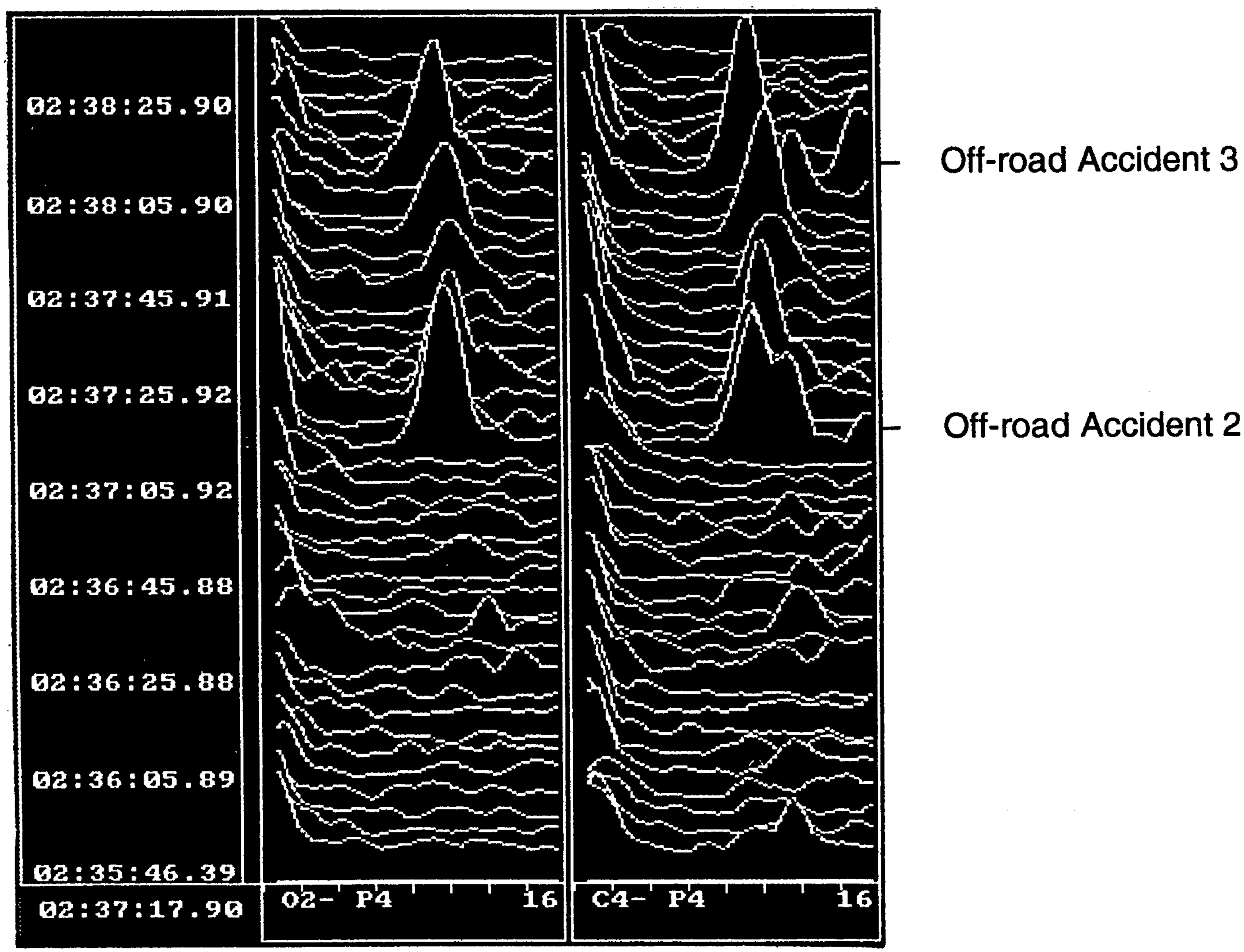


Figure F2c. Amplitude spectra (Subject 5, Off-road Accidents 2 & 3)



APPENDIX G

Figures G1-G2.	Resting EEG records and corresponding amplitude spectra for Subject 5 at two times of day.

Figure G1a. Resting EEG record (Subject 5, 21:30)

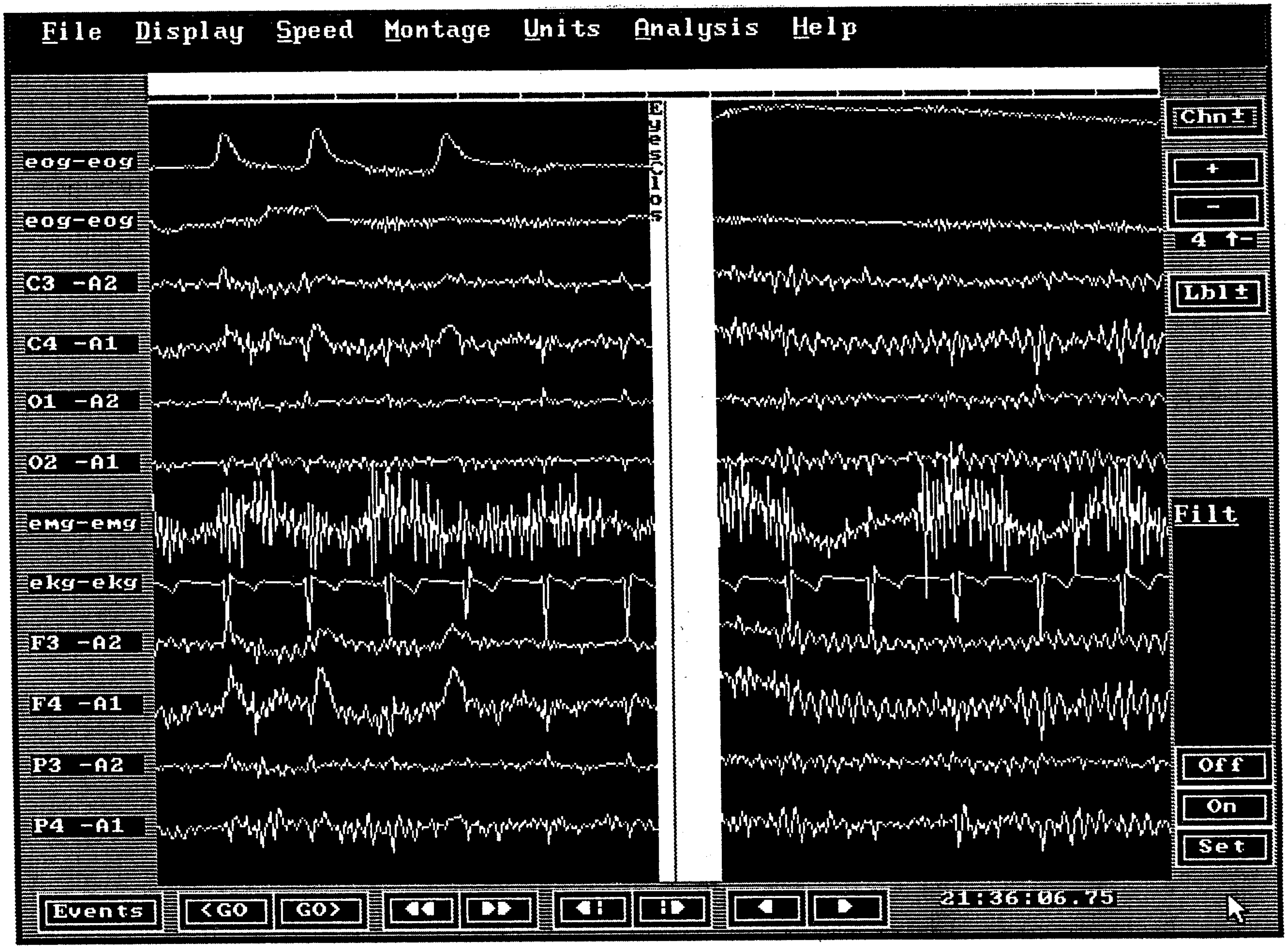


Figure G1b. Amplitude spectra (Subject 5, 21:30)

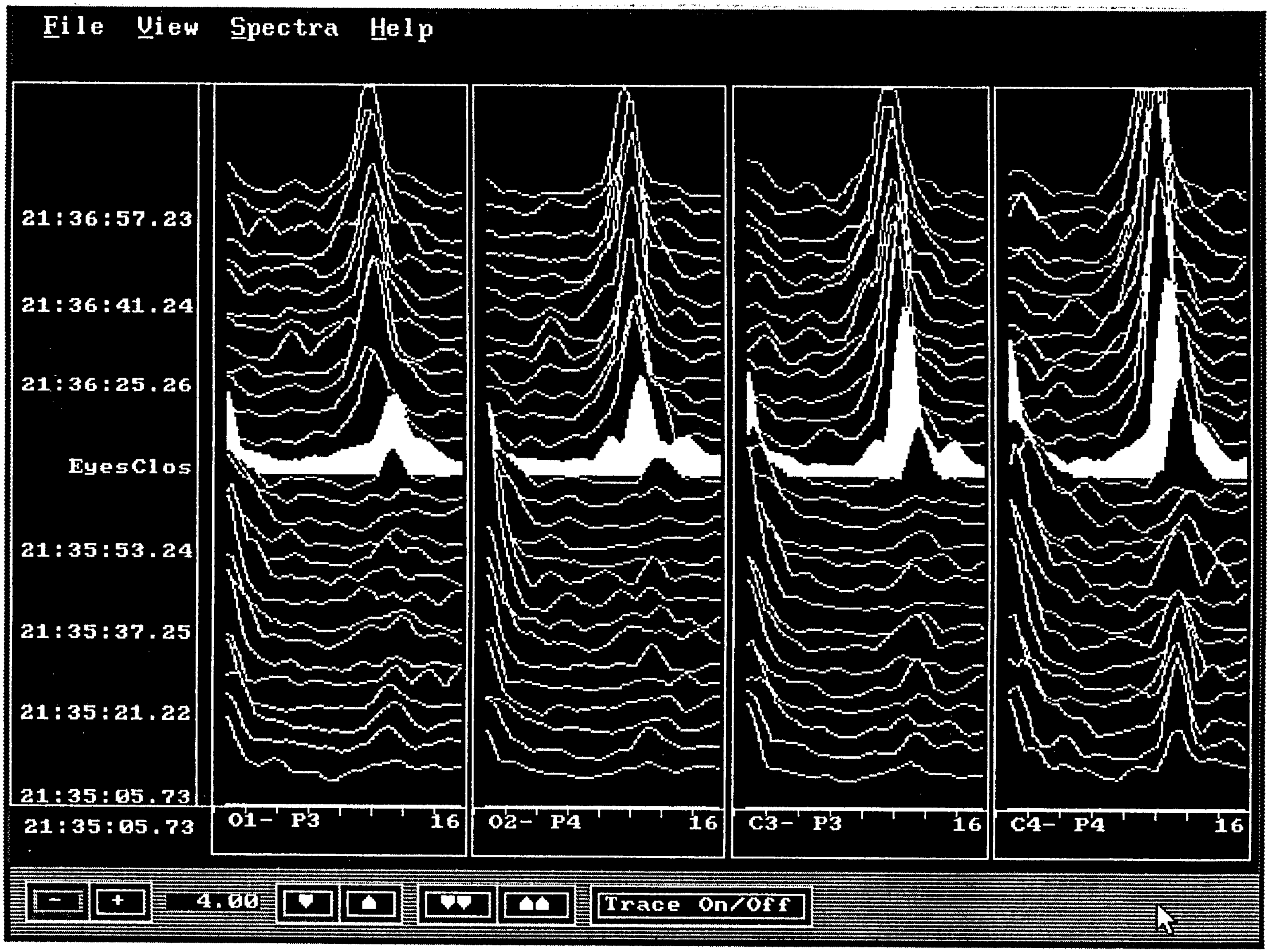


Figure G2a. Resting EEG record (Subject 5, 02:30)

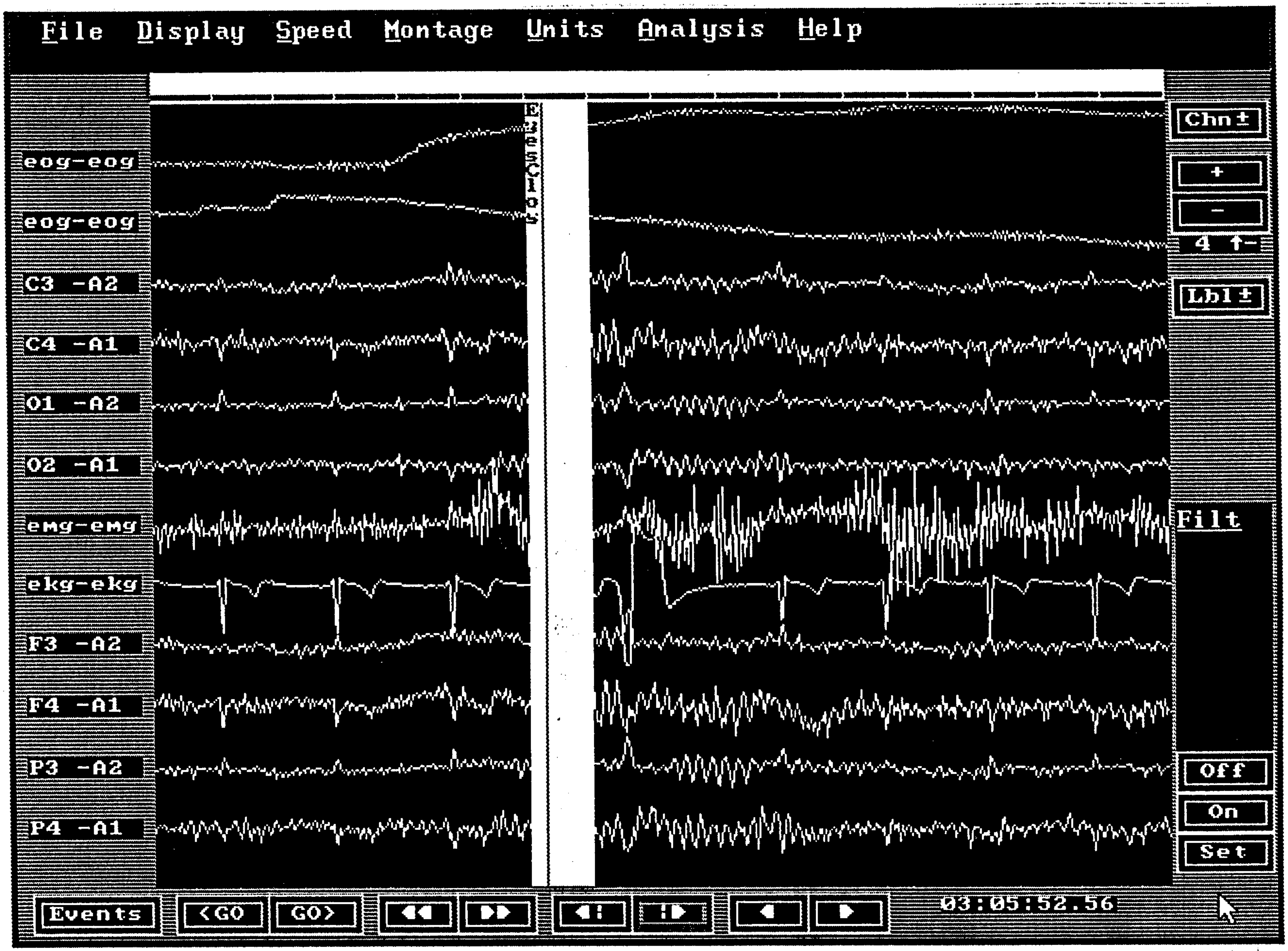
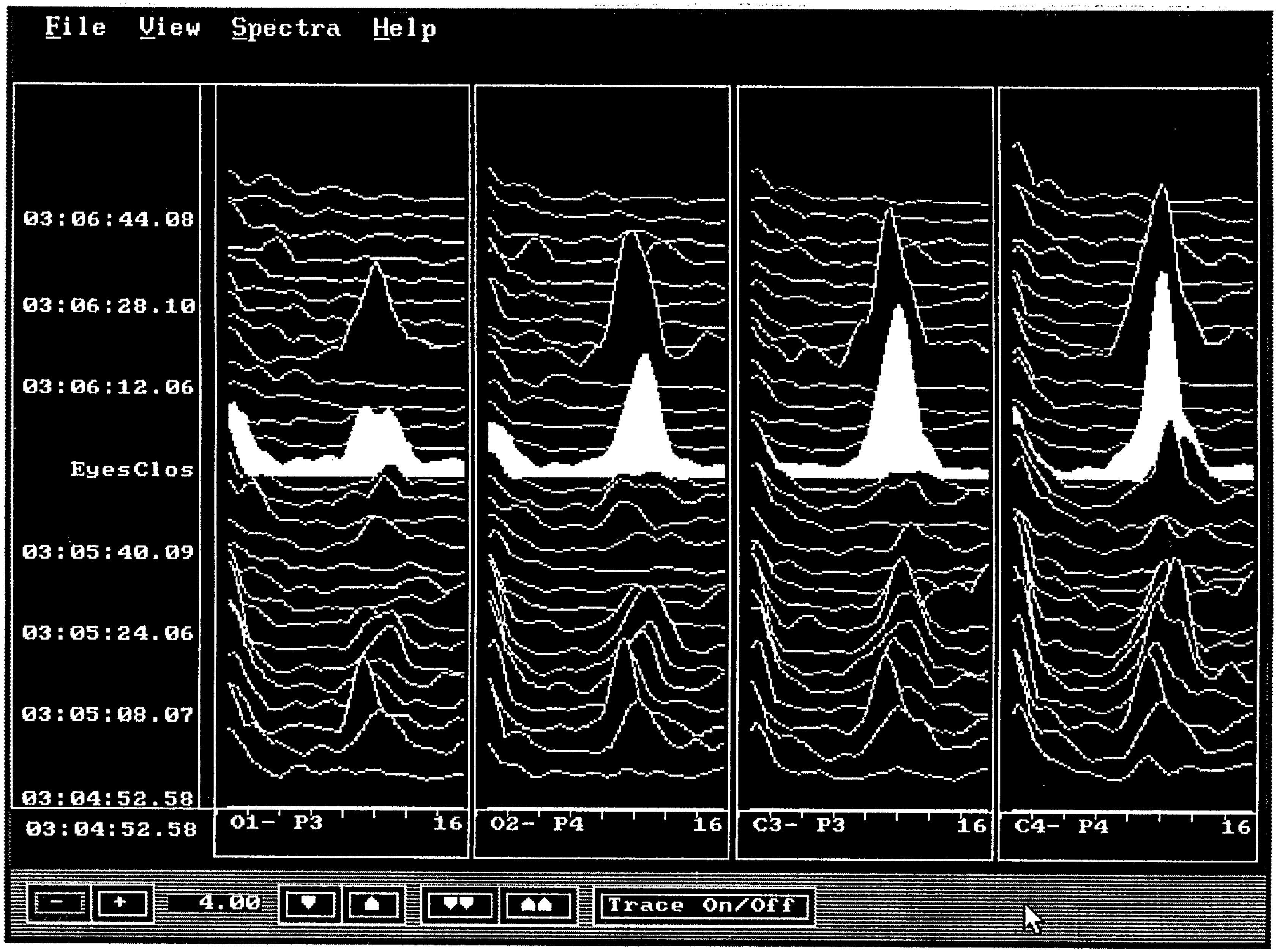


Figure G2b. Amplitude spectra (Subject 5, 02:30)



APPENDIX H

Figures H1-H7. EOG blink measures for individual subjects..

Figure H1. EOG Blink Measures - Sb01

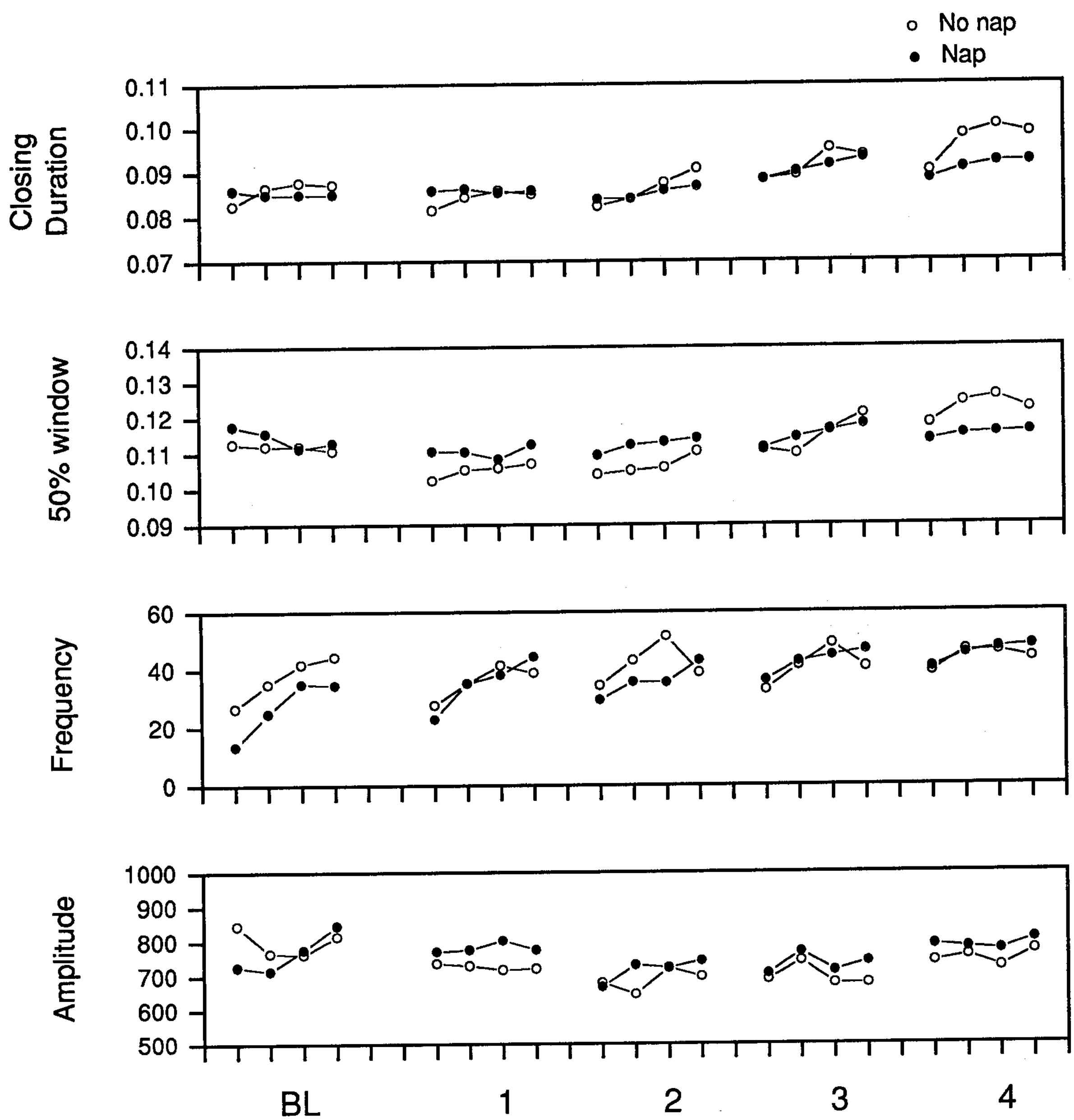


Figure H2. EOG Blink Measures - Sb02

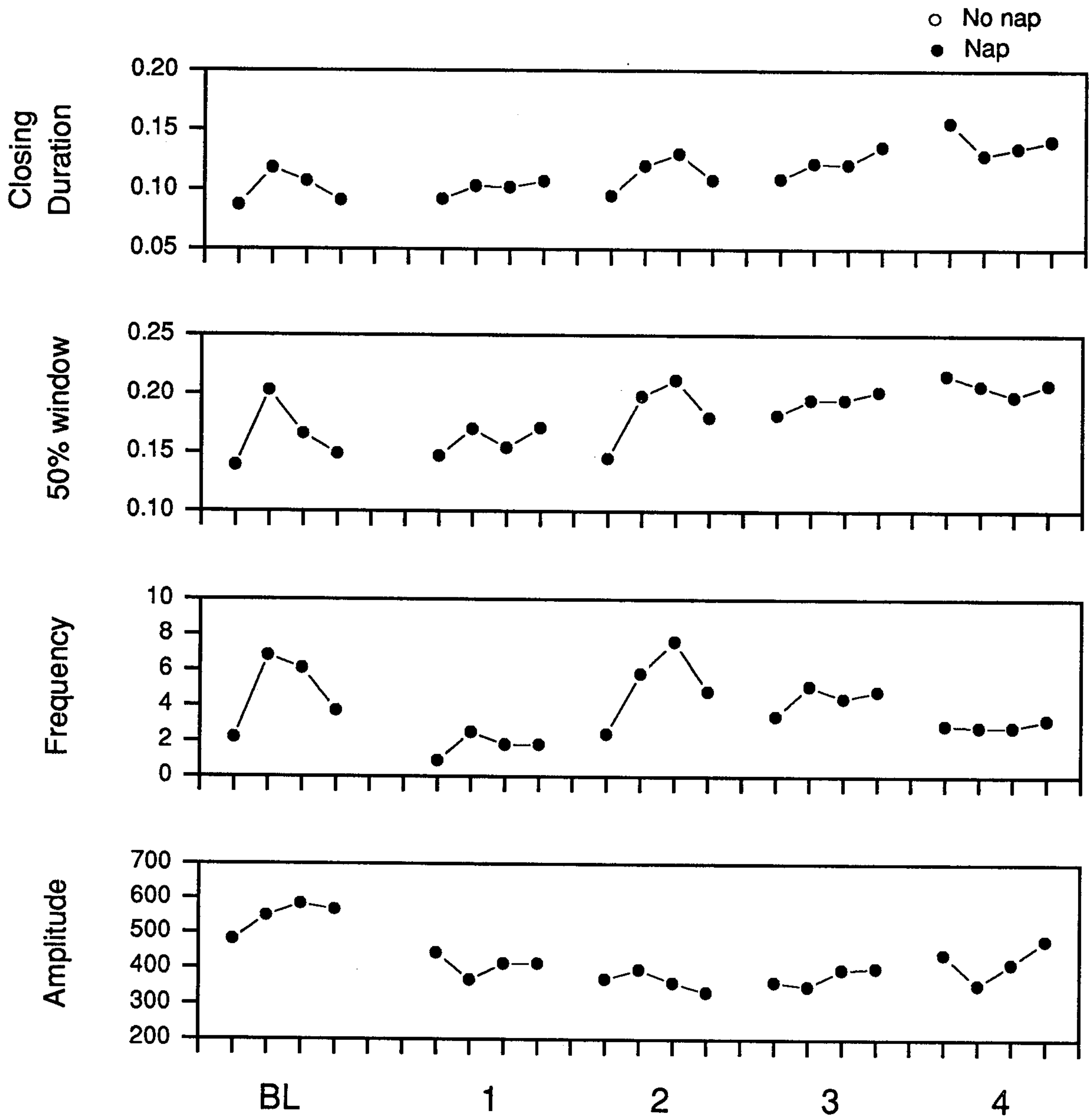


Figure H3. EOG Blink Measures - Sb03

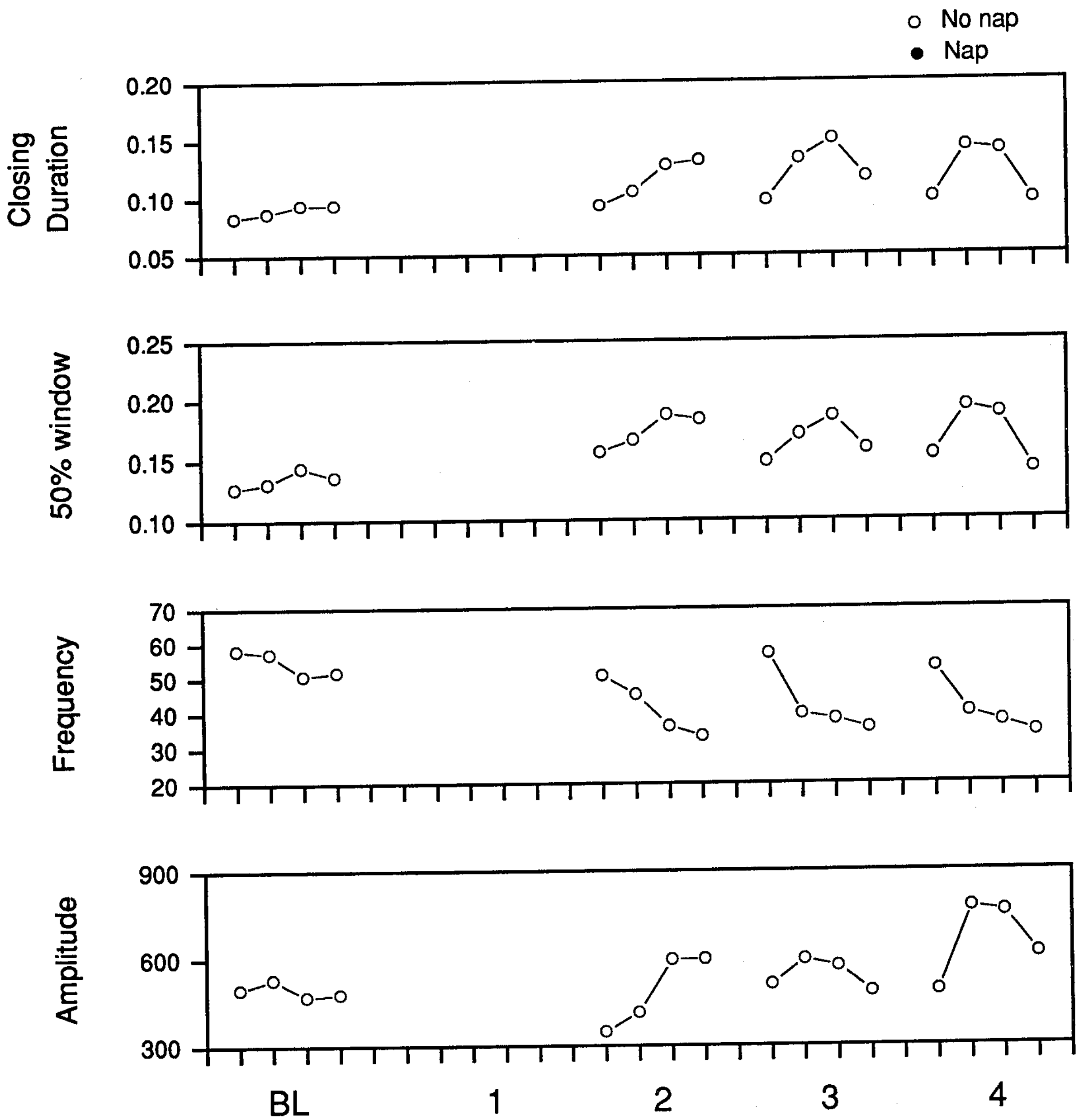


Figure H4. EOG Blink Measures - Sb04

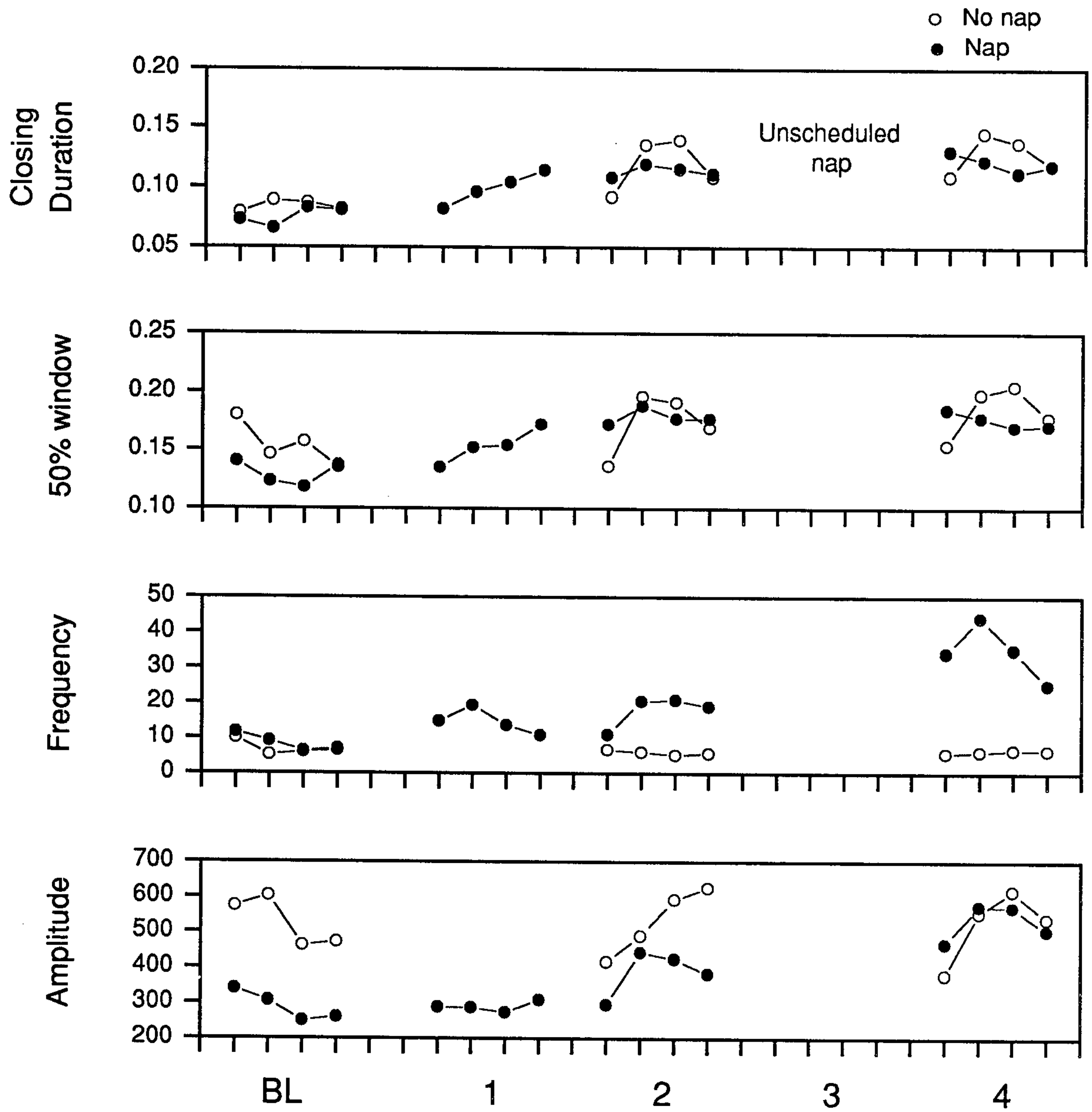


Figure H5. EOG Blink Measures - Sb05

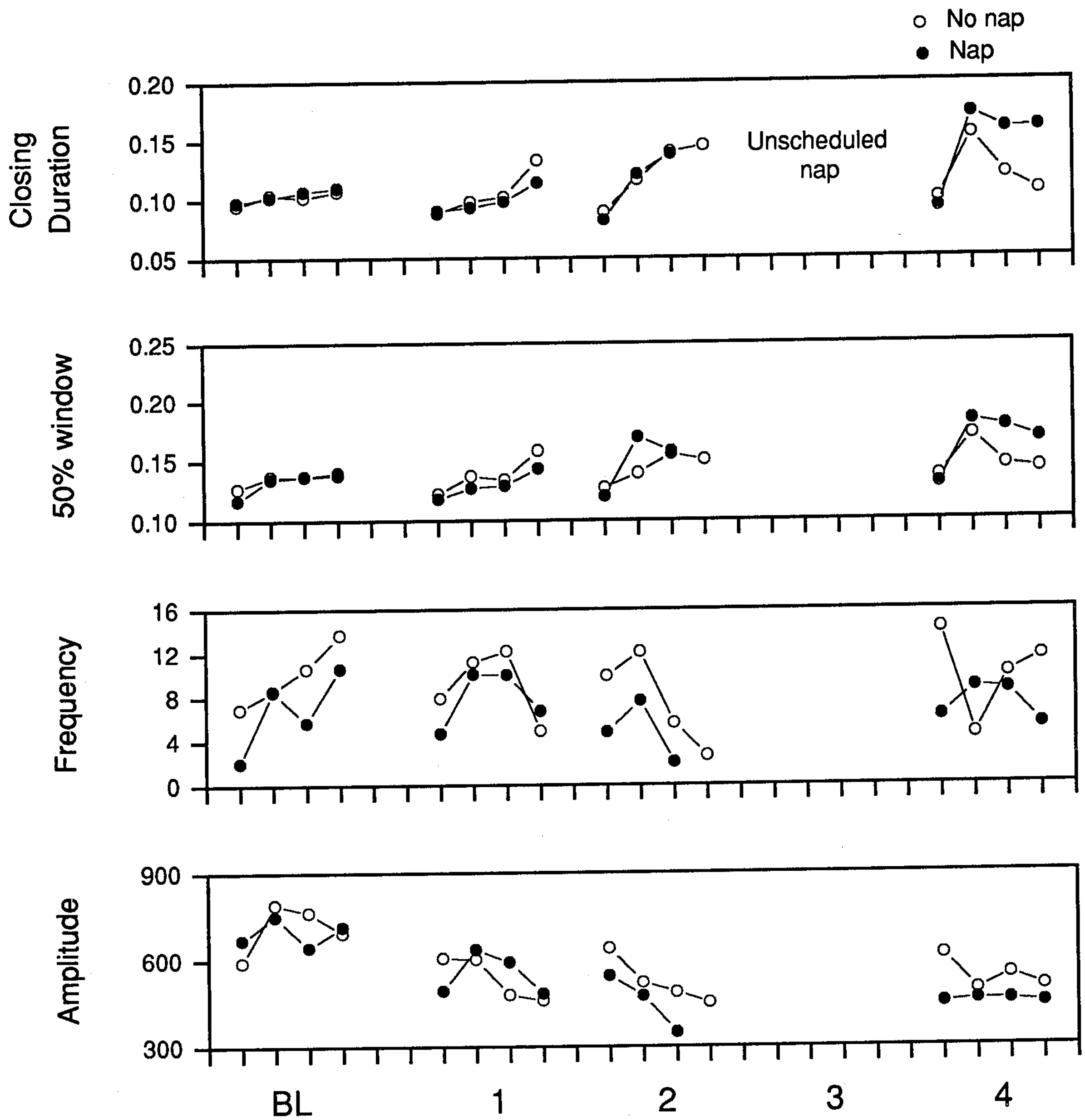


Figure H6. EOG Blink Measures - Sb07

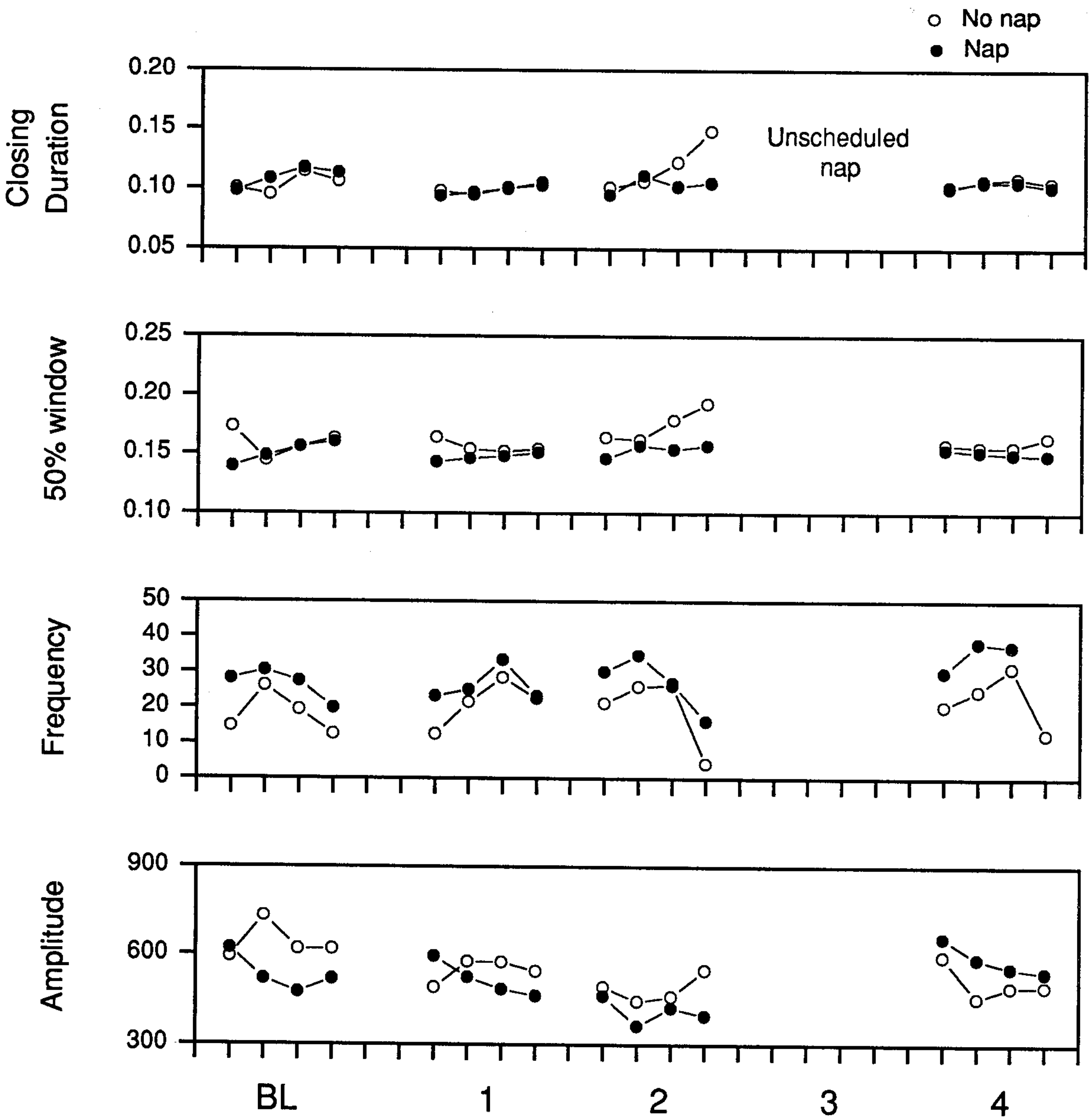
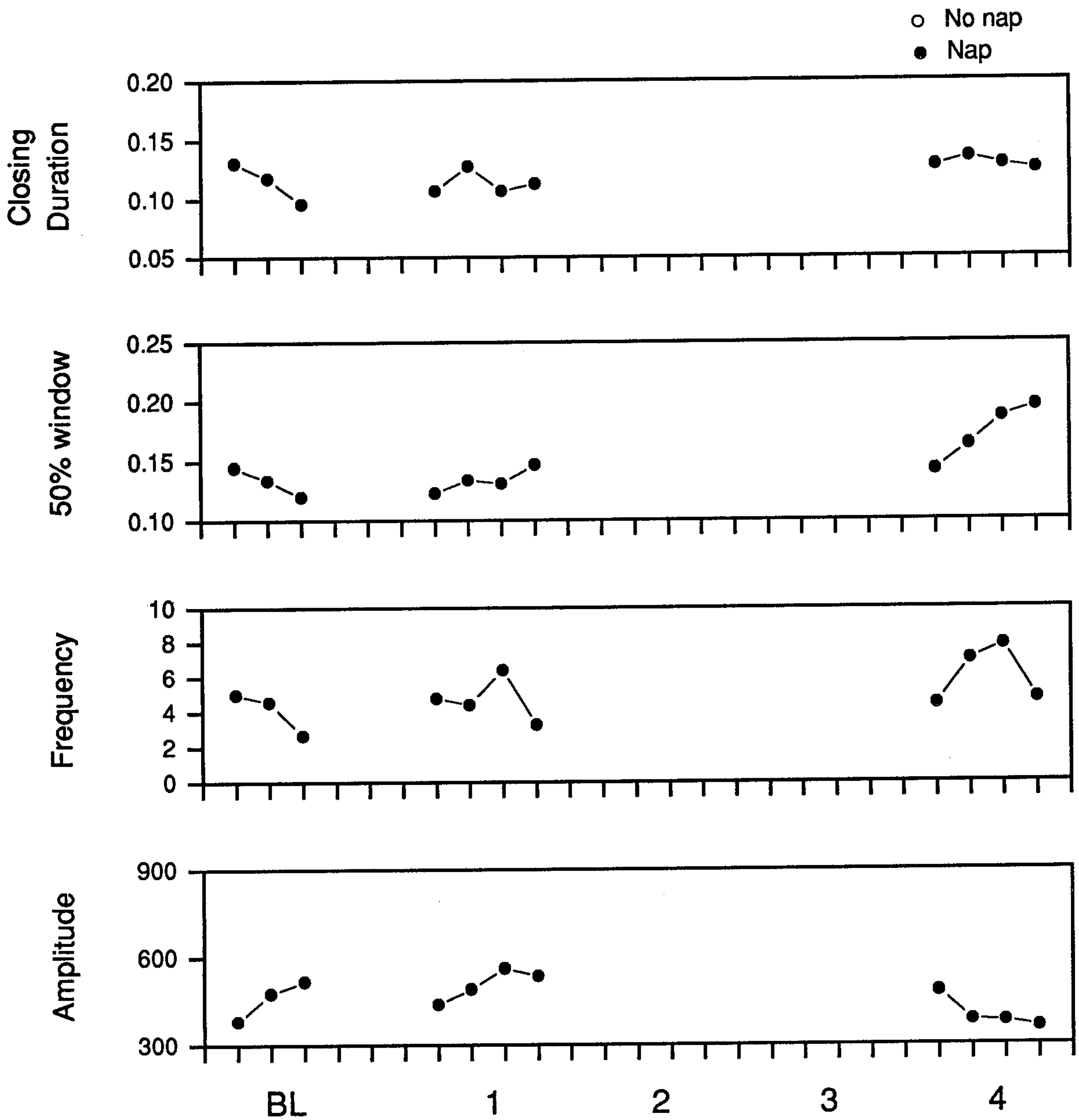


Figure H7. EOG Blink Measures - Sb09



APPENDIX I

Figures I1-I8. Eye tracker blink measures for individual subjects.

Figure I1. Eye Tracker Blink Measures - Sb01

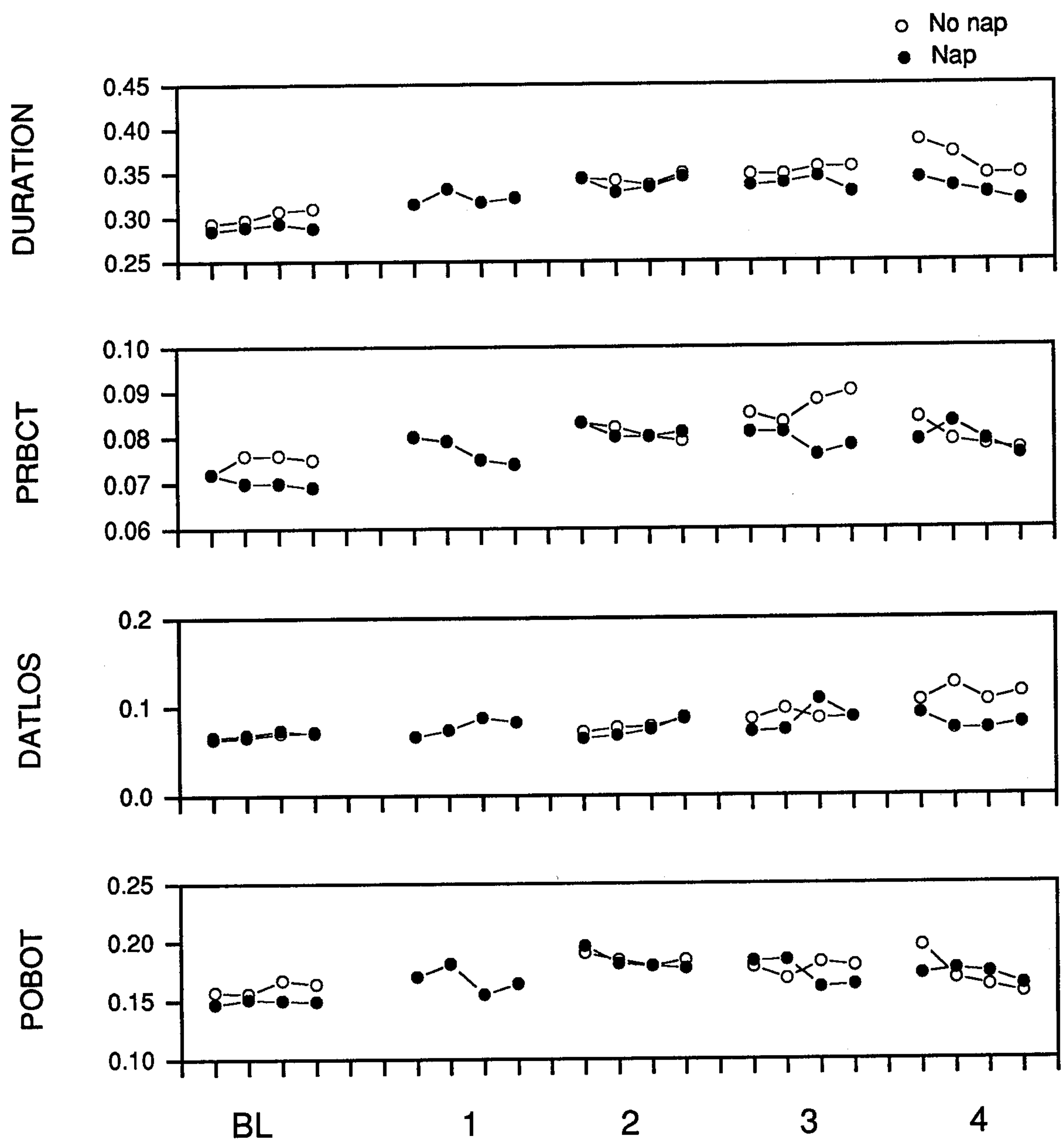


Figure I2. Eye Tracker Blink Measures - Sb02

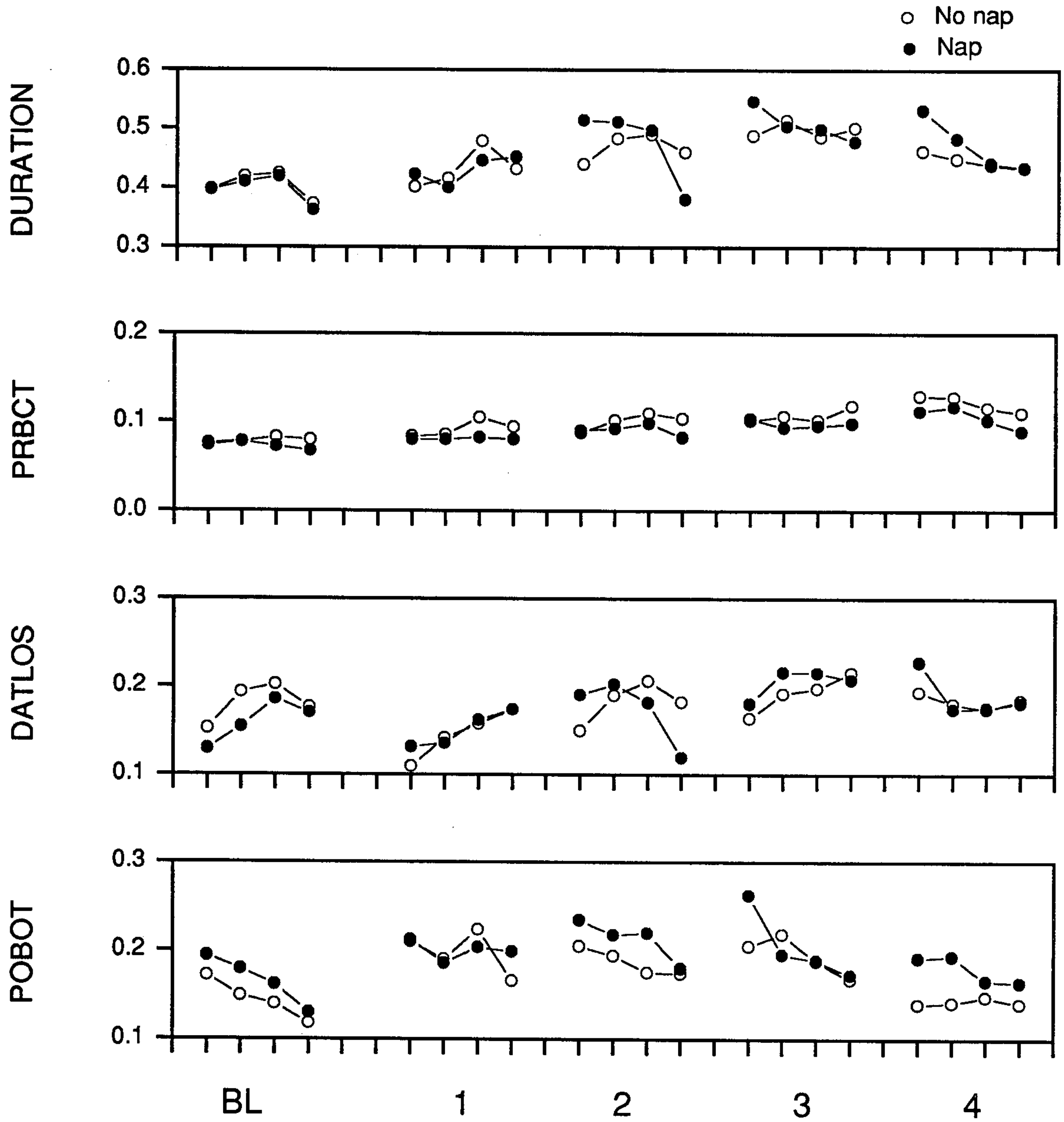


Figure I3. Eye Tracker Blink Measures - Sb03

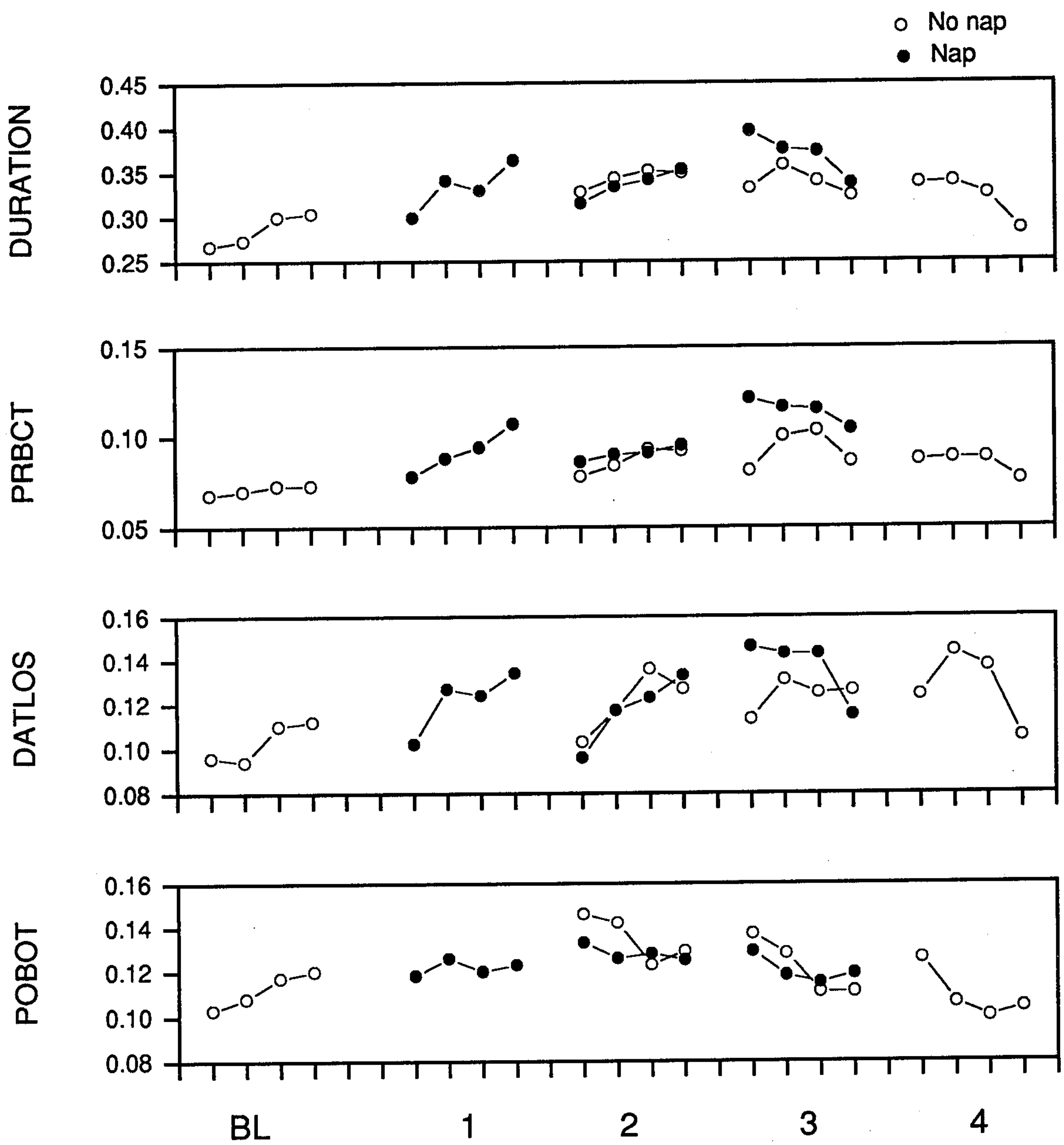


Figure I4. Eye Tracker Blink Measures - Sb05

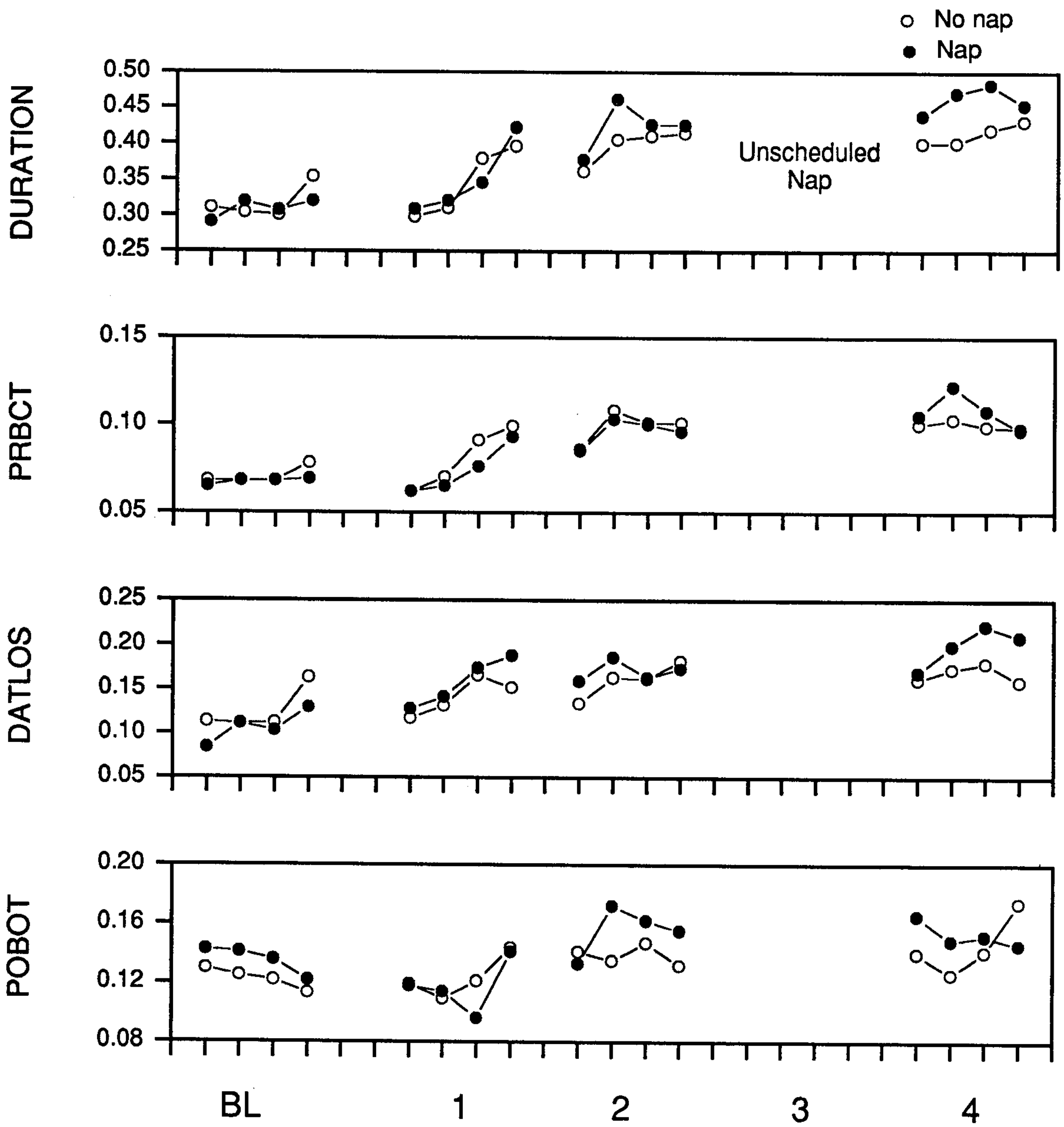


Figure I5. Eye Tracker Blink Measures - Sb06

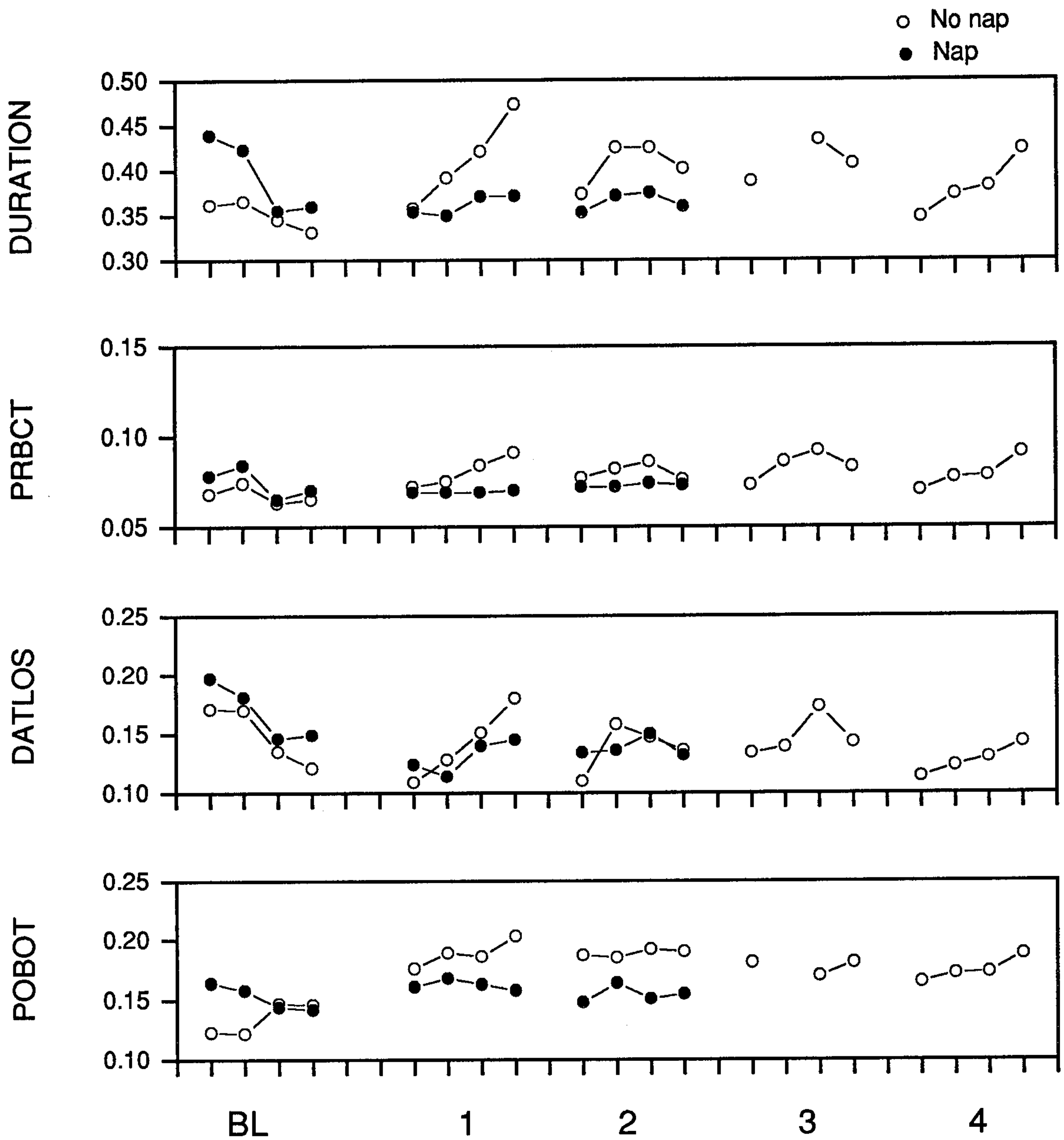


Figure I6. Eye Tracker Blink Measures - Sb07

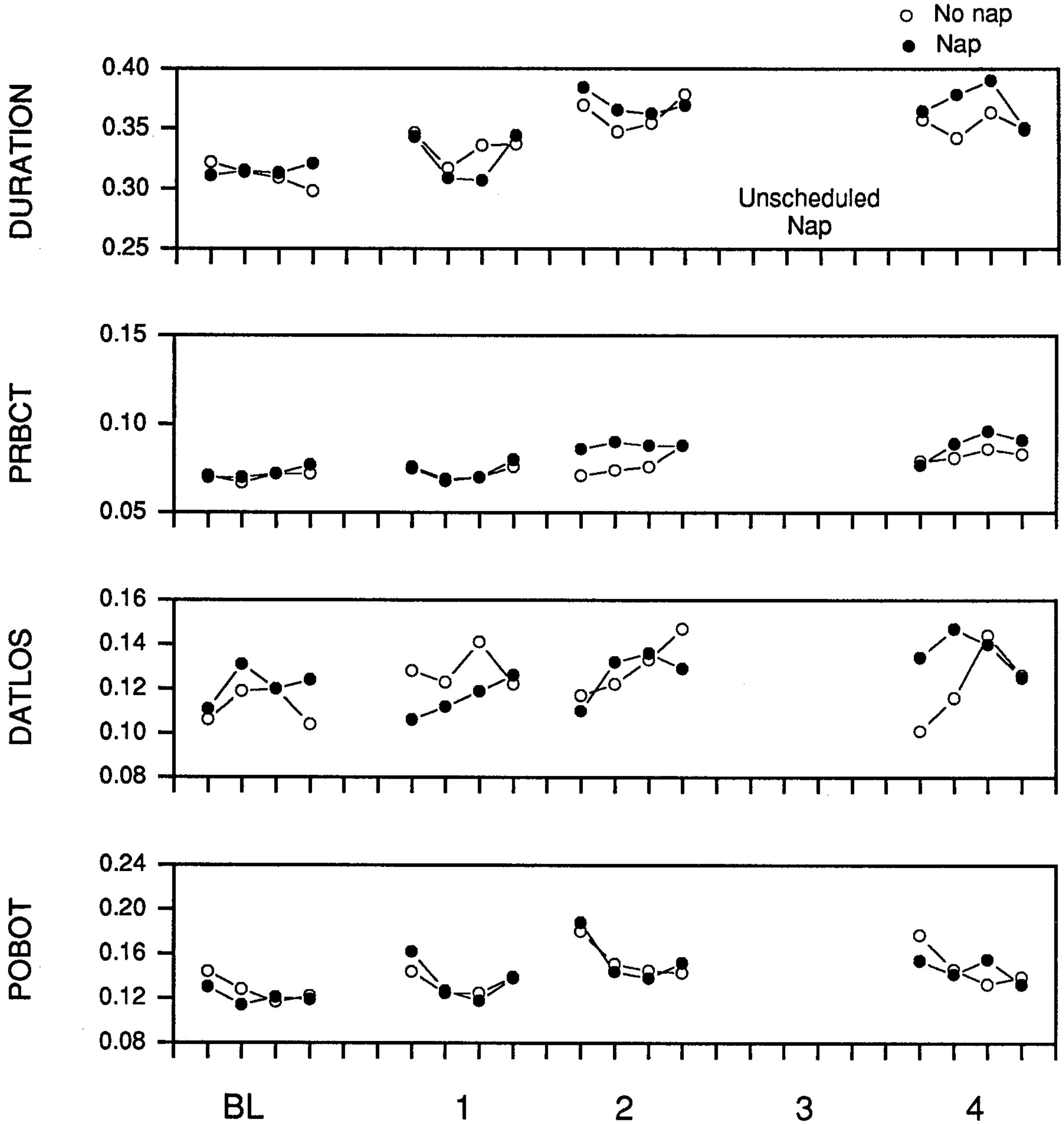


Figure I7. Eye Tracker Blink Measures - Sb08

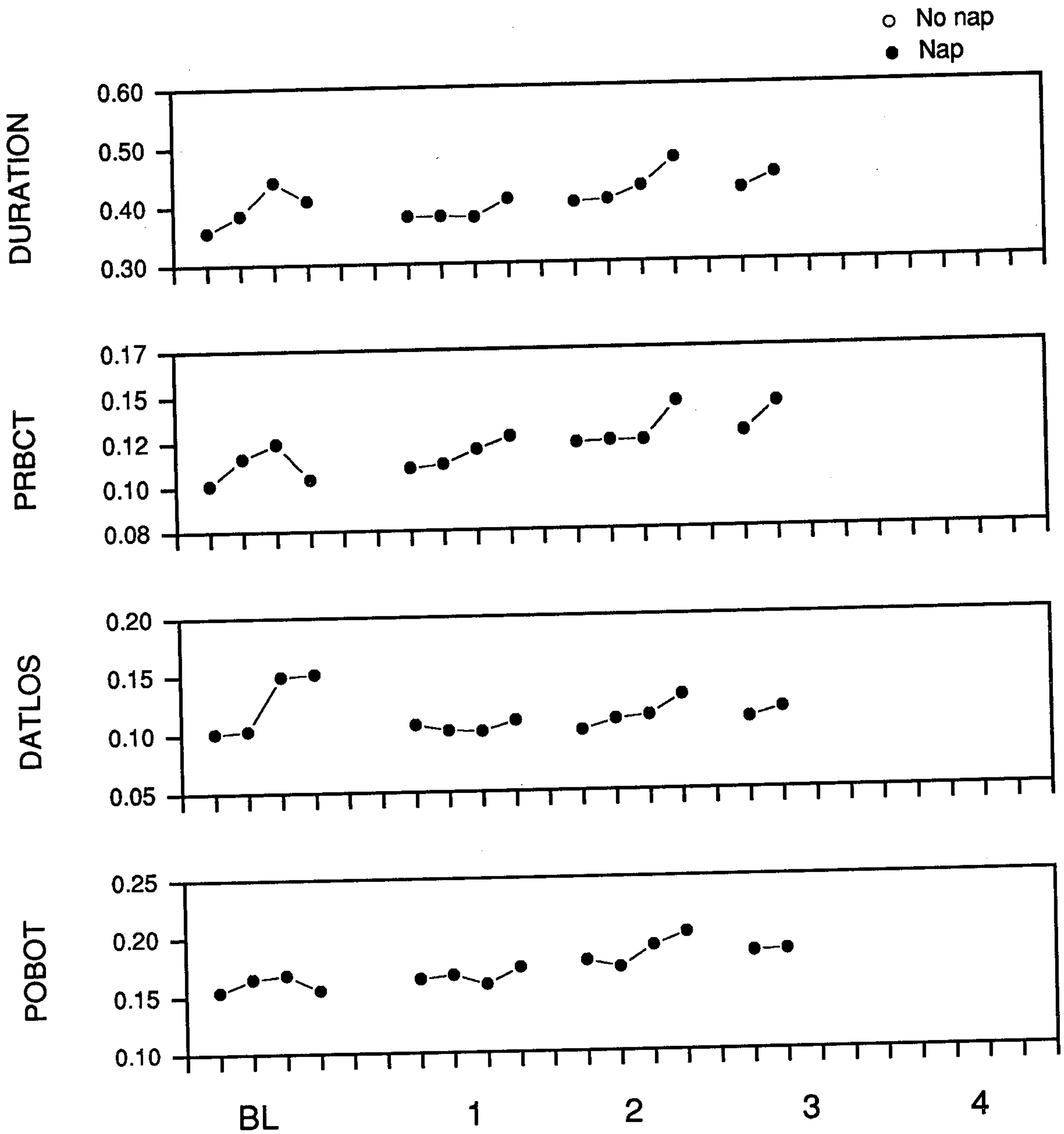
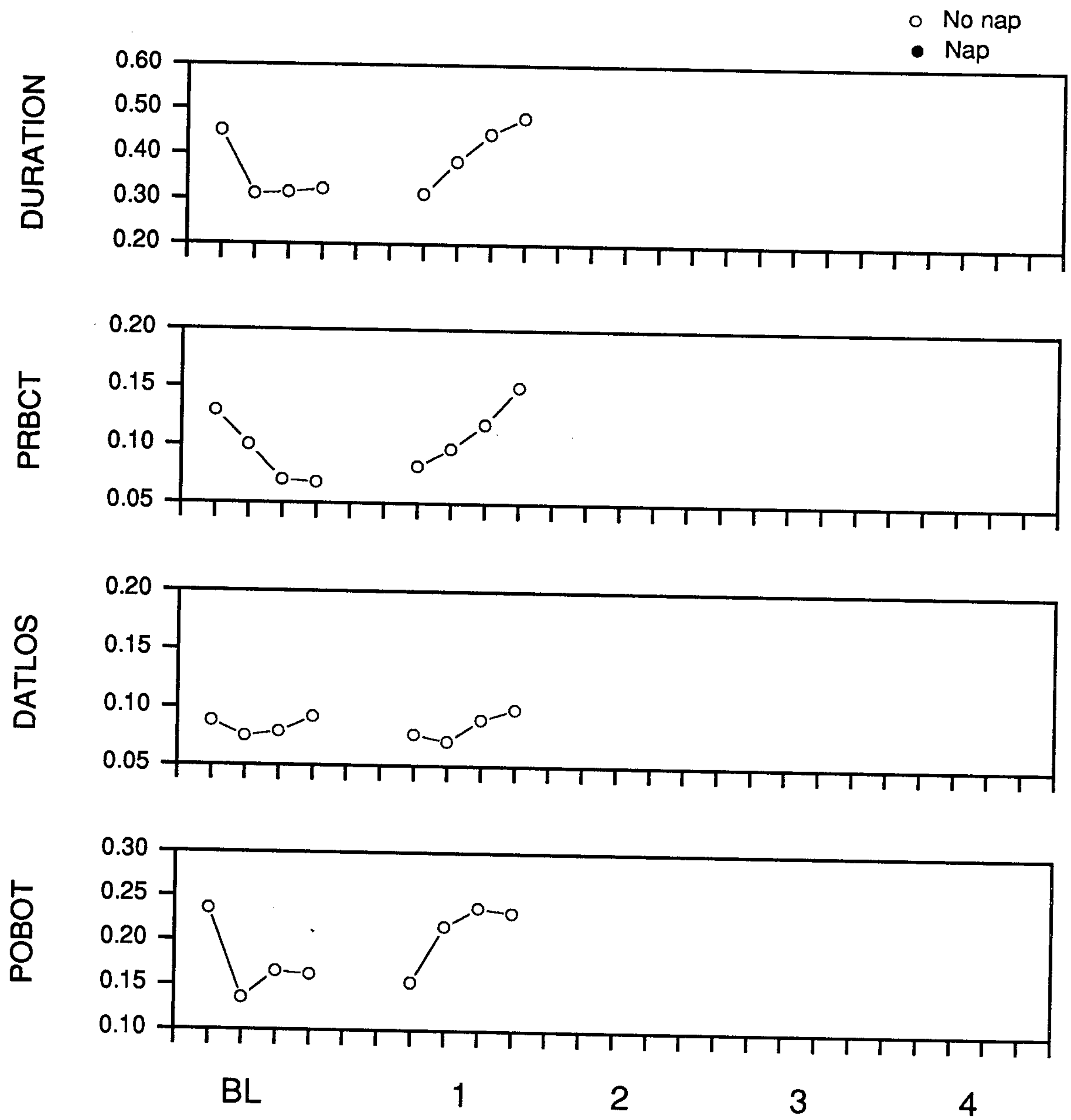


Figure I8. Eye Tracker Blink Measures - Sb09



APPENDIX J

Figures J1-J7. V/H during fixations for 4 subjects.

Figure J1. V/H during fixations - Subject 2 No-nap

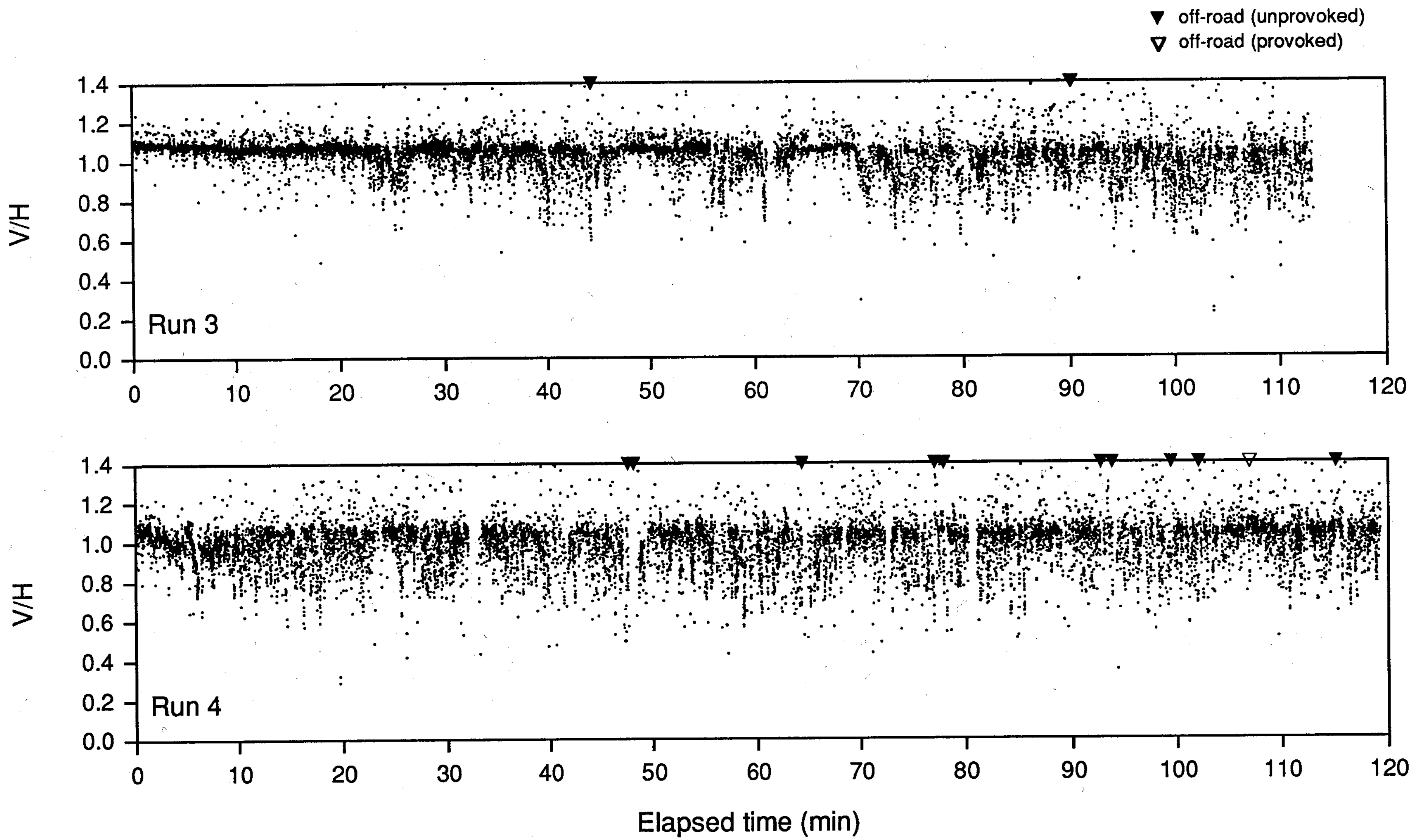


Figure J2. V/H during fixations - Subject 4 Nap

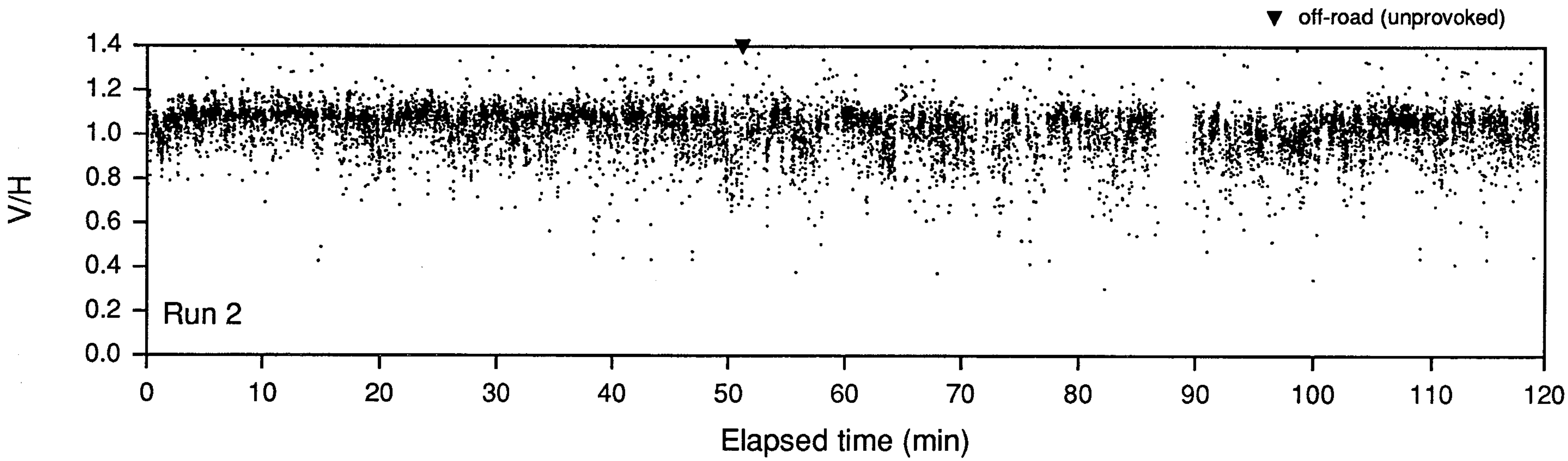


Figure J3. V/H during fixations - Subject 5 No-nap

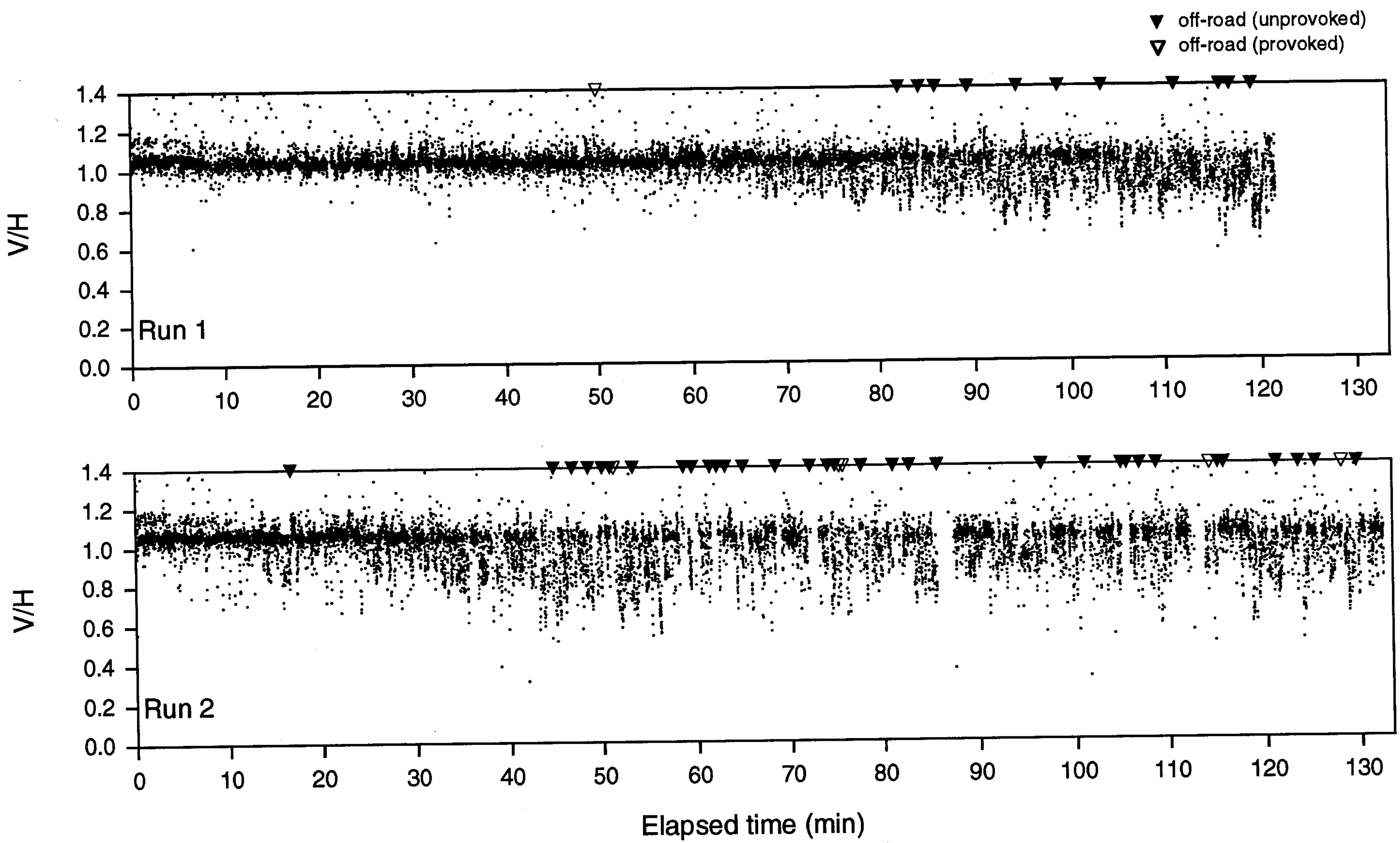


Figure J4. V/H during fixations - Subject 5 Nap

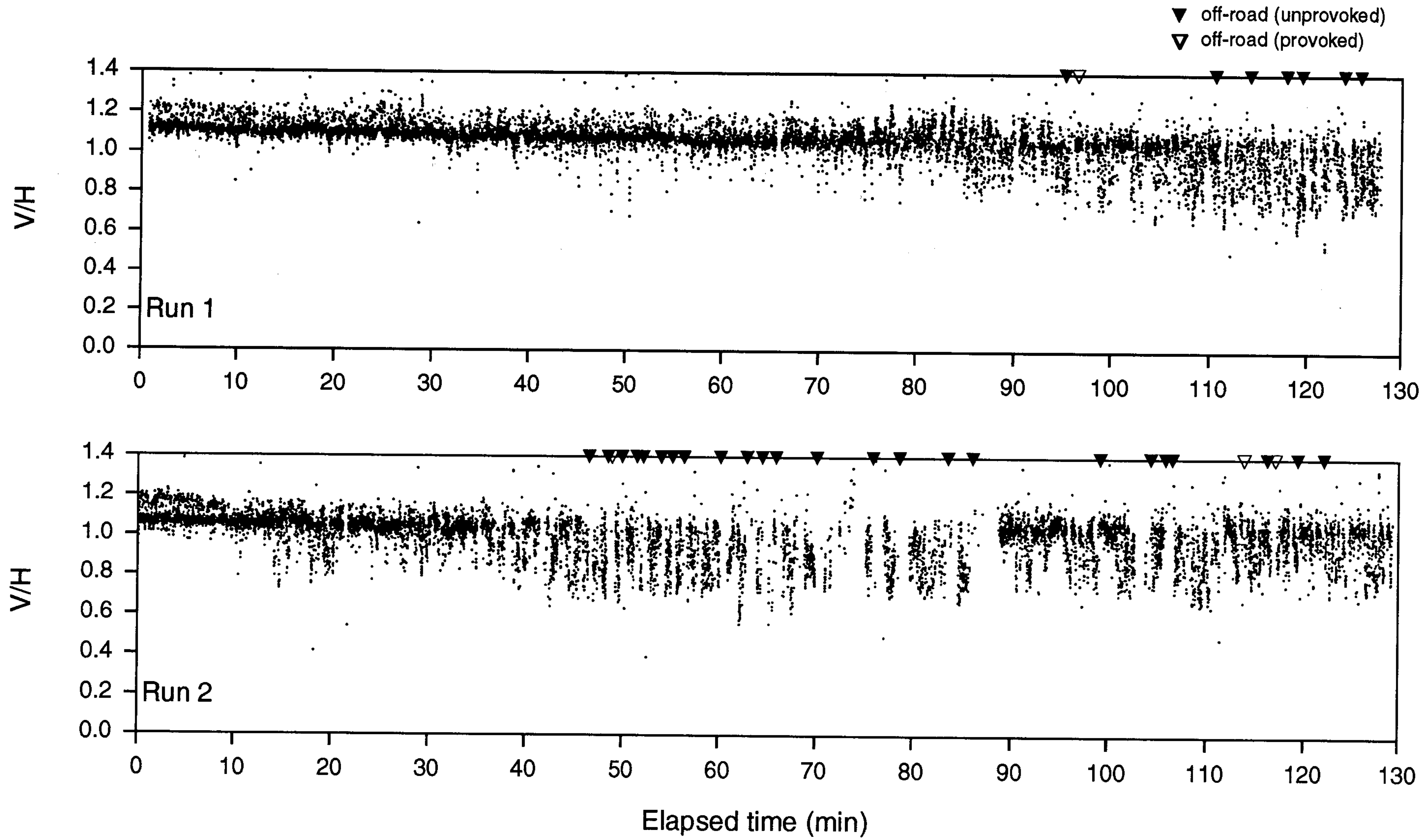


Figure J5. V/H (running average over 80 fixations) for an entire driving session

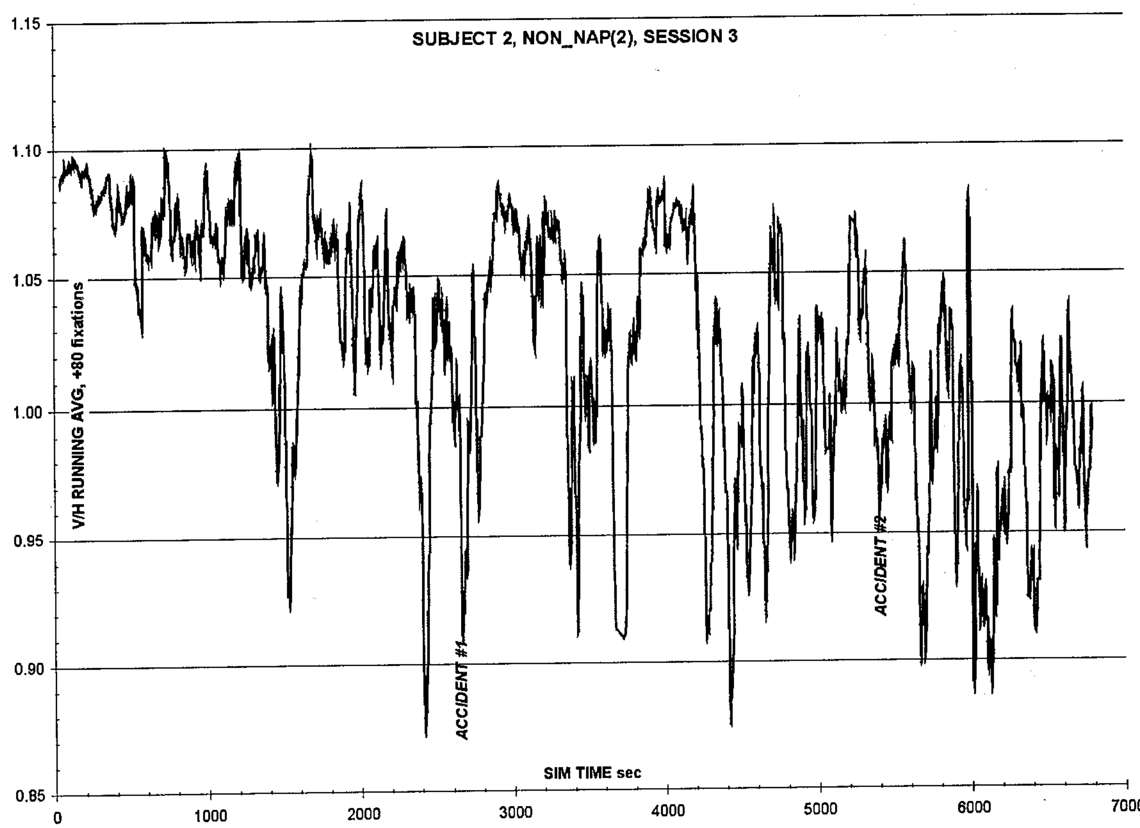


Figure J6. V/H (running average over 80 fixations) early in a driving session (top) and later in the same driving session (bottom) during a period that included an off-road accident.

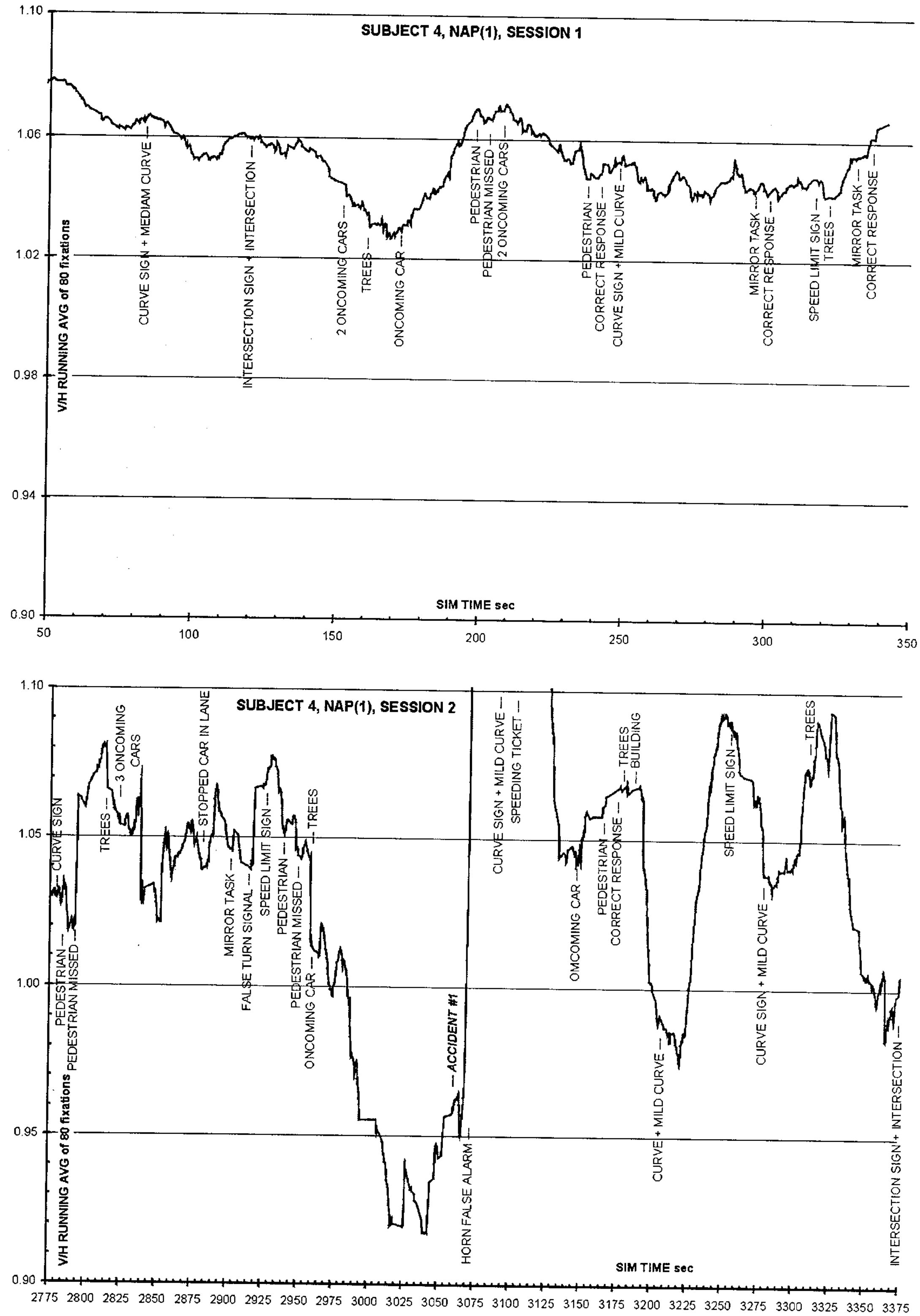
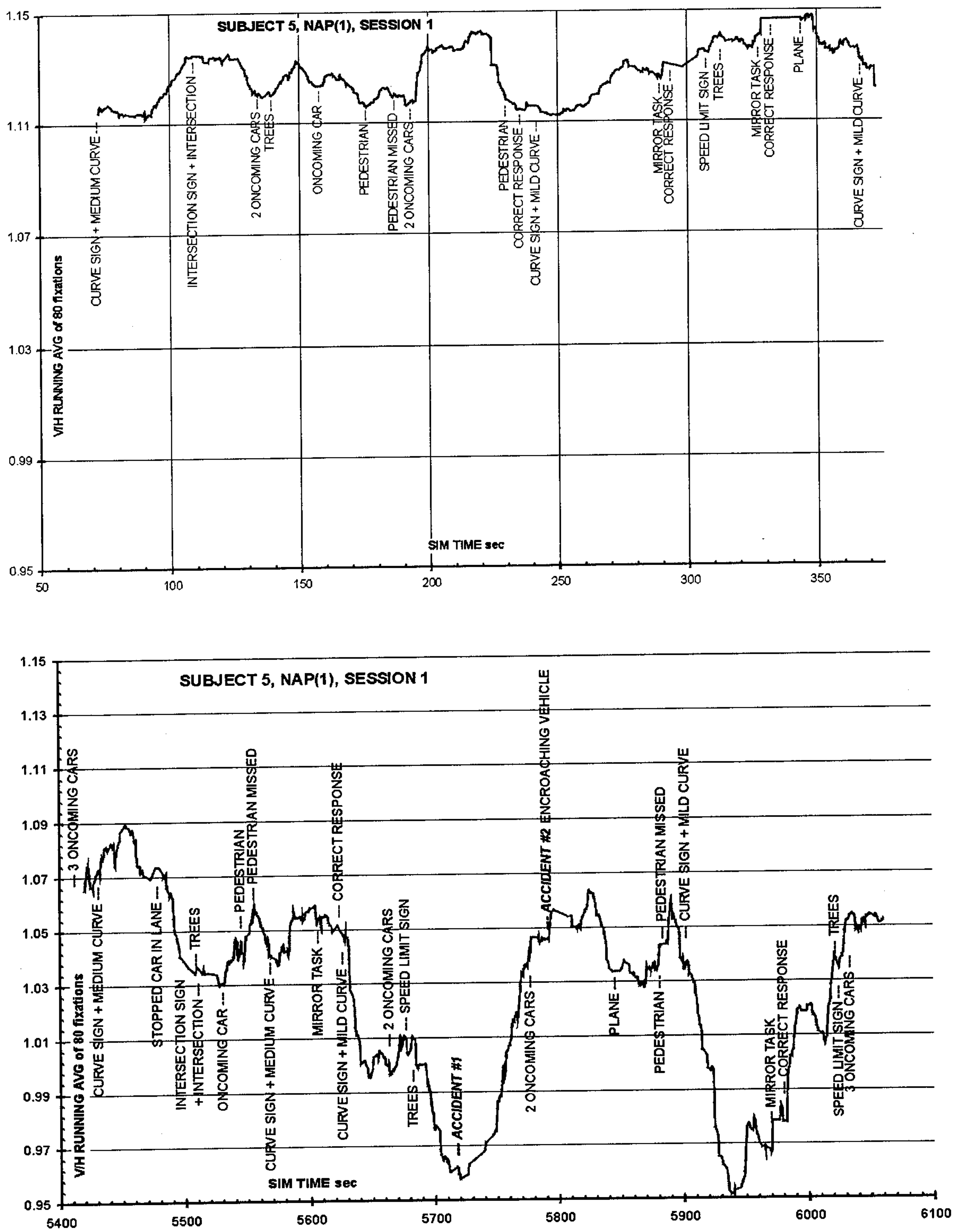


Figure J7. Another example of V/H (running average over 80 fixations) early in a driving session (top) and later in the same driving session (bottom) during a period that included an off-road accident.



APPENDIX K

Figures K1-K4. Fixation durations for 4 subjects.

Figure K1. Fixation duration - Subject 2 No-nap

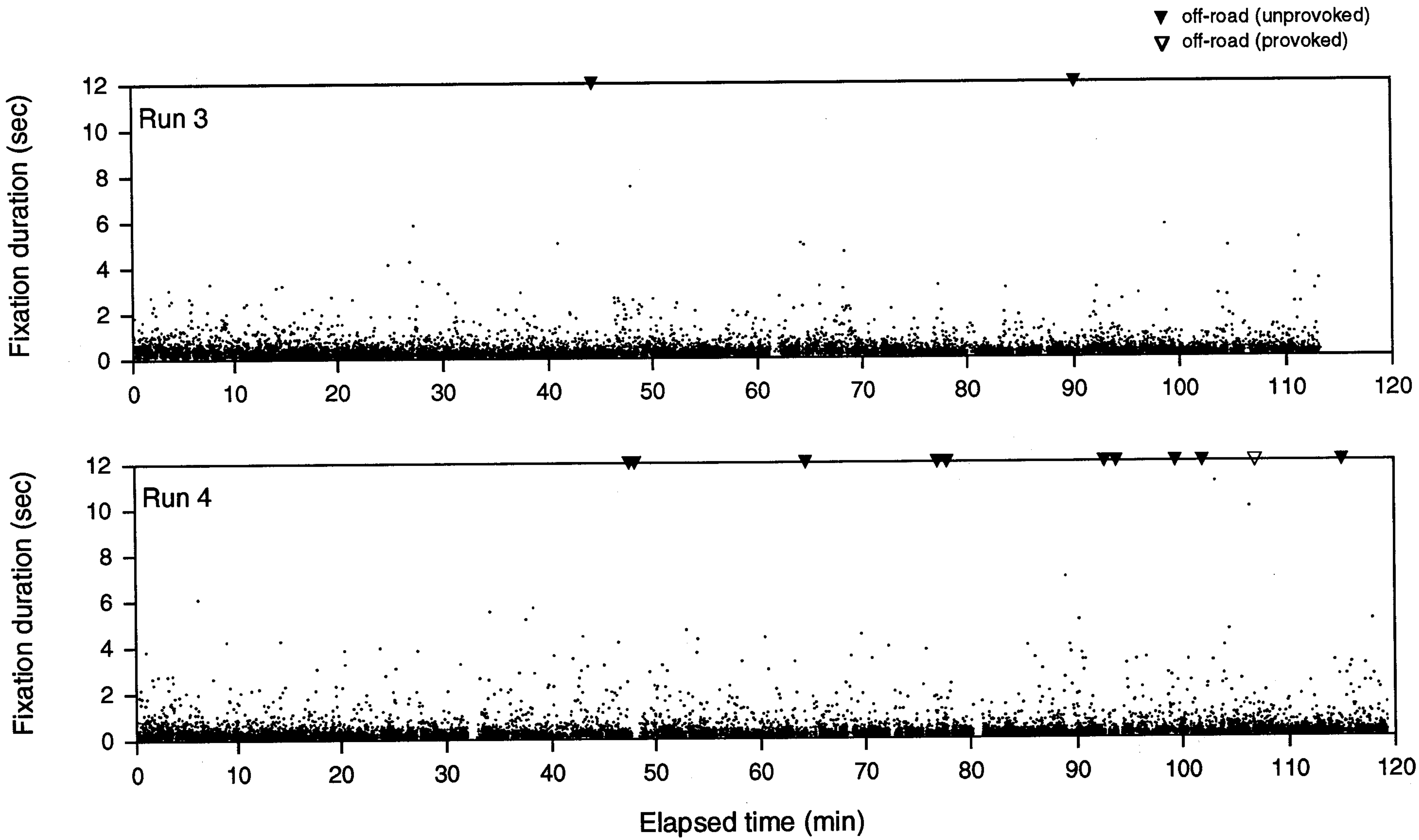


Figure K2. Fixation duration - Subject 4 Nap

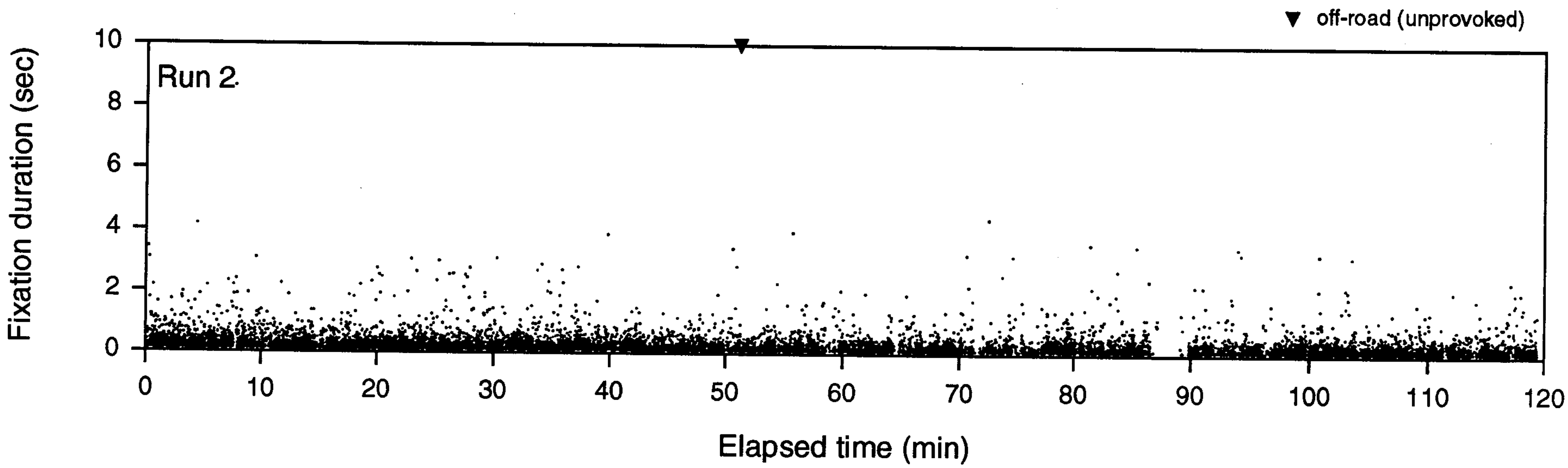


Figure K3. Fixation duration - Subject 5 No-nap

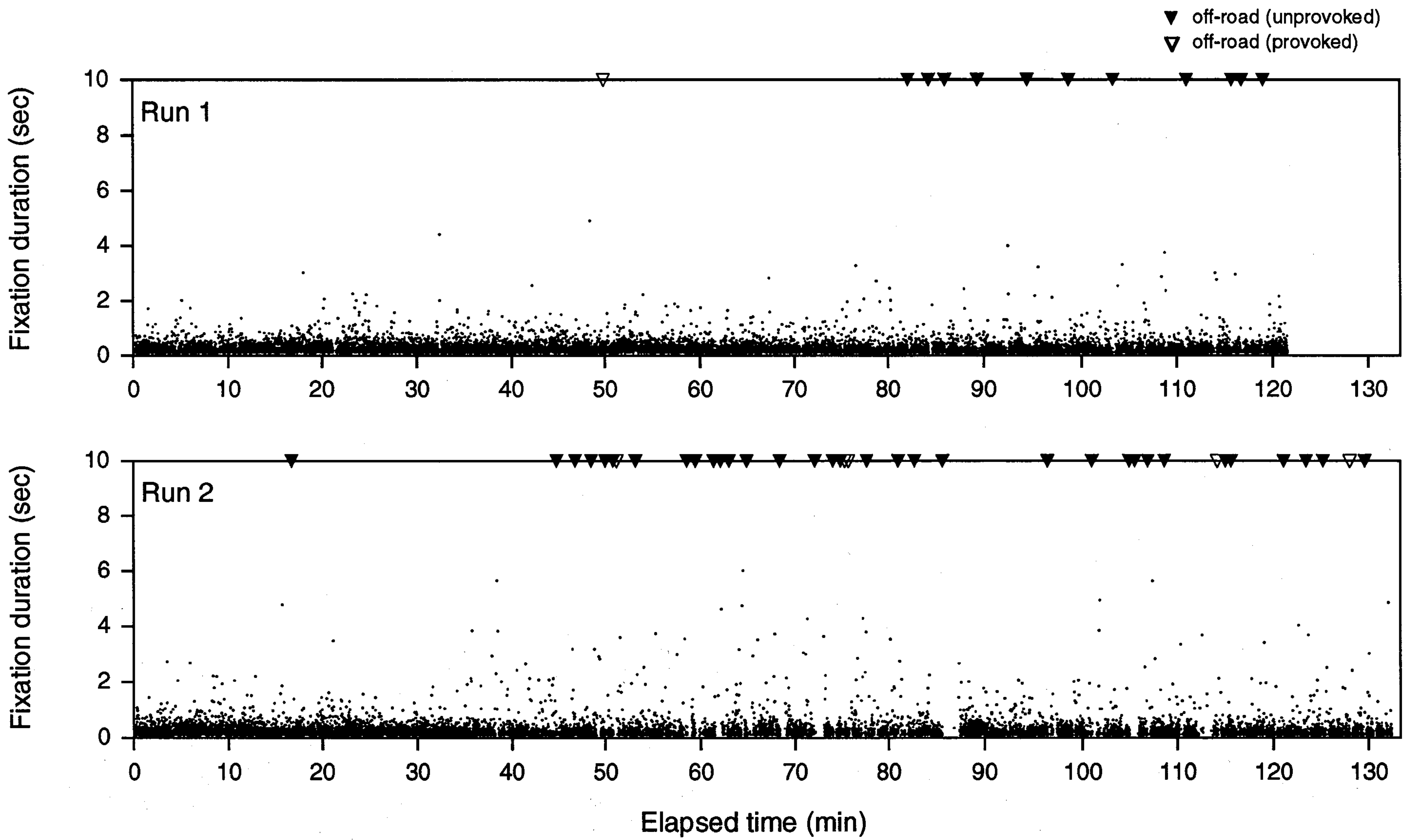
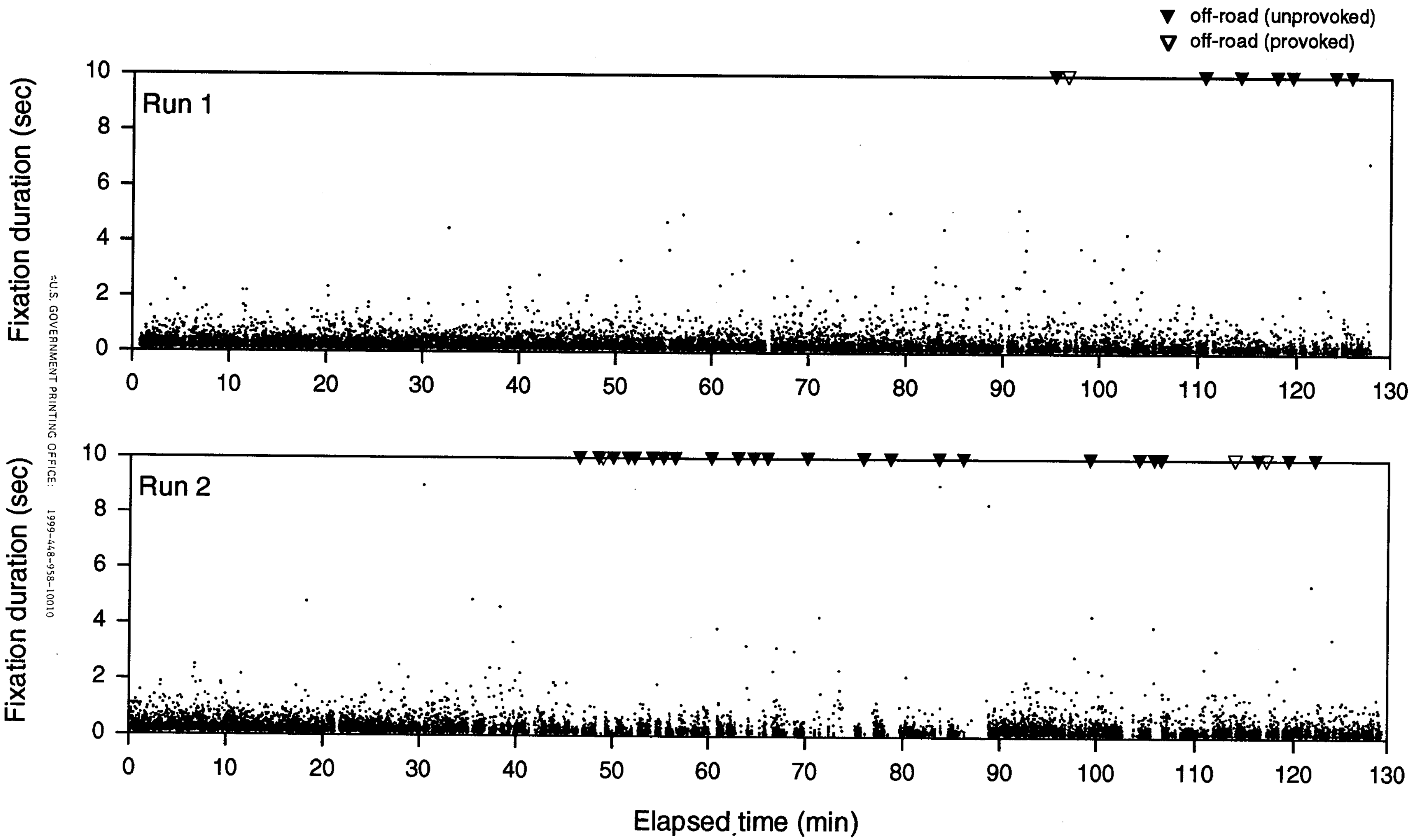


Figure K4. Fixation duration - Subject 5 Nap



☆U.S. GOVERNMENT PRINTING OFFICE: 1999-448-958-10010