AD__________________

MIPR NUMBER 96MM6651

TITLE: The Variability of Electrophysiological and Neurobehavioral Measures of Alertness in Women and Men

PRINCIPAL INVESTIGATOR: David F. Neri, Ph.D.

CONTRACTING ORGANIZATION: Naval Health Research Center
San Diego, California 92152

REPORT DATE: May 1997

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
May 1997

3. REPORT TYPE AND DATES COVERED
Final 22 Dec 95 - 1 Apr 97

4. TITLE AND SUBTITLE
The Variability of Electrophysiological and Neurobehavioral Measures of Alertness in Women and Men

5. FUNDING NUMBERS
96MM6651

6. AUTHOR(S)
David F. Neri, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Health Research Center
San Diego, California 92152

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

The routine requirement of continuous manning in the armed forces means military personnel must frequently perform their jobs for extended periods, at all hours of the day and night, often under monotonous conditions. Much research has shown that such working conditions often produce fatigue, decreased alertness, and compromised performance. As various military operations increasingly involve women there is a need to identify and characterize any gender-related differences in the pattern and dynamics of alertness and performance decrements caused by night work and sleep loss. This study addressed whether men and women show dissimilar patterns of fatigue-related performance decrements. Neurobehavioral and/or electroencephalographic (EEG) measures of alertness and performance were collected in 32 women and 24 men during a night of sleep deprivation. Detailed between-gender comparisons were made of the neurobehavioral measures of low alertness, including patterns of response rate, accuracy, and lapse probability. Gender variability was also explicitly examined in similar neurobehavioral measures previously collected in a series of lengthy sleep deprivation/restriction protocols. Results from both sets of investigations failed to detect any reliable differences in cognitive performance and alertness between men and women as a function of sleep loss. Although there appear to be some gender-related response style differences, the present data do not support distinguishing between men and women in their ability to withstand the effects of sleep deprivation. The EEG data analysis is ongoing and will be described in subsequent reports.

14. SUBJECT TERMS Defense Women's Health Research Program
gender differences, sleep deprivation, performance, vigilance

15. NUMBER OF PAGES
94

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT
Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

X Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

David [illegible] 26 MAY 97
PI - Signature Date

TABLE OF CONTENTS

INTRODUCTION

Night work is prevalent in the military

The routine requirement of continuous manning in the U.S. military, common in a wide range of operations, means that many varieties of shift work, night work, and unusual duty schedules are widespread. Military men and women frequently perform their jobs throughout day and night, often for extended periods. These irregular and suboptimal work schedules have potentially serious consequences. For example, working at night, when the body's natural tendency is to sleep, leads to biological rhythm disruption, sleep loss, and fatigue. Many studies have documented the decreased alertness and impaired performance that result from these physiological challenges. These problems are especially pertinent in the military environment because of the critical around-the-clock monitoring tasks performed by service personnel of both sexes. The difficulties have been recognized by the Office of Technology Assessment (U.S. Congress, 1991): "*Military operations also involve many tasks that require constant vigilance, often under conditions not conducive to alertness. These include air defense radar and electronic surveillance, sonar, sentries, pickets, and so on. The performance of many tasks requiring vigilance (sonar operation, for example) degrades measurably after less than an hour of duty*" (p. 187). In these critical jobs, which by their very nature involve periods of low stimulation and boredom, alertness and performance degradation could compromise mission performance and safety.

Relevant research has focused on men

Historically, the vigilance tasks in the military critical to command and control, communication, monitoring, etc., have been performed by men. As these tasks increasingly involve both male and female participants, there is a need to identify and characterize gender-related differences in the pattern and dynamics of alertness and performance decrements and in the experience of fatigue caused by night work and sleep loss. Unfortunately, the vast majority of laboratory and field studies examining changes in alertness and performance over time have been performed using primarily male subjects. It is unknown to what extent these findings can be generalized without modification to women. Furthermore, recently reported differences in male and female brain function, as documented by functional magnetic resonance imaging, indicate the possibility that men and women may rely on different brain resources in performing tasks requiring sustained attention. Understanding any gender-related differences in sustained attention, the performance of vigilance tasks, and susceptibility to the effects of sleep deprivation could have important implications for selection and training of watchstanders, sonar/radar operators, and others who work at night or with restricted sleep.

Need for research on women and fatigue

The first recommendation of the Institute of Medicine's Committee on Defense Women's Health Research (Institute of Medicine, 1995) is that "*Research is needed to address women's* *physiological, psychological, and behavioral responses to a combination of operational stressors such as ... fatigue ...*" (p. 4). A literature review shows a surprising dearth of data specifically examining the performance of fatigued women compared with men. This research project examining gender differences in the alertness- and performance-related effects of sleep deprivation is a direct response to the committee's recommendation to address women's physiological, psychological, and behavioral responses and compare them with the responses of men to the same operational stressors. This research also fills a void in the literature on the effects of sleep deprivation while addressing a research question that is directly related to the ability of military women to perform newly added mission responsibilities -- responsibilities now including many command and control, monitoring, and watchstanding functions that were until recently the exclusive responsibility of men.

Gender differences in cognitive performance

Although there are considerable similarity and overlap in cognitive abilities between the sexes, there are also well-established male-female differences for several types of cognitive performance, which have been found to hold cross culturally (Mann et al., 1990). Males generally show superior visual spatial abilities, revealed by tasks involving spatial relations and mental rotation, while females generally show superior verbal skill, revealed by performance on tasks of verbal memory and verbal fluency (Mann et al., 1990). Whether these differences increase or diminish under the physiological load of sleep deprivation, or whether the pattern of differences changes when men and women are heavily stressed by fatigue and sleep loss appears never to have been directly investigated.

Fluctuations in vigilance performance

Since the landmark investigations of Mackworth (1948), studies of human alertness have confirmed that, despite sincere intentions, few watchstanders remain "unwaveringly" vigilant while engaged in monotonous monitoring tasks. On average, detection rates in laboratory tests begin to degrade after 2-3 minutes, eventually reaching a plateau during which only 70-80% of targets are detected (Davies & Parasuraman, 1982). While most vigilance research has focused on measuring mean trends in performance, within individual sessions performance tends to fluctuate irregularly. Over the last century, scattered studies have presented evidence for the presence of rhythmic oscillations in performance at cycle lengths ranging from a few seconds to several minutes for measures including time estimation, threshold detection, reaction time, and spontaneous speaking rate (Seashore & Kent, 1905; Stebel & Sinz, 1971; Stroud, 1966; Treisman, 1984; Warner, 1979; Wertheimer, 1953). Although sustained human attentional behavior may be expected to exhibit dynamics on all time scales, knowledge of basic neurophysiological processes that could underlie minute-scale fluctuations in performance is advancing slowly (Churchland & Sejnowski, 1988). In general, no systematic replications or comparisons of these claims have been made.

Most previous vigilance research has used relatively low target presentation rates (1-2/min or less), a sampling rate too low to document minute-scale fluctuations. In experiments designed to simulate the auditory environment of passive sonar, a wide range of performance dynamics was observed, and averaged evoked responses to both task-relevant and task-irrelevant probe tones were found to covary with a measure of local error rate (Makeig et al., 1990). The purpose of those experiments was twofold: first, to better understand the dynamics and electrophysiological correlates of lapses in performance, and second, to evaluate the information available in the EEG that might be used to perform automatic monitoring of operator readiness to detect and respond to sonar contacts. Results have shown that both slow multi-minute waves of drowsiness (Makeig & Inlow, 1993) and more rapid sub-minute cycles in auditory and visual awareness linked to drowsiness (Makeig & Jung, in press) are accompanied by precise changes in the EEG spectrum.

Gender differences in vigilance performance

There has been very little research investigating differences between men and women in the specific area of sustained attention and the ability to maintain vigilance while fatigued. No gender differences were reported in studies of men and women performing a 60-min auditory vigilance task (Tong et al., 1977) or a simulated 120-min radar task (Thackray, Touchstone, & Bailey, 1978). On the other hand, gender differences have been reported on a simple vigilance task (Waag, Halcomb, & Tyler, 1973), in the choice of vigilance tasks undertaken (Dember, Galinsky, & Warm, 1992), in identification of complex visual images (Geheb, Whitfield, & Brannon, 1994), and in the performance tradeoff between speed and accuracy (Blough & Slavin, 1987). It has also been suggested that male and female vigilance performance may be differentially affected by stressors (Holding & Baker, 1987). Yet there have been no *systematic* studies of the effects of fatigue on performance and related neurobehavioral functions in healthy men and women.

Recent preliminary EEG data from the Naval Health Research Center (NHRC), however, indicate that there may be reason to suspect gender differences in the ability to maintain alertness during prolonged vigilance testing in non-sleep-deprived individuals. In a study of the EEG correlates of performance, it was noted that female subjects more frequently showed an insufficient number of lapses in alertness to allow their data to be used to develop alertness algorithms. It is thus an open and important question whether men and women, when stressed from fatigue, sleep loss, and night work, differ in their patterns and dynamics of alertness and performance and in their EEG correlates to these patterns.

Monitoring alertness using EEG

Beatty et al. (1974) reported that periods of lowered performance on a visual task were predicted by lowered electrocortical activation as indexed by a simple measure of occipital theta band EEG amplitude. Townsend and Johnson (1979) reported that errors of omission in sleep-deprived subjects were correlated with changes in the EEG spectrum, although the structure of this relationship varied across subjects. Belyavin and Wright (1987) reported that changes in amplitudes in theta (4-7 Hz) and beta (14-21 Hz) EEG bands predicted changes in performance on a simple vigilance task, while delta (1-3 Hz) and beta band levels predicted fluctuations in performance on a more difficult discrimination task. Torsvall and Akerstedt (1988) observed that slow eye movements and a large increase in alpha EEG precede dozing off during a simple visual task. Recently, Ogilvie et al. (1991), using an auditory response task administered at the beginning of all-night sleep sessions, have claimed that power in all EEG bands increases when sleep-related lapses first occur.

Rapid changes in alertness

Makeig and Jung (in press) have recently reported that tonic and phasic changes in the EEG spectrum accompanying changes in auditory detection performance appear to index the operation of discrete brain mechanisms regulating alertness and sensory awareness. They have shown that sub-minute (15-20 s) cycles in performance by drowsy subjects have particular EEG spectral correlates. Measures of event-related spectral perturbations in the human EEG time locked to target presentations during sustained auditory detection tasks confirm previous observations that intermittent gating of auditory information in drowsiness is correlated with the occurrence of irregular low-level theta band activity (5 Hz). For most subjects, during periods of intermittent performance 4-6 Hz activity begins to increase about 10 seconds before undetected targets occur. Detected targets are preceded by a similar rise in power in the gamma band (above 30 Hz). These phasic perturbations are accompanied by parallel changes in the detection rate for adjacent targets. During extended periods of drowsiness (as evidenced by poor detection performance), these phasic changes are superimposed on slower tonic changes in both performance and the EEG spectrum. In particular, mean power at the human sleep spindle frequency (12-14 Hz) is tonically elevated during sustained periods of poor or absent performance, but its second-to-second amplitude is not coupled to performance fluctuations. The present project allows study of gender differences in slow and rapid alertness cycles and their EEG concomitants.

Alertness monitoring and management system (AMM)

A prototype EEG-based real-time Alertness Monitoring and Management (AMM) system has been developed at NHRC to predict alertness-related decrements on performance based on the EEG spectrum (Makeig, Elliott, & Postal, 1993). Ongoing research is improving that system and generalizing it to different types of performance. Data collection has been completed in a study comparing results with the auditory detection tasks and a compensatory visual tracking task. An inspection of the performance records shows that the visual tracking task can efficiently record fluctuations in alertness over a relatively short period (e.g., less than 1 hr) for most subjects. No work has yet been done on sleep-deprived subjects using this task, and there has been no systematic examination of differences in patterns of performance lapses in women and men.

EEG correlates of alertness

Profound changes in the appearance and spectrum of the EEG with sleep and drowsiness have been noted since the first investigations of Loomis, Harvey, and Hobart (1937). Later, traditional EEG power band measures were shown to correlate with sleep stages defined by visual inspection of the EEG record (Rechtschaffen & Kales, 1968). Matousek and Petersen (1983), using a linear combination of more than thirty band and band-ratio amplitude measures, could reproduce the visual classification of EEG epochs into awake and Stage I sleep states. Penzel and Petzold (1989) observed that hand-scored EEG vigilance estimates of sleep onset periods could be predicted reliably using measures of the cumulative power spectral distribution. The link between EEG and arousal is well enough accepted that many studies of vigilance simply define vigilance using electroencephalographic (EEG) and electrooculographic (EOG) criteria (Fruhstorfer & Bergstrom, 1969).

EEG correlates of performance

Most studies correlating task performance and the EEG spectrum have computed measures on only a few EEG spectral bands defined a priori (Williams et al., 1962), rather than computing full-spectrum correlations between performance and behavior. These reports have generally found that the correlation between EEG band amplitudes and performance varies depending on task, performance measure, subject state, and electrode site. Detailed studies of the appearance of the EEG record in drowsiness have also noted that in different subjects the EEG may exhibit a variety of routes in transition from waking to sleep (Santamaria & Chiappa, 1987). The relation between alertness and the EEG spectrum in most of these studies has been measured in terms of changes in mean power in conventionally-defined EEG frequency bands and the strength of the relationship evaluated by correlating changes in mean performance and EEG power or amplitude across relatively lengthy experimental sessions or blocks. The relationship estimated in this manner should be expected to vary not only with subject state, EEG spectrum, task and instructions, but also with length of the time blocks used. The small amount of evidence leaves some uncertainty of the reliability of gender differences in vigilance performance. However, the response-frequency required by typical vigilance tasks is generally too low to document any minute and sub-minute fluctuations in alertness that may be present. The question of gender differences in vigilance performance remains unanswered.

Recent research conducted at NHRC and elsewhere has documented minute-scale and sub-minute-scale fluctuations in alertness that can significantly impair cognitive performance (Makeig & Inlow, 1993). This research also has shown that information about an operator's cognitive state is continuously available in the operator's electroencephalogram (EEG). Makeig and Jung (in press) have correlated fluctuations in the EEG spectrum with concurrent changes in performance on an auditory detection task and have shown that applying the output of the principal components analysis of the EEG power spectrum to training an artificial neural network results in very accurate estimation of shifts in an operator's level of alertness and detection performance.

Gender differences in EEG manifestations of alertness and performance

We have found no previous investigations of differences between the sexes in EEG manifestations of decreased alertness. Gender differences exist in sleep EEG characteristics, however. (For example, sleep disorders centers use different normal values for sleep stage proportions in males and females). The recently reported differences in the functional organization of male and female brains for language, as documented by fMRI (Shaywitz et al., 1995), suggest the possibility that men and women may rely on different brain resources in performing several tasks, including those requiring sustained attention. The results suggest that men and women also may manifest different brain electrical patterns in association with alertness-related performance decrements. If this were the case, then automated Alertness Monitoring and Management (AMM) technology, currently being developed at NHRC might have to be adapted differently to men and women.

General approach to this research

There are two thrusts to this research program. The first involves experimental research at NHRC on gender differences in behavioral and EEG measures of alertness during one night of sleep deprivation. Ongoing research at NHRC reveals that it is possible to derive real-time estimations of alertness-related performance decrements from fluctuations in the EEG spectrum (Makeig, Elliott, & Postal, 1993; Makeig & Jung, in press). Minute-level fluctuations in alertness, sufficient to compromise vigilance and performance, have been demonstrated in non-sleep-deprived individuals. Preliminary data have suggested more female than male subjects produced an insufficient number of lapses in alertness to allow their data to be used to develop alertness algorithms. This potential advantage for women in maintaining vigilance and avoiding critical lapses needs to examined in a study directly comparing the behavioral and EEG spectral manifestations of low alertness between the genders. Unlike earlier studies, the present project involves one night of sleep deprivation, in which the physiological load on the subject is increased compared with the daytime. These studies include performance measures compatible with those used in the second thrust of this research program, described below, to allow for comparison of the data.

The second research thrust in this proposal is a collaboration with Dr. David Dinges of the Unit for Experimental Psychiatry (UEP), University of Pennsylvania School of Medicine, to analyze data previously collected in a series of sleep deprivation protocols. These analyses measure the gender variability in these neurobehavioral measures, directly addressing whether systematic variability between genders exists in these measures, the nature and form of any such variability, what may account for it, and its magnitude relative to other systematic variables that might limit the performance and drive the selection and training of sonar/radar operators and others who perform critical watchstanding tasks.

Hypotheses

These hypotheses are based on preliminary studies at NHRC of the EEG correlates of alertness in non-sleep-deprived women and men and on preliminary analyses at UEP on women and men undergoing repeated performance and mood testing during controlled laboratory exposure to night work and sleep deprivation. We hypothesize that, although women will report greater fatigue than men, they will maintain or increase certain performance advantages on specific tasks. Specifically:

- During night work and sleep deprivation, female performance on subject-paced cognitive tasks will be better by 5-10% than males in terms of cognitive speed and accuracy. Sleep deprivation is expected to produce marginally greater changes in female cognitive performance than male cognitive performance, owing to this higher baseline performance in females;

- During night work and sleep deprivation, male *optimum* reaction time performance will be faster than that of females by 5-10%. While males are not expected to differ from females in overall psychomotor vigilance performance (i.e., lapse frequency, vigilance decrement slope), sleep deprivation is expected to produce greater increases in male false responding, reflecting the male tendency to compensate for fatigue-induced performance lapses with increased anticipatory responding (i.e., guessing);

- During night work and sleep deprivation, females will report greater fatigue and loss of mental energy than males. These hypothesized differences are expected to interact with sleep and circadian phase, such that females are expected to experience greater fatigue than males near the circadian nadir in body temperature and in the morning, including during morning periods after recovery sleep. These differences are expected to reflect the pathophysiologic tendency for females to experience morning fatigue, possibly due to greater sleep inertia. In preliminary analyses, we found that following recovery sleep from 64 hrs of total sleep deprivation, women experienced significantly higher levels of fatigue, with the difference diminishing as time of day approached the circadian peak in body temperature ($F_{6,132} = 3.43$, $p < .029$);

- Men and women deprived of similar amounts of sleep will show different patterns of EEG spectral changes associated with periods of low alertness;

- Besides the hypotheses described above, the project will explore gender-related differences in accuracy of performance efficiency during night work and sleep deprivation, and perceived effort required to perform, physical and mental exhaustion, other aspects of mood (e.g., tension-anxiety, depression-dejection, vigor, physical health), and motivation.

Technical Objectives

The technical objectives are to:

- Collect simultaneous alertness, performance, subjective, and EEG data on male and female subjects during a single night of sleep deprivation at NHRC analyzing
  a) gender differences in the patterns and dynamics of EEG spectral changes associated with decrements in alertness and performance and
  b) gender differences on a wide range of neurobehavioral measures including performance on cognitive tasks (e.g., continuous performance, calculation performance, mental rotation, digit-symbol substitution, reaction time, and compensatory tracking) and subjective ratings of mood and fatigue.
- Using sleep deprivation data previously gathered at UEP, analyze gender differences in a wide range of neurobehavioral outcomes including performance on cognitive tasks (e.g., serial addition, symbol substitution, perception, psychomotor vigilance, reaction time, short term recall memory, and recognition memory) and subjective ratings of fatigue, self-monitoring, arousal-activation, and mood states.
- Evaluate the profile of each objective and subjective neurobehavioral metric before, during, and after night work and sleep deprivation.
- Compare the circadian profile of body temperature between males and females.
- Perform statistical analyses to reveal the relative contributions and interactions of both the circadian pacemaker and homeostatic sleep drive to waking neurobehavioral functions in men and women.

BODY

Methods: NHRC Study

The research thrust at NHRC was designed to define and compare the electrophysiological (EEG) and neurobehavioral patterns of vigilance decrement in men and women under conditions of one night of sleep deprivation.

Subjects. A total of 65 subjects initiated training. Subjects disenrolling prior to the experiment start and the loss of data for one or more sessions due to technical problems resulted in a maximum of 56 subjects being available for most neurobehavioral analyses: 24 healthy normal males (mean age 28.4 yrs ± 7.6) and 32 healthy normal females (mean age 28.2 yrs ± 9.2). Occasional technical difficulties led to the loss of data for a particular task for a given subject and so the actual number of subjects varied slightly from task to task. These fluctuations were small and the final sample size was in all cases much larger than the 40 subjects promised in the original proposal. EEG data were collected on a subset of 40 subjects (18 males and 22 females). The subjects were recruited as paid volunteers from local colleges. (Paid civilian volunteers were used because experience has shown that insufficient military personnel are available to participate in sleep deprivation studies.) Our recruitment procedures were unbiased and considered all applicants regardless of race, color, creed, or national origin. Every attempt was made to ensure minority representation. Subjects were medically screened before the study.

Subject screening. Approximately one week before the study, subjects reported to the laboratory for a screening, orientation and initial training session. Subjects were given a complete description of the study, tour of the laboratory, completed screening questionnaires, provided written, informed consent, and underwent initial training on the cognitive task battery. The study description included a complete explanation of all study procedures and policies, including the right to discontinue participation in the experiment at any time, without prejudice. The health questionnaire ruled out exclusionary factors (e.g., medication/drug use, birth control medication, smoking, history of significant sleep or fatigue problems, current night work, and limiting types of pathology such as sensory/perceptual disorders, motor dysfunction, epilepsy, and other chronic illnesses). A morningness-eveningness questionnaire (Smith, Reilly, & Midkiff, 1989) was used to quantify the behavioral tendency toward morning or evening behavior. Subjects were also issued Mini Motionlogger wrist actigraphs (Ambulatory Monitoring, Inc.) and sleep logs for recording of sleep/wake activity during the 6-8 days prior to the study start. During this period subjects were asked to keep a regular schedule, obtaining ≥ 7 hours sleep per night, with their bedtimes and waketimes varying by no more than one hour each day. During this time, the wrist actigraph served to monitor the subjects for compliance with this sleep/wake schedule besides providing baseline sleep/wake data. Initial training occurred on all cognitive and subjective tasks. This orientation and training session lasted ~ 3 hr.

Equipment. The equipment for collection of EEG data consisted of a NeuroScan system including NeuroScan 3.0 software with Scan/Stim/Amp (64 channels) hardware system and an

ALR 486 PC connected to the NeuroScan. The cognitive performance battery was presented to the subjects using networked 386 PC computers with ega/vga graphics. The PCs were synchronized to the NeuroScan via a parallel port connection such that the EEG data files could contain markers indicating when stimuli presentations and responses occurred.

Procedure. On the day of the study, subjects were requested to refrain from caffeine use after 1200 and to report to the Sleep Laboratory at 1600. They received dinner and further training on the battery of performance and subjective tasks until near-asymptotic levels of performance were obtained. This required an average of six sessions per task. Subjects were then outfitted with scalp electrodes and performed six iterations of the 90-min test battery beginning at 2000, 2200, 2400, 0200, 0400, and 0600. Three to four subjects were tested simultaneously in the same room. Subjects were visually isolated from each other during testing by partitions. All tasks were administered and the performance data collected using personal computers. EEG data were collected using a modified NeuroScan EEG system. Subjects worked in ordinary levels of room illumination. Short rest periods were scheduled between individual tasks in the battery and 30-min rest periods between sessions of the battery. Subjects were prevented from napping or leaving the laboratory and their activity was monitored for compliance by technicians who were present always. After testing, subjects either were returned to their local college dormitory or home for recovery sleep, or obtained recovery sleep in our lab.

Electroencephalography. EEG data were recorded continuously during task performance. The cognitive test battery was programmed to provide parallel-port sync signals on each stimulus and response delivery. Assembly language routines were embedded in the C programs to insure synchronization. The 4 sync signal cables from the PCs running the tests were level-adjusted and optically isolated through a custom box produced by Neuroscan, then fed into channels 8, 16, 24, and 32 of the 32-channel Synamps EEG amplifier system (Neuroscan). Subject EEG was recorded from 6 scalp channels (three frontal, three parietal) and from a bilateral electrooculographic (EOG) channel. Laboratory standards to prevent possible spread of infection were followed. The EEG and EOG data from the four subject stations were delivered to the Neuroscan amplifiers channels 1-7, 9-15, 17-23, and 25-31. Data were digitized by the Synamps system at a sampling rate of 250 Hz per channel with 12-bit resolution. Data recording was continuous in each 90-minute session. At the beginning of each task, a series of rapid sync signals were sent separately from each PC to mark task onset. Task onsets were self-paced by the subjects, under the verbal direction of the experiment technicians.

This procedure allowed for EEG and performance data to be analyzed to compare EEG spectral parameters, including evoked responses and spectral perturbations to task events, during periods of high and low performance. Presence and relative strength of cycles in alertness, performance, and EEG with lengths from 15 s to several minutes can be compared in male and female subjects through the six 90-min test sessions. Individual models of EEG and performance changes in the visual tracking and continuous performance tasks can be derived from first task sessions and used to process EEG recorded during later sessions to estimate changes in performance levels. Earlier work by Jokeit and Makeig (1994) has shown a group difference in brain processing of simple reaction time stimuli by fast and slow responders, respectively. Possible male and female differences in EEG spectral reactivity and mean performance levels can be explored in data recorded during the simple reaction time, calculation, and mental rotation tasks, and the interaction of accumulating sleep deficit with male/female differences can be determined.

Cognitive performance testing. Subjects performed a battery of cognitive, vigilance, and subjective tasks once every two hours throughout the overnight recording period. The battery was approximately 90 min in length, including short breaks interspersed between some tasks. There was a 30-min break every two hours. All tasks were administered by computer with subjects seated at work stations. Order and duration of the tasks are shown in Table I and brief descriptions are provided below.

Table I. Order, Duration, and Approximate Starting Times of Performance Tasks

Task	**Duration (min)**	**Approximate Starting Times**
Continuous Performance (CPT)	15	2000; 2200; 2400; 0200; 0400; 0600
Short Break	5	2015; 2215; 0015; 0215; 0415; 0615
Calculation Performance (CMT)	4	2020; 2220; 0020; 0220; 0420; 0620
Mental Rotation (MRT)	5	2025; 2225; 0025; 0225; 0425; 0625
Digit-Symbol Substitution (DSST)	5	2030; 2230; 0030; 0230; 0430; 0630
Short Break	5	2035; 2235; 0035; 0235; 0435; 0635
Simple Reaction Time (SRT)	15	2040; 2240; 0040; 0240; 0440; 0640
Short Break	5	2055; 2255; 0055; 0255; 0455; 0655
Compensatory Tracking (TRK)	30	2100; 2300; 0100; 0300; 0500; 0700
Subjective (VAS, KSS)	<1	2130; 2330; 0130; 0330; 0530; 0730
Long Break	30	2130; 2330; 0130; 0330; 0530; 0730
Total for Battery	120	

Continuous performance task. The continuous performance task (CPT) is a high event-rate, machine-paced vigilance task that imposes a high working-memory load on the subject. It is a variant of the Continuous Performance Task first described by Rosvold et al., (1956). The subject is presented with a sequence of letters one at a time in the center of the display. Each letter is presented for 50 msec with an inter-stimulus interval (stimulus onset to stimulus onset) of 2000 msec. The subject's task is to decide whether each letter matches the previous one in the sequence. The subject responds by pressing one key when the current letter matches the previous (target) letter and a different key when the current letter differs from the previous letter. Half the letters are the same as the target letter. Data analyses focus on reaction time and accuracy. (Task duration is 15 minutes.)

Calculation performance task. The calculation performance task (CMT) is a self-paced task modeled after one described by Klein et al. (1976). It presents the subject with a series of randomly generated pairs of 3-digit numbers, one pair at a time. The subject's task is to sum the numbers as fast as possible and enter the answer using the numbers on the top row of the keyboard. The subject is instructed to complete as many pairs as possible in the allotted time. Measures include reaction time, the number of calculations completed in the time allowed, and accuracy. (Task duration is four minutes.)

Mental rotation task. The mental rotation task (MRT), modeled after Cooper (1975), is a self-paced test of the ability of subjects to rotate 2-dimensional objects in space mentally. Subjects are presented with one of six possible numbers and letters (2,5,7,G,J,R) in a circle in the center of the screen. The stimulus can be in any one of six orientations (0°, 60°, 120°, 180°, 240°, 300°) and either reversed (mirror-image) or not. Subjects are instructed to press one key if the letter or number is a mirror image (reversed) and another key if it is not, regardless of orientation. A new letter or number appears immediately after each response. The order of presentation of stimuli is randomized with the constraint that each of the six stimuli appear twice in each of the six orientations and in both regular and mirror-image modes (6 stimuli x 6 orientations x 2 modes x 2 presentations), for a total of 144 stimuli presentations. Subjects are instructed to respond as quickly as possible while maintaining accuracy. The mental rotation task was included because previous research has shown a gender difference in performance, with men usually performing better than women (Mann et al., 1990). Measures include reaction time, the number of responses in the time allowed, and accuracy. (Task duration is about five to six minutes.)

Digit-symbol substitution task. The digit-symbol substitution task (DSST) is a self-paced task modeled after the digit-symbol test of the WAIS (Weschlser, 1958). Subjects are presented with a row of nine symbols in a random order across the top of the screen. Above each symbol is a number (1-9). These symbols and numbers remain present for the duration of the task. Below the row of symbols is a square in the center of the screen in which one symbol appears at a time. The subject's task is to enter the number corresponding to the symbol as quickly as possible. The sequential presentation of symbols continues for five minutes. The digit-symbol substitution task is included because previous research has demonstrated a gender difference in tasks of this type, with women generally showing superior performance (Mann et al., 1990; Smith, 1967). Measures include reaction time, number completed, and accuracy. (Task duration is five minutes.)

Simple reaction time task. The simple reaction time task (SRT) is a machine-paced vigilance task modeled after the Psychomotor Vigilance Task (PVT) of Dinges et al. (1994). Subjects are presented initially with an unfilled rectangle in the center of the screen. After a variable, randomly determined delay from 2-10 sec, the rectangle is filled with white. The subject's task is to respond as quickly as possible to the change in the rectangle by pressing the spacebar. The computer then displays the subject's reaction time (in msec) inside the rectangle. Measures include reaction time, lapses, and anticipations. (Task duration is 15 minutes.)

Compensatory tracking task. The compensatory tracking task (TRK) is a self-paced, two-dimensional compensatory tracking task in which the subject must keep a disk centered on a bullseye target on the screen (Makeig & Jolley, in preparation). The subject uses a trackball to control the disk and compensate for the direction changes in its movement in order to keep it positioned at the center of a ring. The disk is controlled by three forces. The first simulates the force of gravity on a "slippery hat" surface in which the screen center and edges are raised, leaving a circular trough that the disk will settle into in the absence of other forcing. The second force is a pseudo-random ("swirling wind") buffeting force summing six sine waves of incommensurate frequencies and amplitudes. The third force is user input, determined by the current velocity vector (direction, magnitude) of the trackball. Trackball and screen updates are computed approximately 11 times per second. The user must compensate for the effects of the first two forces, by generating an equal and opposite virtual force, to keep the disk centered on the ring. Without user input, the disk settles into a random orbit around the central ring with a mean distance from the origin of 9.4 disk radii. When user input ceases, the disk typically reaches this distance within 5 seconds. Most users can learn to keep the mean radial disk distance within 1.5 disk radii of screen center after 6-12 minutes of practice. The primary measure is radial distance of the disk from the target. (Task duration is 30 minutes.)

Subjective tasks. Two subjective tasks are presented as part of the performance battery to assess the mood, alertness, and physical well-being of the subjects. The Visual Analogue Scale (VAS) consists of a horizontal line drawn on the display with each end of the line labeled with the extremes of a subjective continuum. Four word or phrase pairs were used: sleepy/alert, physically exhausted/energetic, bored/engaged, sad/happy. The subject's task was to use the trackball to move an arrow cursor to the position on the line between the two endpoints that best described how he or she felt at that moment, and then press a response button. The dependent measure is the relative distance of the arrow from one end of the continuum which has a length of 600 pixels. Resolution is one pixel. (The task requires less than 30 seconds to complete.)

The Karolinska Sleepiness Scale (KSS) was incorporated into the battery to assess subject fatigue levels. The KSS is a standard 9-point fatigue scale that has been validated against VAS and EEG measures of sleepiness (Akerstedt & Gillberg, 1990). On this scale the numbers 1, 3, 5, 7, and 9 are associated with statements describing fatigue levels (e.g., 1 = "very alert," 9 = "very sleepy, great effort to keep awake, fighting sleep"). The nine numbered statements are presented simultaneously on the screen and the subject's task is to position the cursor next to the number that best describes how he or she feels at that moment and press a response button. (This task takes less than 30 seconds to complete.)

Cognitive Performance Test Analyses. Multivariate analyses of variance (MANOVAs), regression analyses, and graphical analyses were performed to identify effects of gender, sleep loss, and their interaction in the patterns of performance changes over the night without sleep. Further details of the analyses are provided in the Results and Discussion section.

EEG Analyses. Original plans called for the analysis of EEG data to determine the existence of any gender differences in the patterns and dynamics of EEG spectral changes associated with decrements in alertness and performance. These analyses might then lead to the derivation of algorithms that could predict individual performance. The characteristics and accuracy of the resulting prediction algorithms could then be compared between women and men. Technical problems in extracting the individual EEG files from the combined four-subject EEG recordings have prevented the completion of EEG data analyses in time for inclusion in this report. However, analysis routines have now been constructed to extract and pre-process the EEG data, and in-depth analyses of the EEG data are proceeding, with the results to be documented separately in technical reports. The present report contains a description of detailed analyses of the cognitive performance and subjective data, separate from the EEG correlates. The specific problems encountered in regard to the EEG data collection/analysis are described below.

Altogether, over 5 gigabytes (5×10^9) of EEG was collected from 45 subjects, along with detailed performance histories of each stimulus and response which were recorded in 1 megabyte ascii files by the task PCs, and later transferred to a Unix workstation network for analysis. The first step in EEG analysis, identifying the start and end of each task bout, and exactly matching the sync signal tracks in the EEG data to the performance history files, proved far more difficult and cumbersome than anticipated. Unanticipated problems included:

1. The Synamps "clear" function key not only cleared the screen of the computer running the EEG data collection, but also introduced a large and highly variable perturbation into the EEG and sync signal data itself. As this was a first experience using the Neuroscan unit for large-scale data collection, this problem was not identified until midway through the experiment. The problem of locating, excising and re-synchronizing the EEG data to the performance history had no simple algorithmic solution.

2. It was discovered that the Neuroscan unit sometimes produced perturbations or signal reversals without user input. These also needed to be identified by eye and corrected for.

3. For some tasks (including SRT), the rapid sequences of sync pulses marking task onset were not sufficiently different from the normally rapid sequence of stimulus onsets and subject responses, making it difficult for the custom analysis software to identify the onsets of these tasks without human inspection.

Faced with the task of decoding over 5 gigabytes of EEG information containing the complexities outlined above, an attempt was made to automate as much of the job as possible using a suite of Matlab software. As unanticipated problems with the data were uncovered, this software had to be revised and made more complex. Finally, a detailed script was written that runs a complex graphic user interface that can be used to visually review the data, inspect problem areas, and perform various correction options. (During this process, it was brought to our attention that other highly-regarded EEG scientists have had similar problems). The continuing FY97 efforts to separate the EEG data have been funded solely by an ONR-sponsored project. Over 400 hours of FY97 analyst time, plus many hours of co-investigator time have gone into this effort.

The task now is to use the suite of developed Matlab software to go through the collected data to produce synchronized EEG and performance records allowing the detailed time-frequency analyses outlined in the SOW. Unfortunately, as the (FY96-only) funding for the present project has expired, time must be borrowed from other tasks to continue this process.

It is expected that combined analysis of the EEG and performance data should be scientifically highly fruitful in at least two ways. First, the data may reveal whether the trends towards differences in performance and subjective measures for men and women subjects already reported are accompanied by quantifiable physiological differences. Second, the data are the first in our possession in which second-to-second fluctuations in performance linked to drowsiness and loss of vigilance have been observed at different times of the night. Therefore, the data provide a chance to test the accuracy of our published alertness monitoring algorithm during periods of sleep deprivation. Unfortunately, the budget for fatigue research has been cut significantly for FY99, making it impossible to budget sufficient analysis time for this project to carry out most of the possible analyses. In future years, efforts will be made to secure US Navy as well as other outside funding for this process.

Results and Discussion: NHRC Study

Training Sessions. Training session data were analyzed for the presence of a priori gender differences on the cognitive tasks and to determine whether asymptotic performance levels had been reached prior to the start of the experiment. Task results were visually inspected on computer during the training sessions and especially high error rates or reaction times by a subject led to additional sessions being performed until performance appeared to reach average levels. Additional sessions were required for less than 10% of the subjects. The training sessions were conducted on two different days. The first three training sessions occurred on the subjects' first visit to the laboratory as part of the orientation session. The second three training sessions occurred six to eight days later, just prior to the start of the formal experiment. The first set of training sessions were performed in the afternoon. The second set were performed between 1600 and 2000 hr. With a week between the two massed training sessions, one might expect initial learning, evidenced by improvements from sessions 1-2 and 2-3, and some relearning, evidenced by improvements from sessions 4-5 and 5-6. Any drops in performance or reductions in the slope of the learning curve between sessions three and four may be due to forgetting during the intervening week.

Mixed-model MANOVAs were performed on the training session data for each task. In the model, gender served as the between-subjects variable (two levels) while session was the repeated measure with six levels. Where significant gender by session differences were found, the Tukey HSD post hoc test was performed on gender groups at each session. Differences between adjacent sessions were also used as a post hoc test of session differences and gender by session interactions, as was a polynomial test for trends (up to a third order). Because the training data were analyzed primarily to determine whether sufficient learning had taken place prior to the start of the experiment, only mean response times and mean accuracy measures were examined for each of the tasks. The tasks will be described in the order in which they were presented to the subjects.

An examination of the continuous performance task reveals evidence of a learning effect that is largely dissipated by the final training session. Response times were initially between 0.6 and 0.65 sec but appeared to asymptote at about 0.5 sec before the start of the experiment (Figure 1). This improvement in response latency resulted in a significant effect of training session ($F_{5,46} = 19.99$, $p < .0001$). The large adjacent session differences indicate initial learning (large reductions in reaction time from session 1-2 and 2-3), a small amount of forgetting during the intervening week (small increase from session 3-4), and relearning (reduction from session 4-5). There was no significant difference between the genders, nor a gender by session interaction, despite consistently faster response times (by about 40 msec) for the women.

The practice effect, as measured by the percentage of correct responses on the same continuous performance task (Figure 2), is consistent with the pattern for reaction time. Again there is evidence of an initial learning effect up through session three but a subsequent leveling-off at 92-94% correct responses from sessions four through six. The initial learning resulted in a significant effect of session ($F_{5,46} = 4.03$, $p < .01$) but no effects of gender or gender by session.

The pattern for response times on the calculation performance task is shown in Figure 3. As with the continuous performance task, these data show evidence of most of the learning taking place in the first three sessions. For the present task there is evidence of forgetting occurring over the 6-8 days between sessions three and four for the men but not the women. By the last training session it appears that reaction time has achieved a stable value. These learning-related changes result in a significant gender by session interaction ($F_{5,46} = 2.37$, $p = .05$). The accuracy data on the calculation performance task (Figure 4) are consistent with the reaction time data. The women are fairly stable throughout the training sessions, averaging between 90% and

95%. The men show initial improvement through session three, a drop-off with the break in between sessions three and four, and further learning over the last three sessions. These changes are quite small, however, with the men also exhibiting variation between 90% and 95% correct responses. None of these changes resulted in any statistically significant effects.

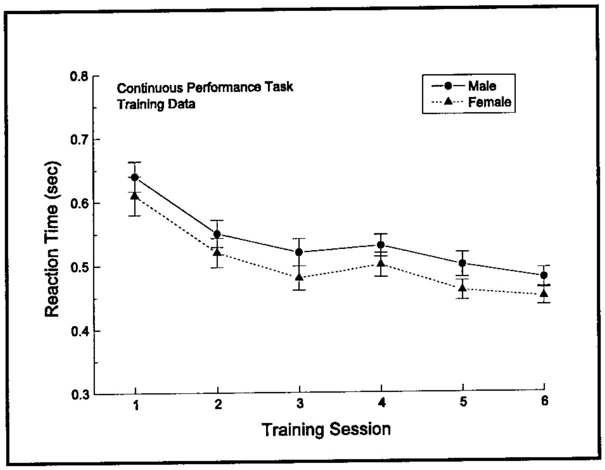


Figure 1: Mean response latency (sec) for the six training sessions on the continuous performance task. Error bars represent ± 1 sem.

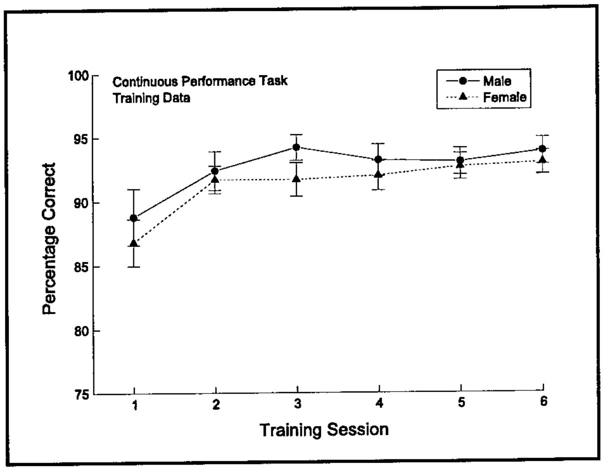


Figure 2: Mean percentage correct for the six training sessions on the continuous performance task. Error bars represent ± 1 sem.

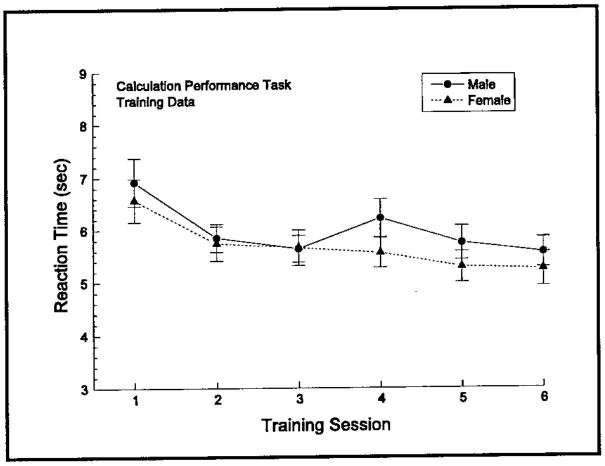


Figure 3: Mean response latency (sec) for the six training sessions on the calculation performance task. Error bars represent ± 1 sem.

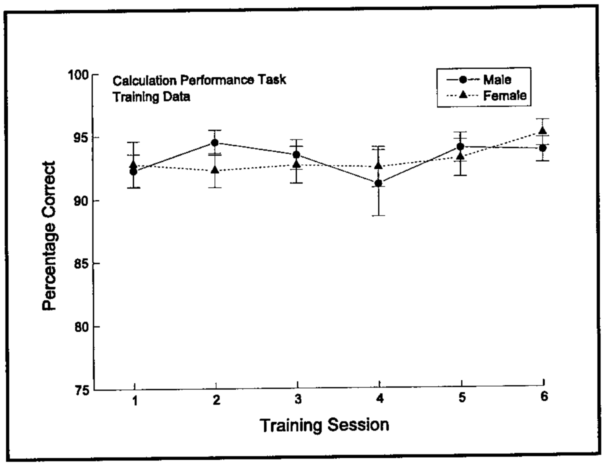


Figure 4: Mean percentage correct for the six training sessions on the calculation performance task. Error bars represent ± 1 sem.

Reaction times on the mental rotation task continue to decrease across all six training sessions but the bulk of the improvement occurred between sessions one and two with smaller improvements thereafter (Figure 5). The improvement across sessions was large enough to be statistically significant ($F_{5,45} = 18.40, p < .0001$). There was no statistically significant difference between genders and no gender by session interaction. The accuracy data for the MRT reveal that asymptotic accuracy levels have been reached by session three (Figure 6). This session effect is significant ($F_{5,45} = 5.14, p < .001$). The trend of better performance by men does not reach statistical significance ($p = .26$), nor is there a significant gender by session interaction.

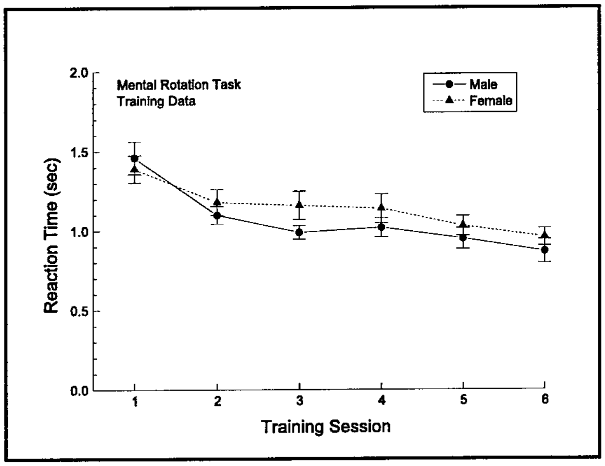


Figure 5: Mean response latency (sec) for the six training sessions on the mental rotation task. Error bars represent ± 1 sem.

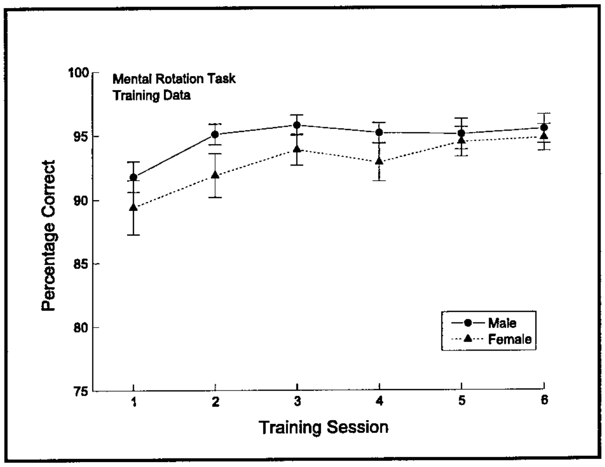


Figure 6: Mean percentage correct for the six training sessions on the mental rotation task. Error bars represent ± 1 sem.

Response times on the digit-symbol substitution task are shown in Figure 7. There is evidence of a slight tendency to respond more quickly with practice, as evidenced by a significant effect of session ($F_{5,45} = 2.79, p < .05$). However, examination of Figure 7 reveals that these changes were quite small. There were no differences between the genders and no gender by session interaction, although the men again showed a drop in performance after the break between sessions three and four. Analysis of changes in the percentage of correct responses during training on the DSST (Figure 8) reveals that there were no statistically significant effects of session, gender, or session x gender. The Tukey HSD reveals a significant drop from session one to two ($p < .05$) for unknown reasons. Nonetheless, the trend is for mean performance to hover at about 98% for all training sessions.

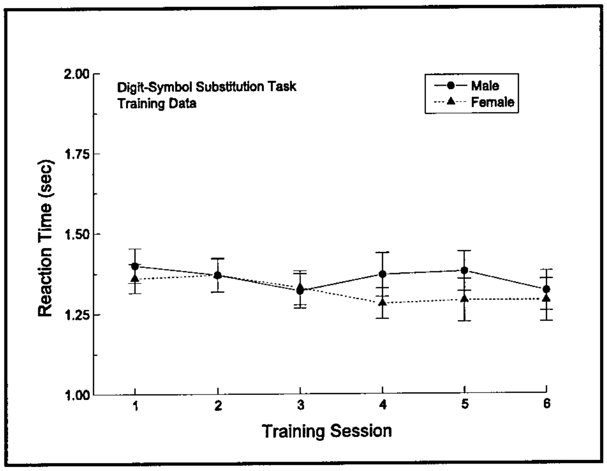


Figure 7: Mean response latency (sec) for the six training sessions on the digit-symbol substitution task. Error bars represent ± 1 sem.

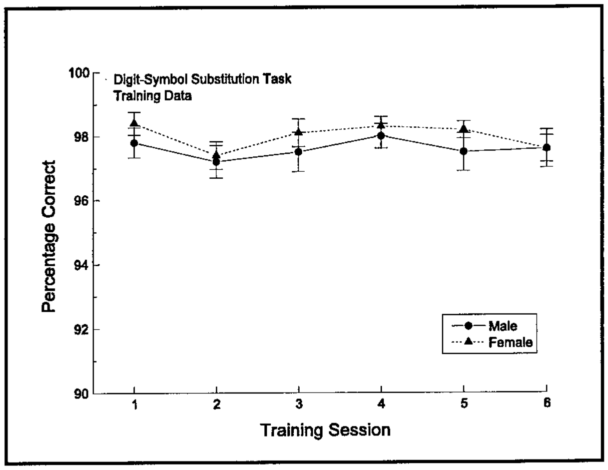


Figure 8: Mean percentage correct for the six training sessions on the digit-symbol substitution task. Error bars represent ± 1 sem.

Examination of mean reaction times in the simple reaction time task reveals little evidence of any practice effect across the six training sessions (Figure 9). Note, however, that this sample of males (mean of 280 msec) is consistently about 20 msec faster than the females (mean of 302 msec) for all the sessions, although this difference does not reach statistical significance. The slight trend toward slower responding with additional training may indicate the presence of boredom or monotony from repeated administrations of a task that subjects can successfully train on in probably just a few trails, as has been demonstrated for similar simple reaction time tasks (Dinges & Kribbs, 1991).

On the tracking task, training performance differed slightly between genders. "Best" training performance was better (p = .01) for men (normalized distance = .0991) than for women (normalized distance = .0996). (See the discussion of the TRK results during the experiment proper for a detailed description of this performance measure.)

Considered all together, the training data indicate that stable performance levels were attained for all six cognitive tasks. Most of the performance improvement occurred by the third session with near asymptotic levels achieved by the sixth session. Interestingly, while none of the differences between men and

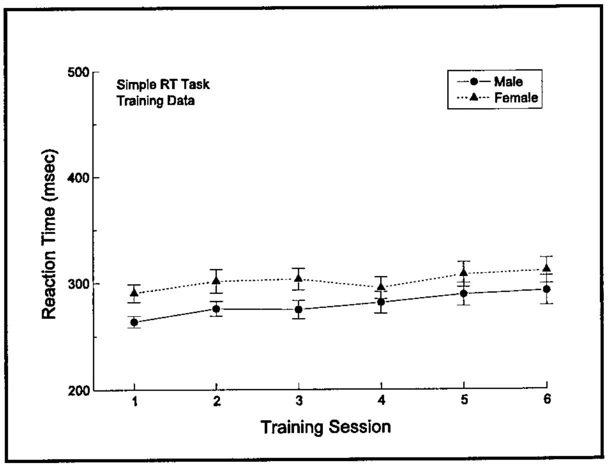


Figure 9: Mean response latency (msec) for the six training sessions on the simple reaction time task. Error bars represent ± 1 sem.

women reached statistical significance, the two groups did not perform identically. A number of the measures showed the two groups to have parallel yet slightly separated learning curves. For example, the women showed trends of being slightly faster but less accurate on the CPT, slightly faster and more accurate on the DSST, slightly slower and less accurate on the MRT, and slightly slower on the SRT. Again, these differences did not reach statistical significance, but were consistent across most of the six training sessions. Trends such as these might be expected on the MRT and DSST, tasks structured to favor men and women, respectively.

Test Sessions. Analyses for performance changes across the course of the night without sleep used the same mixed-model MANOVA as that used for the training session data. Again there were two levels of the between-subjects factor (gender) and six levels of the within-subjects factor (test session). Test sessions are labeled on all figures as one to six. The specific times at which these sessions started are provided in Table I. As mentioned previously, the number of subjects varied somewhat from task to task due to technical problems and these differences will be evident in the degrees of freedom associated with the F values reported below.

Simple reaction time task. Reaction time tasks, when analyzed using the proper metrics, are remarkably sensitive to the lapsing and other effects of sleep deprivation (Dinges, 1992; Dinges & Kribbs, 1991; Kribbs & Dinges, 1994). The SRT was the primary reaction time task in this study and will be examined first and in detail. We calculated five measures of performance (Dinges, 1992). First, reaction times were examined for evidence of any differences in cognitive slowing between men and women. Figure 10 shows median reaction times for the men and women at each test session over the night. (The median, rather than the mean, was chosen to describe the reaction time data in this and several of the other measures to avoid allowing relatively infrequent lapses [response times ≥ twice the mean value] to unduly influence the measure of central tendency.) There is a strong effect of session, with reaction times increasing significantly over the course of the night ($F_{5,45} = 25.34$, $p < .0001$). This trend had a significant linear component ($F_{1,49} = 133.76$, $p < .0001$). There was also a significant difference in degree of cognitive slowing between men and women, with men showing faster overall median reaction times ($F_{1,49} = 4.85$, $p < .05$). Tukey post-hoc tests revealed that the men showed significantly better performance than the women ($p < .05$) during the second, third, and fifth test sessions. The gender by session interaction was a borderline effect ($F_{4,45} = 2.32$, $p = .06$).

Second, the simple reaction time data were examined for gender differences in the optimum response domain -- that subset of responses that represents the best psychomotor effort. Following the procedure recommended by Dinges (1992), this optimum response domain was defined as the set of 10% fastest reaction times. The data are plotted in Figure 11 and show a very similar pattern to the complete set of responses shown in Figure 10. Once again there is a significant session effect, indicating generally poorer performance as the sleepless night progressed ($F_{5,45} = 20.69$, $p < .0001$). The trend of a slowing in the fastest reaction times is significantly linear ($F_{1,49} = 133.76$, $p < .0001$). Here the gender difference shows a strong trend toward significance, with men again showing faster optimum response times ($F_{1,49} = 3.36$, $p = .07$). The men were faster than the women at sessions two and three, according to the Tukey

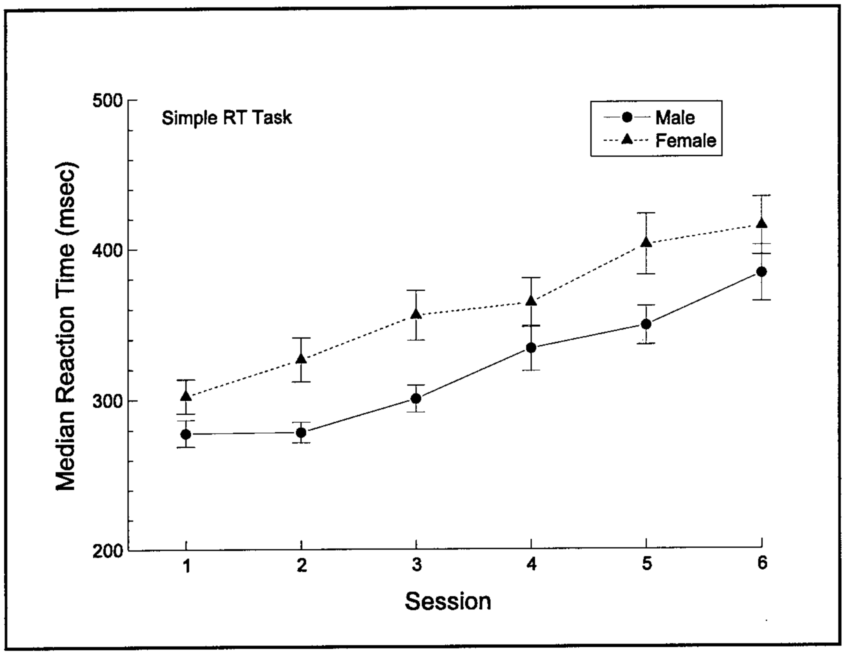


Figure 10: Median reaction time (msec) for the six experiment sessions on the simple reaction time task. Error bars represent ± 1 sem.

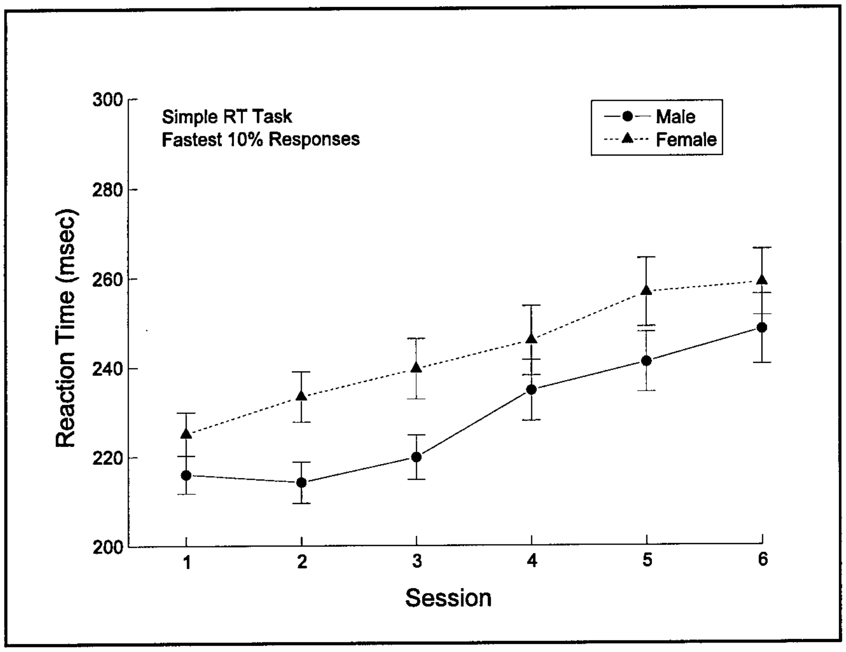


Figure 11: The mean fastest 10% reaction times (msec) for the six experiment sessions on the simple reaction time task. Error bars represent ± 1 sem.

HSD post-hoc test ($p < .05$). The gender by session interaction also showed a strong trend toward significance ($F_{4,45} = 2.37$, $p = .055$).

Third, reaction time data were examined for evidence of performance lapses, defined as a response with a latency greater than twice the mean latency. Response times during training sessions were used to determine the average latency values. Recall that there was no evidence of a practice effect during training sessions (Figure 9) and so values were included from all six sessions in determining the mean. The training data showed that the grand mean reaction time was 292 msec (280 msec for men and 302 msec for women). Consequently, 600 msec (approximately twice the mean) was selected as the lapse criterion value. The number of responses per test session that exceeded this value are plotted in Figure 12. Before submitting the data to statistical analysis the number of lapses per test session (x) were transformed using the formula $(x)^{1/2} + (x + 1)^{1/2}$ to reduce the proportionality between the session means and variances. This proportionality is found for frequency data (such as lapses) with a small probability of occurrence (Kirk, 1982). There is clearly an effect of sleep loss, evidenced by a significant increase in lapsing over the course of the night (significant effect of session: $F_{5,46} = 40.00$, $p < .0001$). The trend was linear and statistically significant ($F_{1,50} = 206.39$, $p < .0001$). Lapses averaged less than 5 during the first two sessions of the night but increased monotonically and exceeded an average of 20 per session by the end of the night of sleep deprivation. There were no differences in lapse frequency between the first two sessions, but by the third session the number of lapses in each session was significantly greater than the number in the first. Furthermore, the magnitude of the increase was such that starting with the third session the lapse frequency for each session was significantly higher than all the preceding sessions. However, the four-fold increase in lapsing across the course of the night was similar for men and women. Although the women exhibited an average of about four more lapses than the men during every test bout, this consistent gender difference was not large enough to reach statistical significance ($p = .27$). Nor was there a significant gender by session interaction.

Any additional evidence of gender differences in the speed of responding was sought by examining the "lapse domain," or the slowest 10% of the responses for each subject during each test session (Dinges & Kribbs, 1991). The mean 10% slowest reaction times are shown in Figure 13. As with the lapse frequency data, the reaction time values were transformed before submitting them to statistical analysis. In this case a simple reciprocal transform (1/RT) was used in accordance with the recommendation of Kirk (1982) for situations where the dependent variable is response time. The typical pattern emerged, with an overall increase in the very slowest response times with increasing duration of wakefulness through the night. This session effect was statistically significant ($F_{5,45} = 40.81$, $p < .0001$). Once again the linear trend was the primary one ($F_{1,49} = 202.08$, $p < .0001$). The men again showed faster reaction times than the women but the difference failed to reach statistical significance for this measure ($p = .18$). There was no significant gender by session interaction.

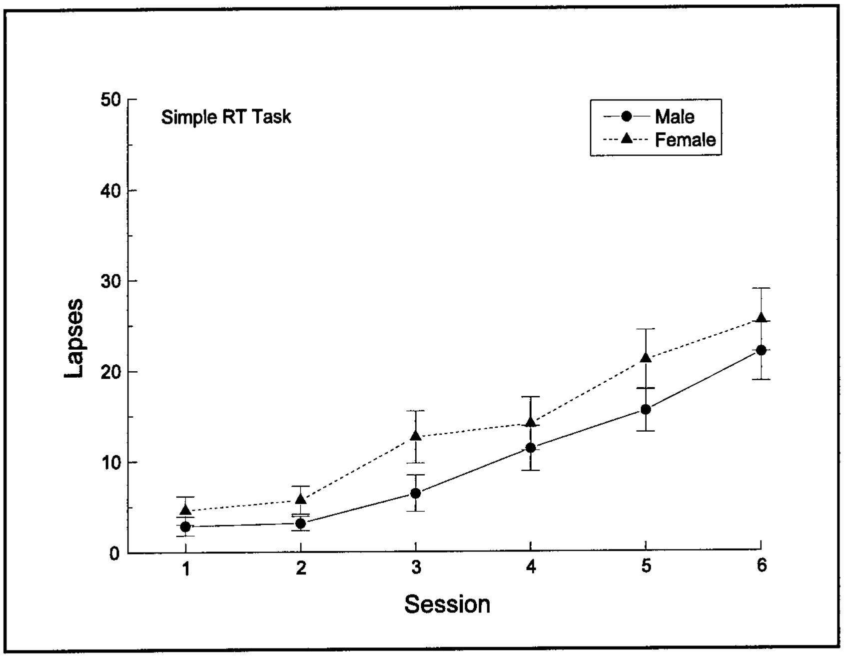


Figure 12: Mean number of lapses for the six experiment sessions on the simple reaction time task. Error bars represent ± 1 sem.

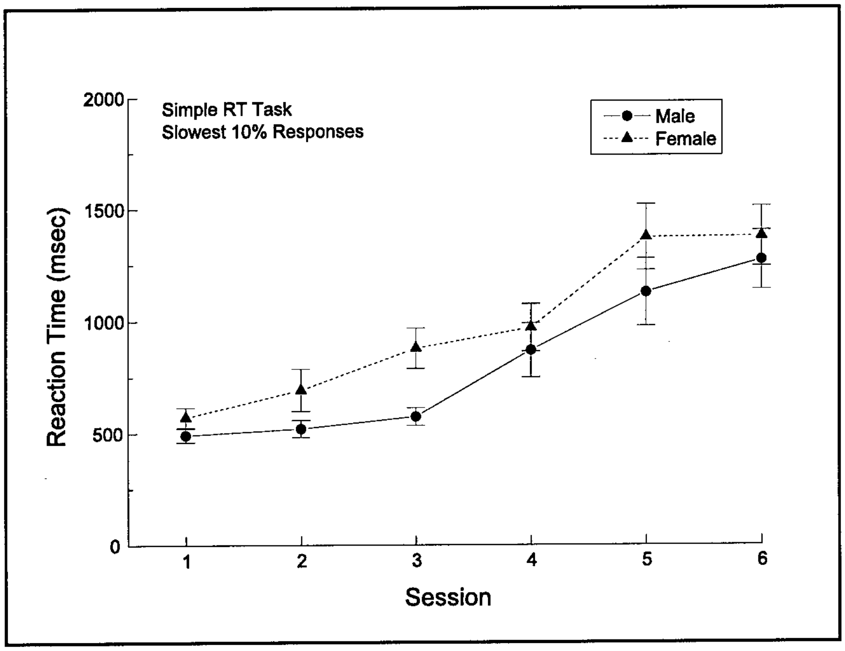


Figure 13: The mean slowest 10% reaction times (msec) for the six experiment sessions on the simple reaction time task. Error bars represent ± 1 sem.

Finally, anticipations (false starts) were analyzed for gender differences. Anticipations are premature responses, occurring before the stimulus appears on the screen. Dinges and Powell (1989) showed that, with sleep loss, anticipations actually tend to increase rather than decrease. Careful inspection of these data revealed several subject sessions with a high number of consecutive anticipatory responses, indicative of an attempt by a subject to avoid performance on the task by making multiple keypresses in a row, rather than responding strictly to individual stimuli. In order to prevent this action from contaminating the data, only nonconsecutive anticipations were included in this analysis. These anticipations are plotted in Figure 14. The data were subjected to the same square-root transform $[(x)^{1/2} + (x + 1)^{1/2}]$ as the lapse data before being submitted to statistical analysis. Increasing time without sleep clearly resulted in increasing anticipatory responses, as evidenced by the effect of session ($F_{5,47} = 19.38$, $p < .0001$). Again the linear trend was highly significant ($p < .0001$), with anticipations starting at less than four in the first fifteen-min session but increasing two- to three-fold by the end of the night. Significant increases in anticipatory responses occurred from session two to three ($p < .0001$) and five to six ($p < .001$). Both men and women show a tendency to make more false responses over the course of the sleep deprivation period but with a nonsignificant trend toward more false responses by the women than the men that became exacerbated as the night progressed. However, the gender by session interaction failed to reach statistical significance.

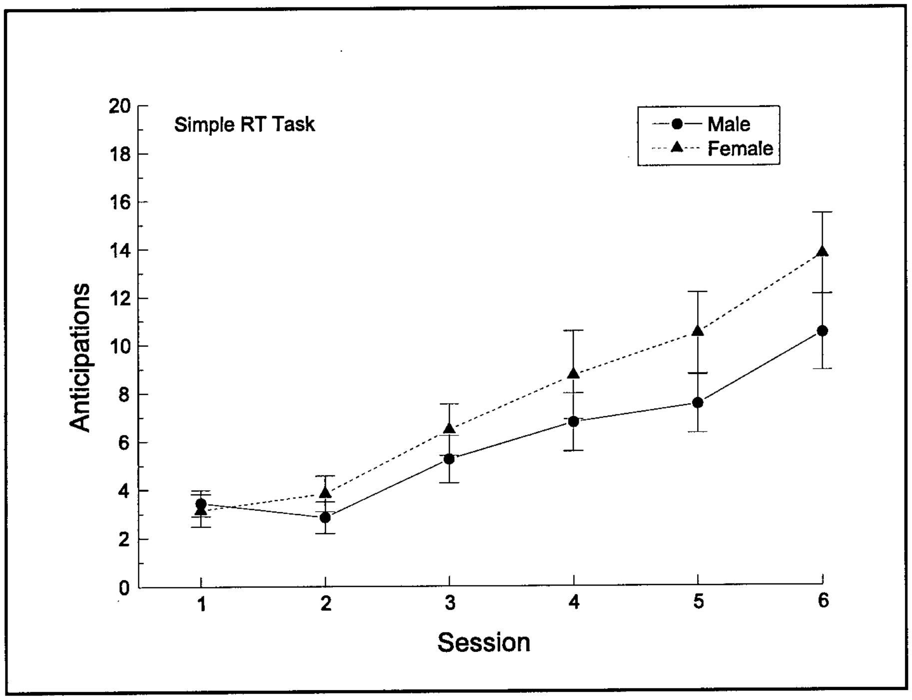


Figure 14: Mean number of anticipations (false starts) for the six experiment sessions on the simple reaction time task. Error bars represent ± 1 sem.

In sum, there were clear and consistent effects of sleep deprivation for the simple reaction time task. Both men and women showed significantly degraded performance over the course of the night on all five SRT measures. Examination of the data in Figures 10 through 13 shows that the men consistently outperformed the women but the difference between the genders reached statistical significance only for median reaction time, although it was borderline ($p = .07$) for the optimum response domain. The trend toward degraded performance through the night was similar for the men and women, as evidenced by the lack of any significant gender by session interactions (and the fairly parallel lines in the figures), although the interaction approached significance for the median response times and optimum response domain. Thus, while the men tended to perform better, there was no significant difference in the effect of sleep loss between the two groups. Furthermore, the trend toward better performance by the men is a carry-over from the training sessions (Figure 9). In other words, the borderline difference in reaction time between men and women was present before the start of the experiment and remained fairly consistent throughout the experiment.

Continuous performance task. The other task for which reaction time was a primary component was the CPT. While sharing many characteristics with the SRT, the CPT also has important differences. Among them is the fact that the CPT has letter recognition and choice components to it with subjects asked to make a rapid same-different judgment before responding. The CPT was included primarily because its fixed stimulus presentation rate allows better comparisons to the EEG data than the SRT. Nonetheless, it is still informative to examine two CPT measures for comparison with the SRT. Figure 15 shows the trend toward increasing mean reaction times with successive test sessions as sleep loss accumulates. (Mean reaction time was chosen for the CPT, rather than median, because the relatively brief interstimulus interval of 2 sec limits the duration of "lapses" or long response times.) The effect of session on reaction time is highly significant even though the total increase is about 0.1 sec ($F_{5,50} = 18.99$, $p < .0001$). The primary trend is one of a linear increase in reaction time with session ($F_{1,54} = 82.09$, $p < .0001$). Unlike for the SRT, the gender difference this time favors women, who are consistently responding a little faster than the men, although this difference does not reach significance ($p = .21$), nor does the gender by session interaction. However, the difference is also not an effect of sleep loss because it is present right from the start of the experiment, in the first two to three test sessions, before much sleep loss has accumulated.

The accuracy results on the CPT (Figure 16) show the women starting at the same level as the men during the first test session. While performance degrades over the course of the night for both groups, resulting in a significant session effect ($F_{5,50} = 14.97$, $p < .0001$), the women's performance drops off at a slightly faster rate than that for the men. However, both the gender by session interaction and the main effect of gender failed to reach significance. The trend in performance degradation was primarily linear ($F_{1,54} = 4.40$, $p < .05$).

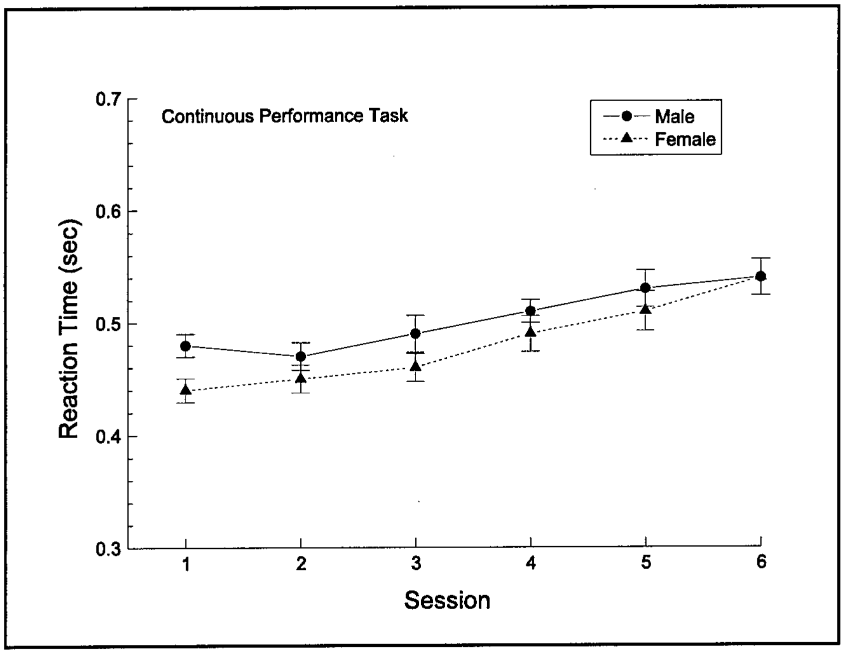


Figure 15: Mean reaction time (sec) for the six experiment sessions on the continuous performance task. Error bars represent ± 1 sem.

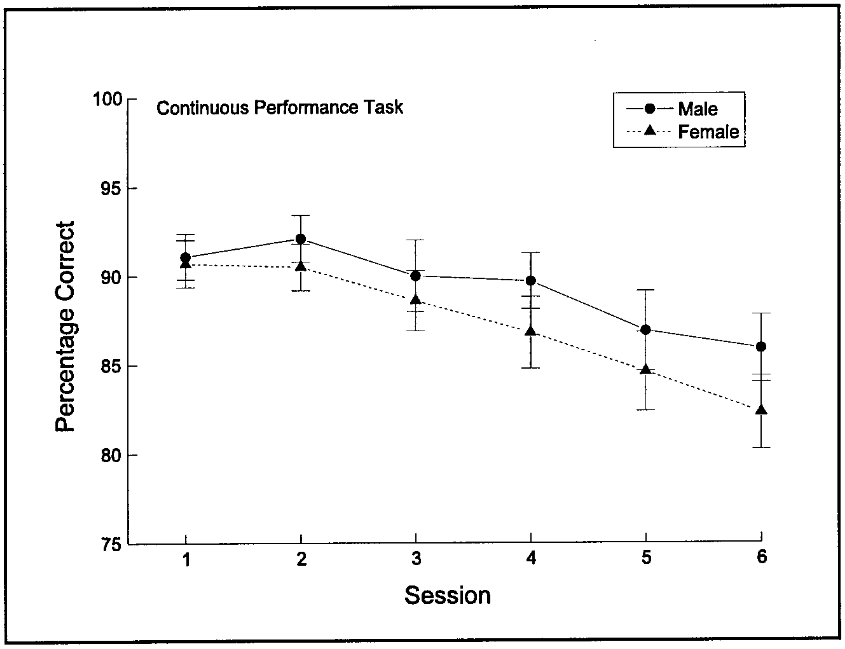


Figure 16: Mean percentage correct for the six experiment sessions on the continuous performance task. Error bars represent ± 1 sem.

As with the SRT, one night of sleep loss appeared to affect men and women similarly, resulting in significantly slower response times and reduced accuracy. While women showed a nonsignificant pattern of responding faster than the men while being slightly less accurate, this difference was present from the early sessions and was therefore not due to lack of sleep. As with the SRT, these very same nonsignificant differences were also present in the training session data (Figures 1 & 2), again indicating the presence of group differences in responding that existed prior to the start of the experiment and also demonstrating the reliability of the task in consistently detecting them.

Mental rotation task. The MRT and the next task to be described, the DSST, were included because the literature indicates that men usually perform better than women on the MRT and worse than women on tasks such as the DSST (Mann et al., 1990; Smith, 1967). The mental rotation task data were examined for the effects of degree of stimulus rotation as well as the effects of gender and session and their various interactions. Therefore the data were analyzed using a mixed model repeated-measures MANOVA with gender as the between-subjects variable (two levels) and both rotation (six levels: 0°, 60°, 120°, 180°, 240°, and 300°) and session (six levels) as the within-subjects variables. The mean reaction times for the men and women in all six test sessions are shown in Figure 17. Overall, the men averaged 1.03 sec and the women 1.17 sec, a difference that failed to reach statistical significance. The women's performance degraded relatively more than that of the men (and became much more variable) in the late night/early morning hours, during sessions five and six. However, this divergence in the performance curves was not large enough to result in a significant gender by session interaction. However, there was a significant effect of session ($F_{5,47} = 4.75$, $p < .001$). The only other significant finding was an effect of rotation ($F_{5,47} = 23.01$, $p < .0001$).

The effect of rotation has been found with similar versions of this task (e.g., Cooper, 1975) and this effect, along with the effect of sleep loss, can be seen clearly in Figures 18 and 19. Figure 18 reveals how the task becomes more difficult, resulting in longer reaction times, as stimulus rotation varies from 0°. Longest reaction times are associated with 180° (complete inversion), followed by 120° and 240°, which are similar. Those rotations (60° and 300°) closest to the unrotated stimulus (0°) did not result in response times very different from that for 0°. The increases in reaction time from 60° to 180° and from 300° to 180° are clearly linear, a finding that is consistent with the literature on this task. Poorer performance during the later sessions (five and six) can also be clearly seen. Figure 19 is a different plot of the same information. Here the data clearly show the effect of session and the lack of any interaction with rotation. Again, however, there were no differences between men and women in their response times on this task.

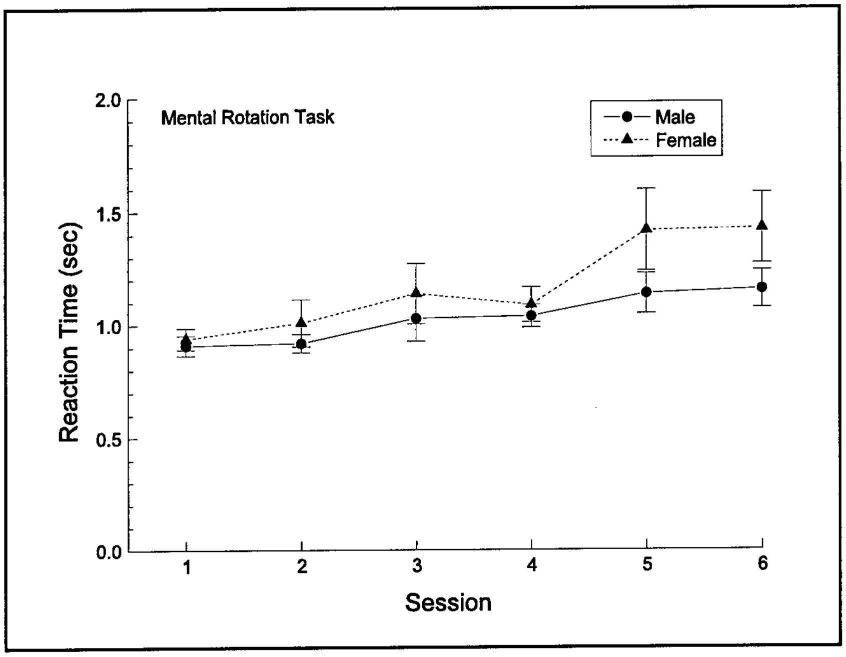


Figure 17: Mean reaction time (sec) for the six experiment sessions on the mental rotation task. Error bars represent ± 1 sem.

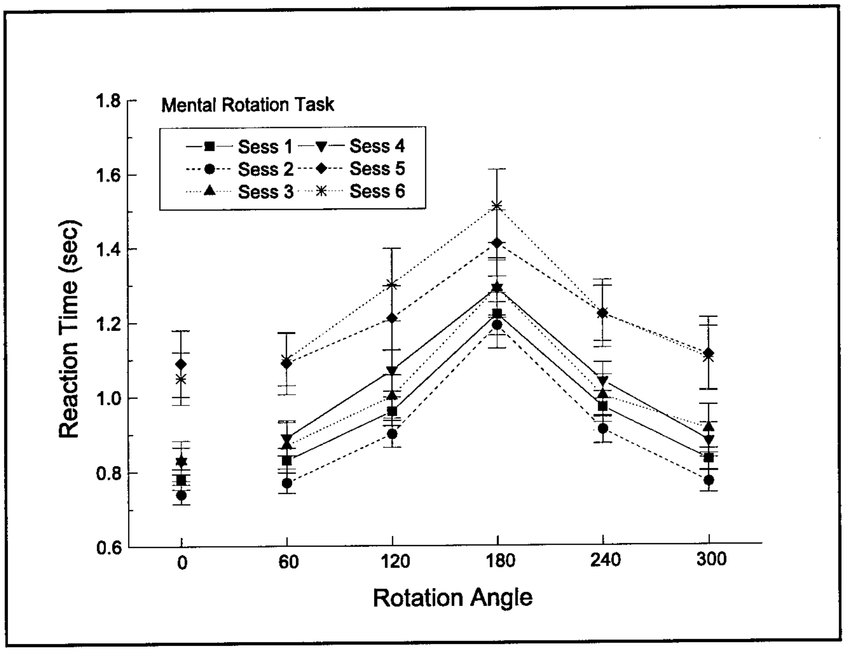


Figure 18: Mean reaction time (sec) on the mental rotation task for the six experiment sessions and the six stimulus rotations. Error bars represent ± 1 sem.

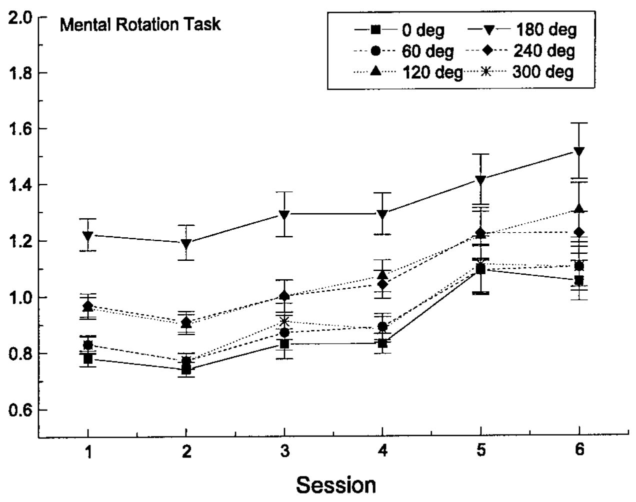


Figure 19: Mean reaction time (sec) on the mental rotation task for the six stimulus rotations and the six experiment sessions. Error bars represent ± 1 sem.

Accuracy data for the MRT are shown in Figure 20. The effect of test session is significant ($F_{5,50}$ = 4.65, $p < .01$) but neither the effect of gender nor the gender by session interaction reach significance. Nevertheless, with sleep loss, the women once again show a trend toward poorer and more variable performance on the MRT than the men.

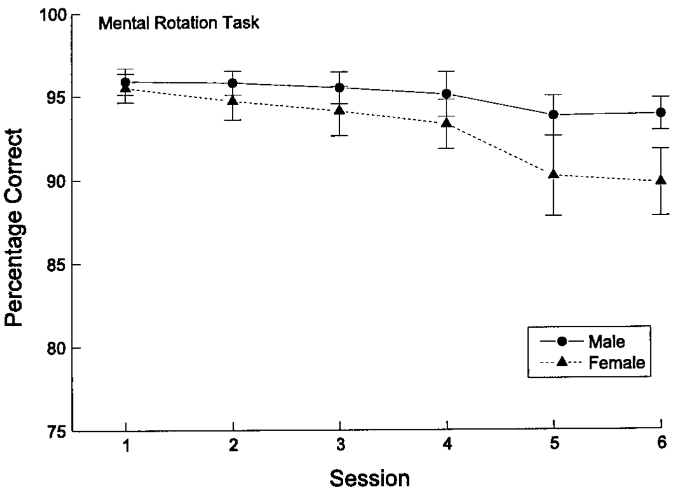


Figure 20: Mean percentage correct for the six experiment sessions on the mental rotation task. Error bars represent ± 1 sem.

As with the reaction time data on this task, the accuracy data can be plotted so as to clearly see effects of stimulus rotation as well as sleep loss. Also, because there were a fixed number of stimuli presented to each subject, the number of correct responses can be used as a simpler measure of accuracy. The MRT data were analyzed for cognitive throughput differences between the genders by submitting the number of correct responses per trial to the same 3-way MANOVA described previously. As with reaction time, there was no effect of gender. The only significant effects were session ($F_{5,47} = 4.77$, $p < .01$) and rotation ($F_{5,47} = 9.72$, $p < .0001$), with none of the interactions reaching significance. These effects are plotted in Figures 21 and 22. One can see the effects of both session and rotation in Figure 21, with inverted stimuli (180°) resulting in a fewer number of correct responses. The results are not exactly symmetrical with respect to degree of rotation, in that fewer correct responses were made (particularly in the later sessions) at 240° than at 120°. It was at the 180°, 240°, and 300° rotations that the session effect is most noticeable. However, this session by rotation interaction did not reach statistical significance. Figure 22 is a different plot revealing the same effects of session and rotation, again showing the difficulties that subjects had with the inverted stimuli and the significantly poorer performance as sleep loss accumulated. Yet the changes in the absolute number of correct responses show that performance degraded by only a few additional errors from the beginning to the end of the night.

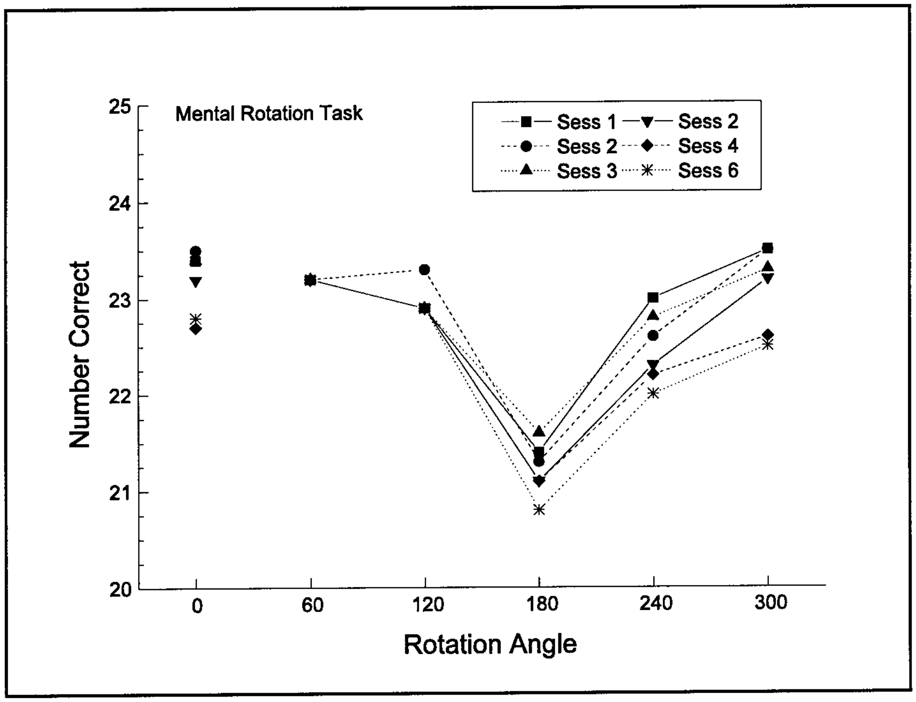


Figure 21: Mean number of correct responses on the mental rotation task for the six experiment sessions and the six stimulus rotations.

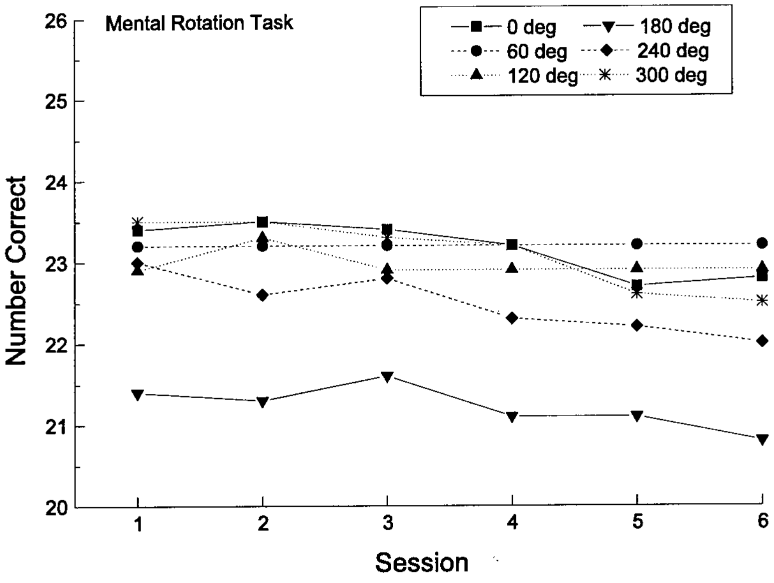


Figure 22: Mean number of correct responses on the mental rotation task for the six stimulus rotations and the six experiment sessions.

While the pattern of performance as a function of stimulus rotation angle (Figures 18, 19, 21, and 22) is what one would predict from the literature, the differences between men and women, while suggestive, do not reach statistical significance. Figures 17 and 20 reveal that, while men and women start out with virtually identical response times and accuracy levels on this task, men maintain higher levels of performance on both measures throughout the period of sleep deprivation. The slight advantage that the men showed during the training period (Figures 5 & 6) was not carried into the experiment on this task as evidenced by the nearly identical performances for the two groups during the first experimental test session. It appears that, rather than starting with any significant advantage relative to women, men may be more resistant to the effects of sleep loss on this task that has traditionally favored men. It should be remembered, however, that these results are only trends in the data.

Digit-symbol substitution task. The DSST was chosen as a task likely to result in better performance by women. Median reaction time results are plotted in Figure 23. Consistent with findings from the other tasks with reaction time measures, both groups slowed down as the night progressed, resulting in an effect of test session ($F_{5,47} = 12.64$, $p < .0001$). The increases in median reaction time from sessions two to three and four to five were statistically significant ($p < .001$ and $p < .01$, respectively). As expected, the women showed consistently faster mean reaction times but the difference was small (~ 0.1 sec) and did not reach statistical significance (p = .37). There was no gender by session interaction. The drop in response time was accompanied by a reduction of almost 10% in the number of trials completed over the course of the night (Figure 24), providing additional evidence of cognitive slowing. This effect of sleep loss on trial completions was significant ($F_{5,48} = 17.57$, $p < .0001$). There were no significant effects of gender or gender by session despite the trend for the women to very slightly outperform the men, particularly early in the experiment. The tendency by both groups to respond more slowly and complete fewer trials was not accompanied by any statistically significant changes in the percentage of correct responses (Figure 25), indicating the presence of speed-accuracy trading. In other words, both groups of subjects tended to slow down in order to maintain accuracy -- a strategy often used by sleep-deprived individuals. The small advantages shown by the women on this task during training appeared to be maintained throughout the night in the form of faster response times and more trials completed. As with the MRT, these differences were not statistically significant but represent only trends.

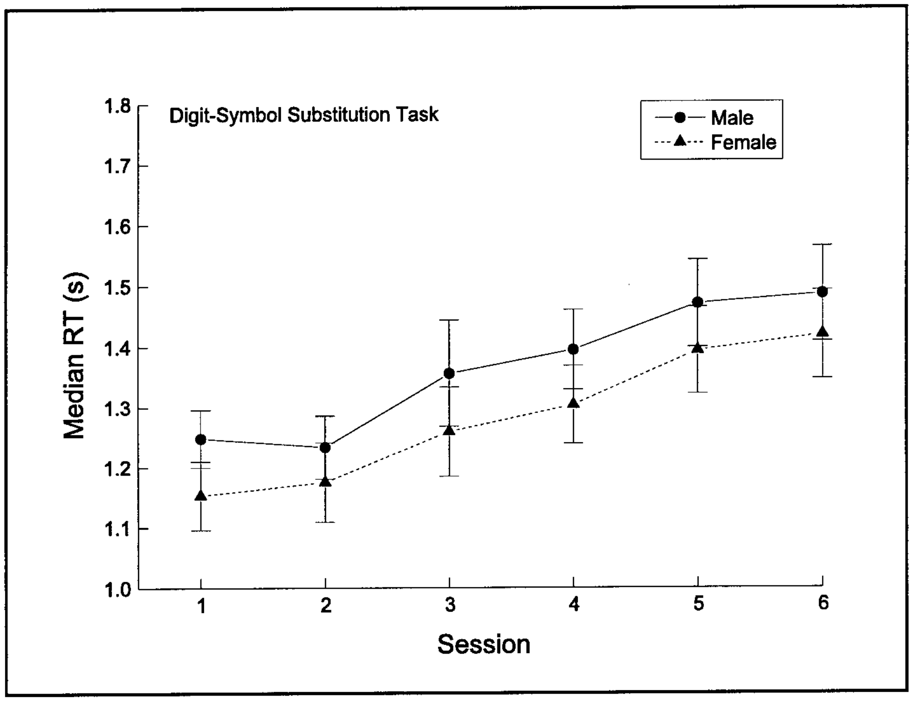


Figure 23: Median reaction time (s) on the digit-symbol substitution task for the six experiment sessions. Error bars represent ± 1 sem.

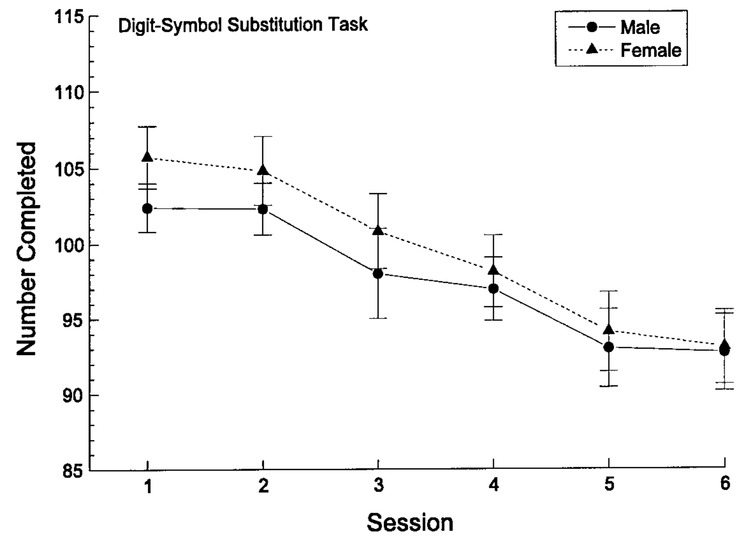


Figure 24: Number of completed trials on the digit-symbol substitution task for the six experiment sessions. Error bars represent ± 1 sem.

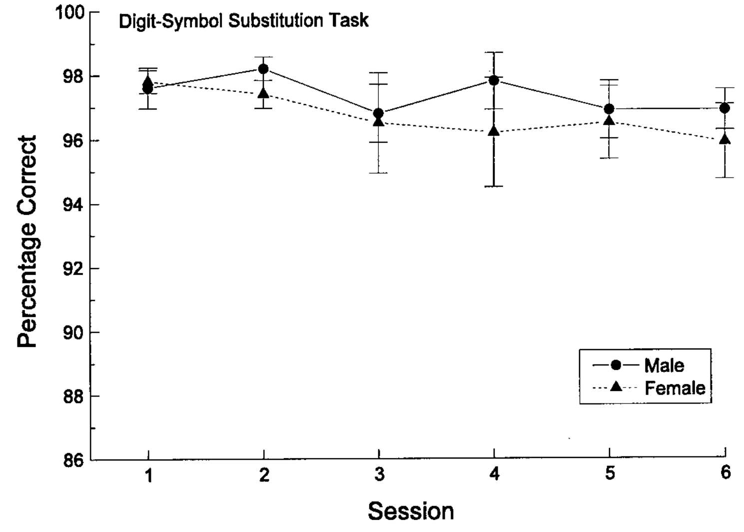


Figure 25: Percentage correct responses on the digit-symbol substitution task for the six experiment sessions. Error bars represent ± 1 sem.

Calculation performance task. Median response times for the CMT are plotted in Figure 26. There is a significant effect of test session ($F_{5,45} = 9.47$, $p < .0001$) with the primary trend again being one of a linear increase in the time to respond ($F_{1,49} = 30.38$, $p < .0001$). The only significant increase in reaction time between successive sessions is the increase from sessions four to five. The trend of faster responding by the women does not reach significance ($p = .15$). The completed trials data (Figure 27) show a similar result with a significant reduction in completed trials over the course of the night ($F_{5,50} = 8.46$, $p < .0001$) but no significant differences between the groups ($p = .21$) and no group by session interaction. The primary trend was the linear decline in completed trials ($F_{1,54} = 31.97$, $p < .0001$). As with the DSST, these results show evidence of cognitive slowing with a nonsignificant advantage for the women in terms of faster responding and more problems completed. The practice data show no real advantage for the women in response time by the last training session (Figure 3). Similar to the DSST, the percentage of correct responses did not change with time nor differ between men and women (Figure 28), further evidence for the presence of a speed-accuracy tradeoff. In sum, performance on the CMT parallels that for the DSST, not an unexpected finding for two similarly structured self-paced tasks.

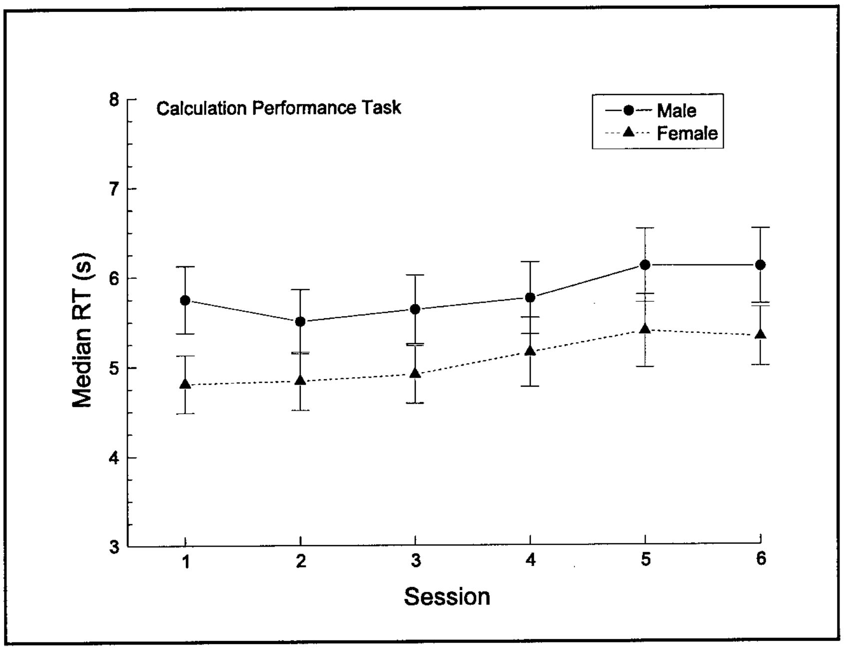


Figure 26: Median reaction time (s) on the calculation performance task for the six experiment sessions. Error bars represent ± 1 sem.

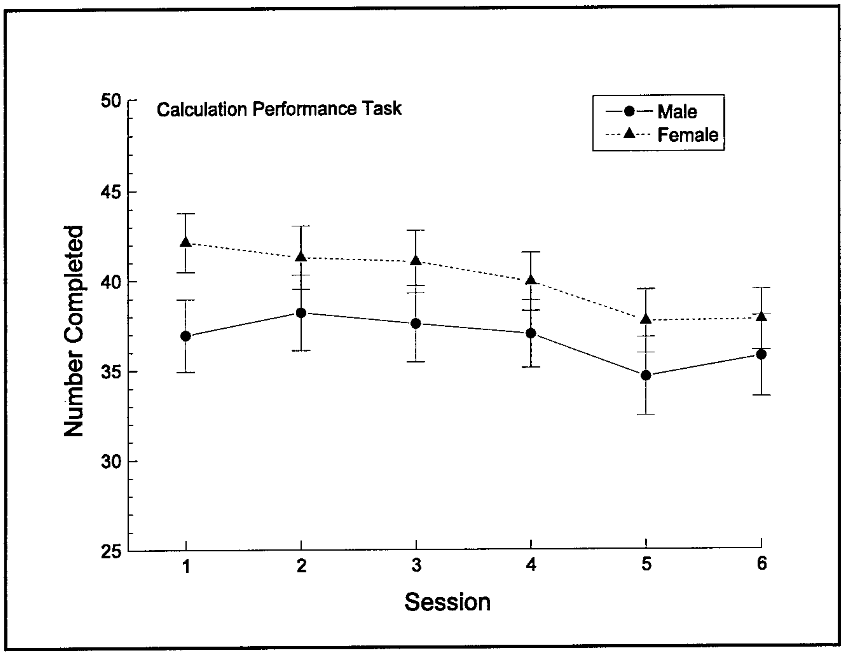


Figure 27: Number of completed trials on the calculation performance task for the six experiment sessions. Error bars represent ± 1 sem.

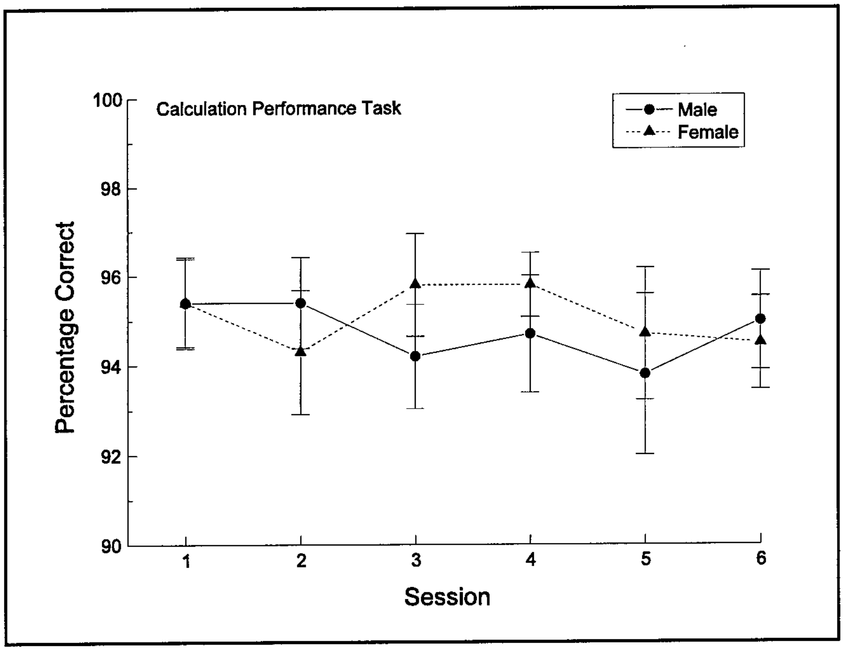


Figure 28: Mean percentage correct responses on the calculation performance task for the six experiment sessions. Error bars represent ± 1 sem.

Compensatory tracking task. Performance during each 30-minute TRK session was quantified by averaging the radial distance from the origin of the disk in a one-minute moving window which was advanced through the session in 1.648-sec steps. The first smoothing window was centered on time 0; the last was centered on 30 minutes, giving a total of 1025 time windows in each session. The smoothed performance measures were normalized by passing them through a logistic sigmoid function whose maximum and minimum were mean radial distance in zero-input (asleep) simulations and a standardized "best" performance (1.0) baseline. The normalized distance measures varied between 0.0 (alert) and 1.0 (asleep/random). Figure 29 shows the time course of this TRK performance measure in the six sessions for a representative subject. Grand means of the normalized smoothed distance measures were then computed for males and females. "Best" training performance was measured for each subject as the tenth-percentile of mean distance of the disk from the origin after one-minute smoothing and sigmoidal correction. Significance levels were computed by analysis of variance on data from the 52 subjects (24 males, 28 females) from whom TRK performance data were available for all six sessions.

Mean normalized distance on the tracking task is plotted separately for men and women over the course of the night in Figure 30. Across the six test sessions, overall average performance strongly decreased, from .11 to .58 ($F_{5,250} = 96.42$, $p < .0001$). This decrease amounted to about a two-diameter increase in the mean distance of the disk from the center of the target ring. Overall, performance of men (.28) was considerably better than that of women (.43; $F_{1,50} = 11.75$, $p = .001$). This difference was significant in each of the six sessions ($p \leq .026$). The gender by session interaction across the six test sessions was not significant ($p = .10$), but showed a trend toward greater performance declines in women after session one. The gender group difference was largest during sessions three and four (.22 and .19, respectively) and smallest in session one (.06). Analysis of variance on data from sessions one through four showed only a significant gender by session interaction ($p = .019$). The effect of gender was also significant in data from session one only ($p = .019$). In data from the first ten minutes of session one only this gender difference was smaller (.060 men versus .099 women), but still significant ($p < .05$). Combined analysis of "best" training and test session performance also indicated a significant gender by session difference ($p = .001$). In each of the six sessions, there was greater performance variance between women than between men.

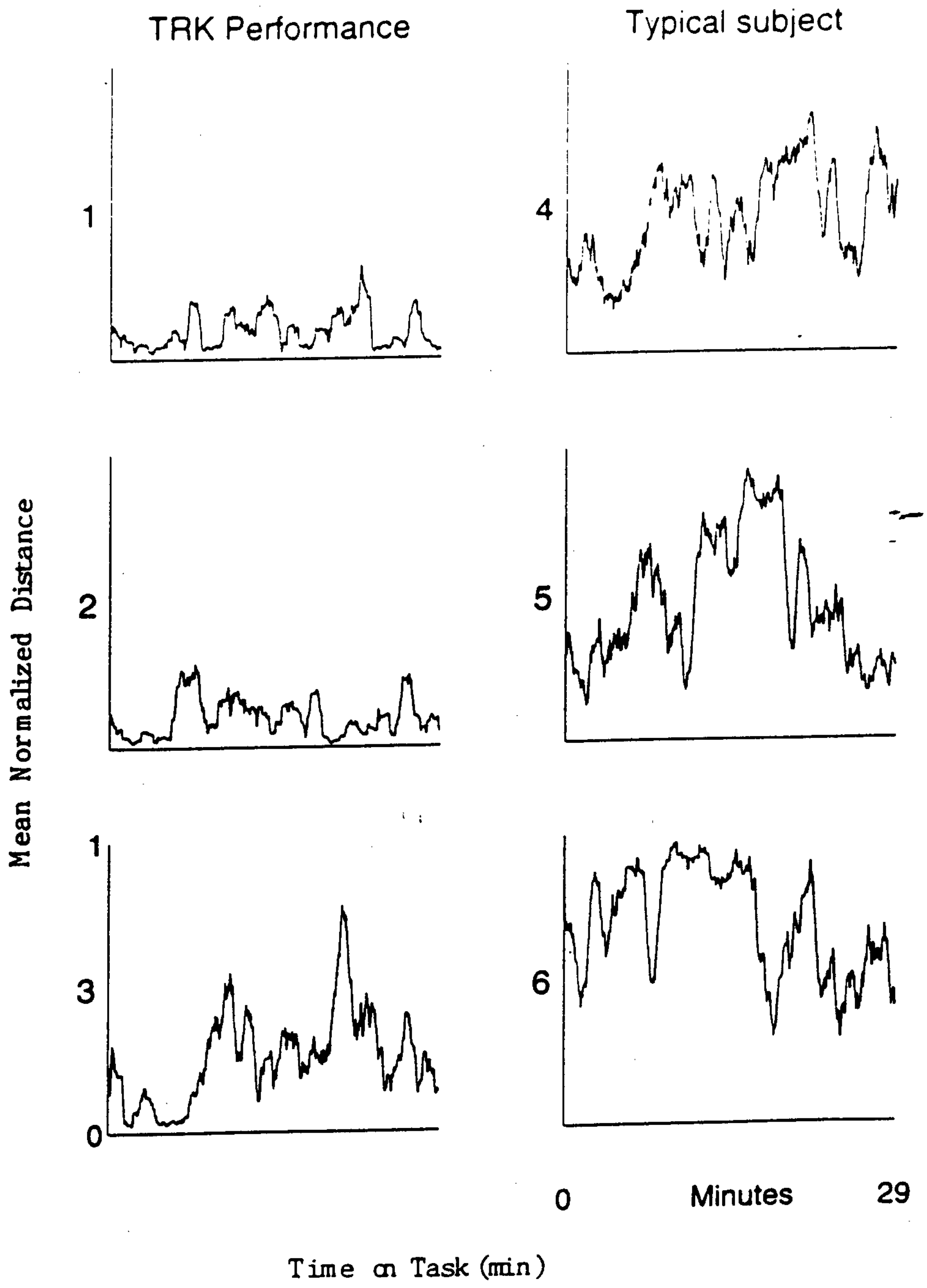


Figure 29: Mean tracking performance during the six test sessions by a typical subject. Performance is measured as mean distance of the disk from the center of the target ring, smoothed using a one-minute moving window and passed through a sigmoidal correction to account for the relatively large random fluctuations of mean distance in the absence of subject input. The normalized distance measure varies from 0.0 (optimum performance) to 1.0 (no subject input).

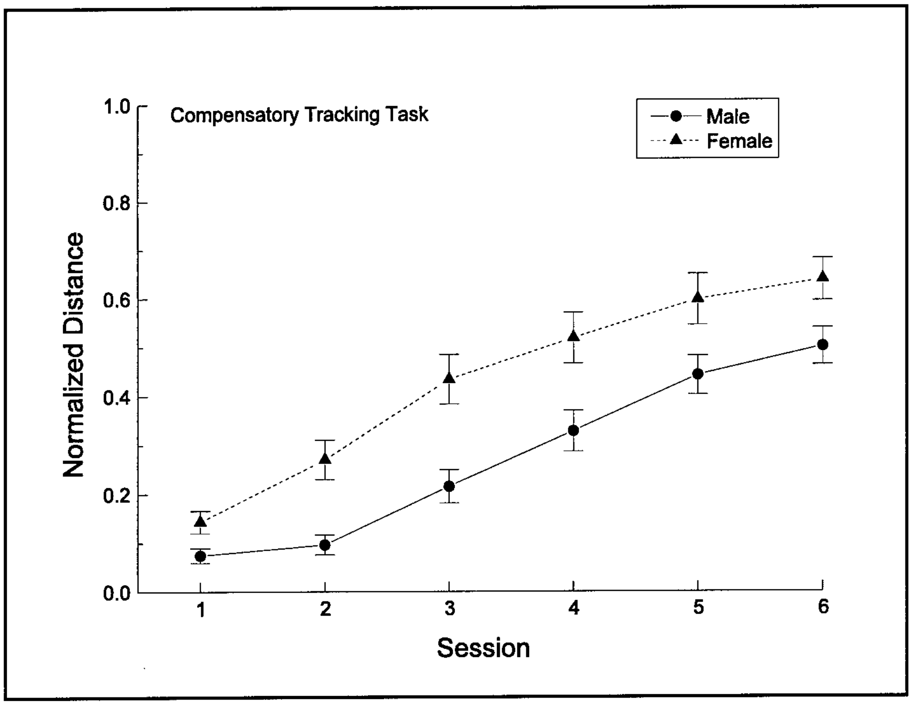


Figure 30: Mean normalized radial distance on the compensatory tracking task for the six experiment sessions. Error bars represent ± 1 sem.

To summarize, the data indicate that, overall, tracking performance was better in men and showed a trend toward declining more in women than in men. The variance between women was greater than the variance between men. Of the top eight performers in session six, four were women and four men, while all 11 poorest performers in session six were women. The tracking data suggest that performance in monotonous and demanding visual tracking may differ in men and women across the course of one night of sleep deprivation. However, the reason for this difference is not clear, nor is the cause of the larger performance variance in women. It is possible that the observed gender differences may have been caused from differences in group motivation, although the experimental protocol contains no obvious source of such a difference. If tracking performance of men and women does indeed differ following mild sleep deprivation, this fact might be considered in recruitment into monitoring occupations.

Subjective tasks. The responses on the Karolinska Sleepiness Scale (Figure 31) were initially near the midpoint of the scale during session one ("Neither alert nor sleepy") and declined significantly over subsequent test sessions ($F_{5,48} = 46.45$, $p < .0001$) with a very strong linear trend ($F_{1,52} = 152.94$, $p < .0001$). With each successive test session there was a significant increase in reported sleepiness compared to the previous one ($p < .001$), with the sole exception of session six. There was a trend of the men reporting less sleepiness than the women ($p = .12$), particularly during the first four sessions. Both groups reported being quite sleepy during sessions four and five. The gender by session interaction was not significant ($p = .17$).

The Sleepy-Alert dimension of the VAS (Figure 32) showed the usual effect of sleep loss ($F_{5,48} = 33.00$, $p < .0001$) but this time the gender difference was statistically significant ($F_{1,52} = 5.63$, $p < .05$), with the men reporting higher levels of alertness. The decline was linear ($F_{1,52} = 127.95$, $p < .0001$). As with the KSS, with each successive test session there was a significant increase in reported sleepiness compared to the previous one ($p < .01$), with the exception of session six. There was no gender by session interaction. On the Physically Exhausted-Energetic dimension of the VAS (Figure 33) the results were similar with a main effect of test session ($F_{5,47} = 24.50$, $p < .0001$). With each successive test session there was a significant increase in reported sleepiness compared to the previous one ($p < .05$). The main effect of gender was significant ($F_{1,51} = 4.47$, $p < .05$) but there was no gender by session interaction. The direction of the gender difference was in agreement with the previous scales in that men reported being consistently more energetic than the women. The primary trend was one of a significant linear decline ($F_{1,51} = 122.28$, $p < .0001$), again consistent with the other scales. The Bored-Engaged dimension of the VAS (Figure 34) showed the typical effect of session ($F_{5,47} = 9.14$, $p < .0001$) and the usual linear trend ($F_{1,51} = 39.56$, $p < .0001$). Only the successive decreases from sessions one to two and three to four were significant. The men showed a tendency to report being less bored than the women, with the difference being of borderline statistical significance ($F_{1,51} = 3.62$, $p = .06$). The Sad-Happy dimension (Figure 35) was included with the expectation that responses would not vary over the course of the experiment, except perhaps as the end neared. In fact, there was an effect of session ($F_{5,47} = 7.23$, $p < .0001$) and a significant linear decline in happiness ($F_{1,51} = 8.12$, $p < .01$) but the response changes over the course of the night were quite small compared to the other scales. Some of the session effect was probably due to the slight increase in reported happiness at the end of session six (the VAS is the last measure taken in each test session) as would be expected upon finishing the last task in the last session after a night without sleep. This time the gender difference in the direction of men reporting higher levels of happiness did not reach statistical significance ($p = .15$). It is notable that on all scales except Sad-Happy, the initial differences in subjective sleepiness and fatigue between men and women actually became smaller with increasing sleep deprivation.

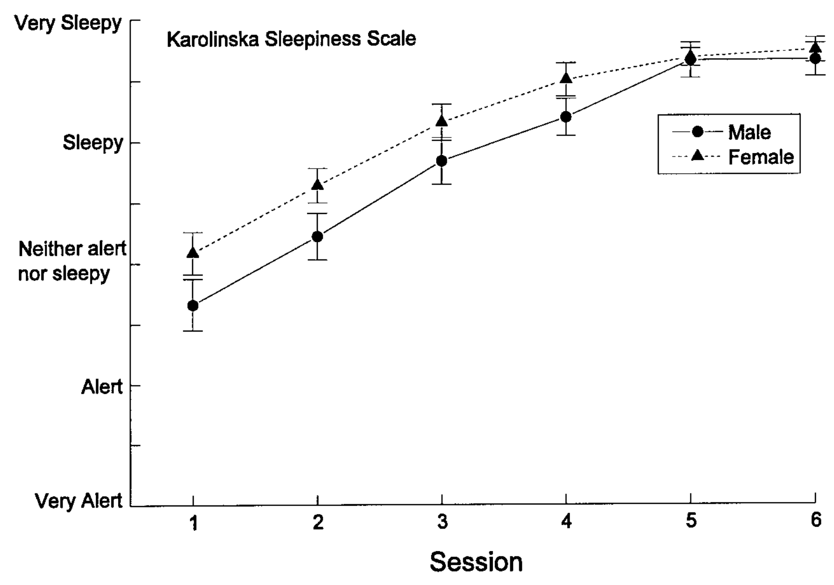


Figure 31: Mean scores on the Karolinska Sleepiness Scale for the six experiment sessions. Error bars represent ± 1 sem.

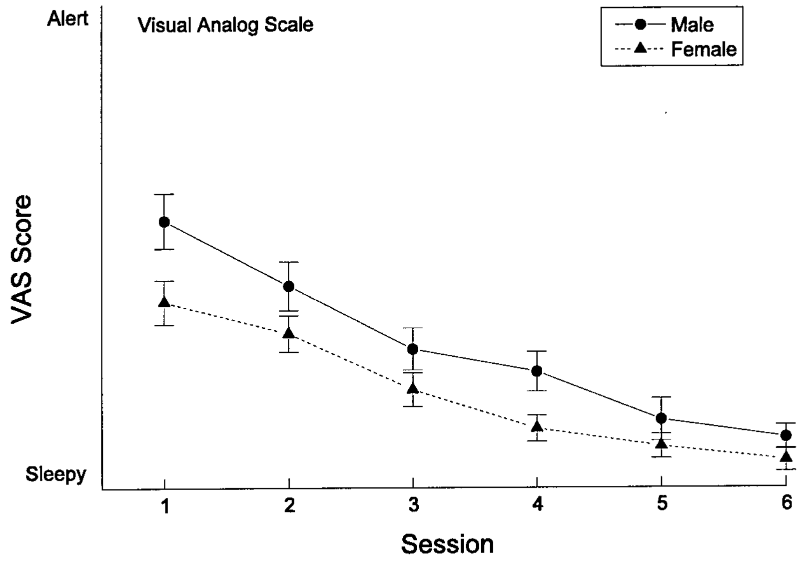


Figure 32: Mean scores on the Sleepy--Alert dimension of the visual analog scale. Error bars represent ± 1 sem.

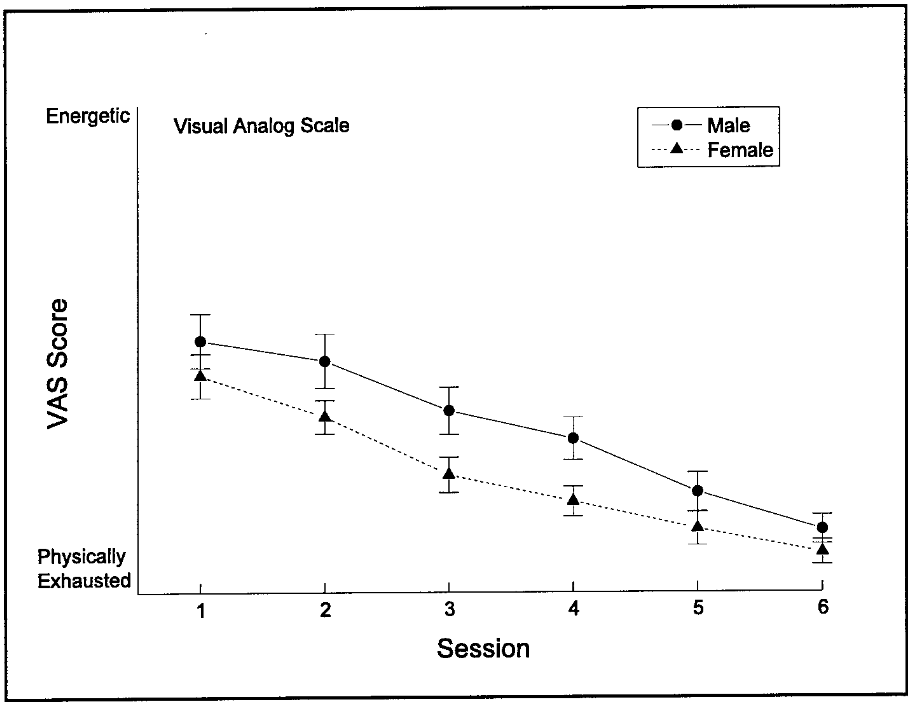


Figure 33: Mean scores on the Physically Exhausted--Energetic dimension of the visual analog scale. Error bars represent ± 1 sem.

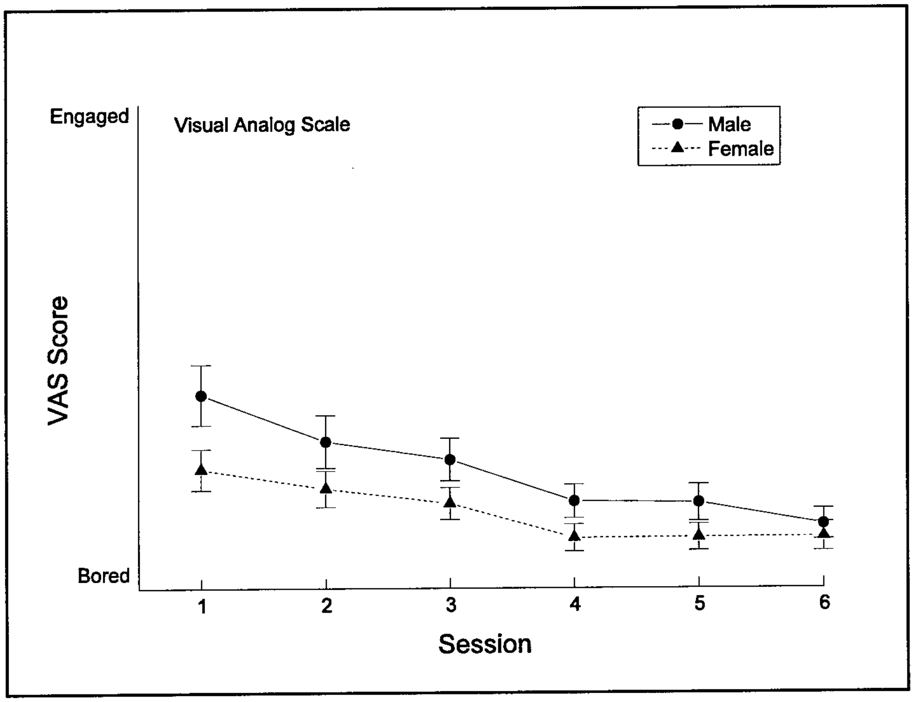


Figure 34: Mean scores on the Bored--Engaged dimension of the visual analog scale. Error bars represent ± 1 sem.

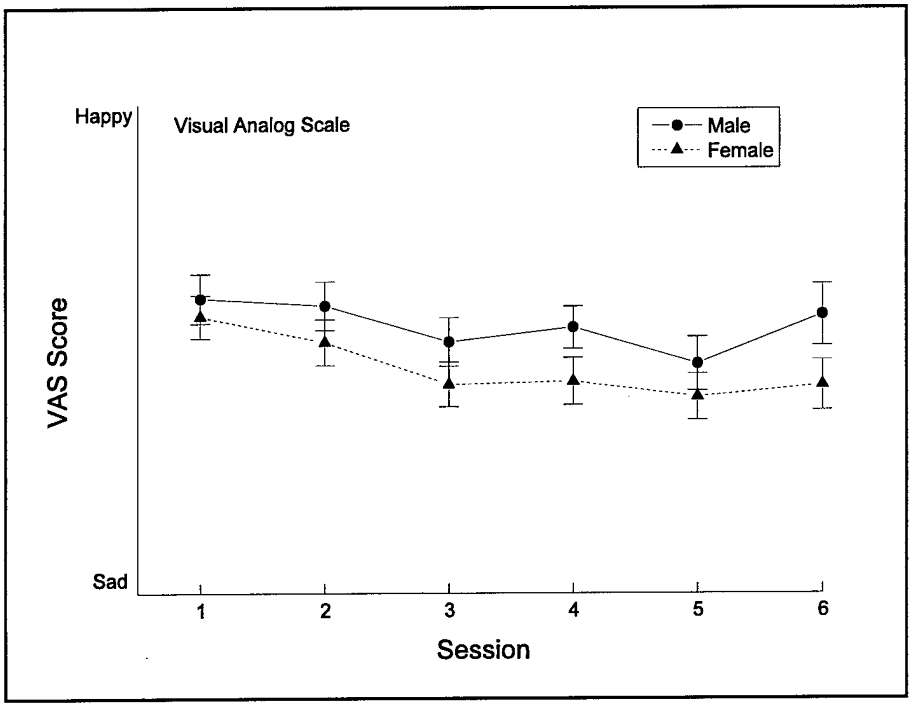


Figure 35: Mean scores on the Sad--Happy dimension of the visual analog scale. Error bars represent ± 1 sem.

Summary. The results of all the statistical analyses described above are summarized in Table II. From the table it can readily be seen that the major and most consistent effect on the various aspects of cognitive performance and subjective sleepiness was that of test session, i.e., increasing sleep deprivation. There was an almost universal linear decline in cognitive performance and subjective mood over the course of the one night without sleep. Clearly the experimental protocol was successful in inducing performance changes as a function of increasing wake duration.

There were but a few significant effects that involved gender, the primary issue of this study. Several of the subjective measures revealed that the women reported being more sleepy, more physically tired, and more bored than the men, consistent with one of our hypotheses. Yet, unlike our expectation, the women responded more like the men near the circadian trough and as sleep loss accumulated, rather than reporting relatively more fatigue. Further, the differences in these few subjective measures, while consistent across sessions, were not especially large.

Of the cognitive measures, gender differences appeared on the median reaction time and fastest 10% reaction time measures of the 15-min simple reaction time task and on the 30-min compensatory tracking task. Both tasks require extended periods of vigilance. The finding for the SRT is consistent with our hypothesis that the *optimum* reaction time performance for the men will be 5-10% faster than that for the women. The data revealed a difference of 6% in the predicted direction. (The men also were faster than the women by 12.5% on the median reaction

Table 11. Summary of Statistical Results on Cognitive and Subjective Tasks

Task	Gender Effect	Session Effect	Session Trend	Gender by Session	Comments
SRT					
Median RT	$p < .05$	$p < .0001$	Linear	$p = .06$	men faster
Fastest 10%	$p = .07$	$p < .0001$	Linear	$p = .055$	men faster (trend)
Lapses		$p < .0001$	Linear		
Slowest 10%		$p < .0001$	Linear		
Anticipations		$p < .0001$	Linear		
CPT					
Mean RT		$p < .0001$	Linear		
% Correct		$p < .0001$	Linear		
MRT					
Mean RT		$p < .001$	Linear		effect of stimulus rotation ($p < .0001$)
% Correct		$p < .01$	Linear		
# Correct		$p < .01$			effect of stimulus rotation ($p < .0001$)
DSST					
Median RT		$p < .0001$	Linear		
# Completed		$p < .0001$	Linear		
% Correct			Linear		
CMT					
Median RT		$p < .0001$	Linear		
# Completed		$p < .0001$	Linear		
% Correct					
TRK					
Distance	$p = .001$	$p < .0001$	Linear	$p = .10$	men better; women more variable
SUBJECTIVE					
KSS	$p = .12$	$p < .0001$	Linear		men more alert (trend)
VAS-Sleepy	$p < .05$	$p < .0001$	Linear		men more alert
VAS-Tired	$p < .05$	$p < .0001$	Linear		men more energetic
VAS-Bored	$p = .06$	$p < .0001$	Linear		men less bored (trend)
VAS-Sad	$p = .15$	$p < .0001$	Linear		men more happy (trend)

time measure.) Unlike our expectation, however, sleep deprivation did not produce statistically significant increases in male false responding even though Figure 14 reveals a trend in that direction in the last two sessions. The only other gender difference occurred with the lengthy TRK on which men performed considerably better than women. Yet a similar and significant (though small) difference was present during the TRK training sessions. For the median reaction time measure on the SRT, the training session difference was also trending in the same direction favoring men. For both the TRK and median reaction time of the SRT, similar small differences were present in the first testing session, growing somewhat as the night progressed. Nevertheless, the differences still failed to result in a significant gender by test session interaction. In fact, there were no gender by session interaction effects for any of the many measures, indicating that, even in the few instances where the groups differed, *men and women did not show different response patterns as a function of sleep loss.*

Our hypothesis predicting better performance by women on the self-paced cognitive tasks (CMT, MRT, DSST, TRK) did not turn out to be the case universally. The TRK results represented a reversal of this prediction. The other three tasks did not show statistically significant gender differences. The trends are interesting nonetheless. Women did show a trend toward faster performance on the CMT and more accurate performance midway through the experiment. On the DSST, women also showed a trend toward faster performance but not better accuracy. An advantage for women on this task was predicted due to its verbal encoding component. This may also underlie the trend toward faster responding by the women on the CPT. However, on the MRT, it was the men who showed the trend toward outperforming the women, especially late in the experiment. This was expected due to the advantage for men in tasks involving spatial relations and mental rotation. Thus there are reasonable interpretations for the trends in the data. In the final analysis, however, it should be emphasized that these trends were not strong enough to reach statistical significance.

In sum, given the facts that in this sleep deprivation experiment (a) so few gender differences occurred, (b) that in many cases these differences were present prior to the start of the experiment and the loss of sleep, (c) that the few gender differences were not nearly as robust as the effects of the sleep loss itself, and (d) that there were no interactions between gender and sleep loss, it appears on the basis of this data set that there are no significant performance differences between men and women in their response to the effects of one night of sleep deprivation.

Methods: UEP Study

Neurobehavioral data (i.e., performance, subjective activation, mood, body temperature) from two major experiments completed on the effects of 54 hrs of sleep deprivation (with 2-hr nap opportunities) and 64 hrs of total sleep deprivation, and seven days of partial sleep deprivation (5 hrs per night), in 52 healthy young adult males and 30 healthy young adult females were statistically analyzed to determine if gender differences were evident. Prior work has not included detailed and explicit analyses aimed specifically at detecting gender differences. Specific performance metrics (e.g., lapse duration, vigilance decrement slope, cognitive throughput) and subjective metrics (e.g., fatigue, perceived effort, motivation) were extracted from each variable class for analysis. Using graphic analyses, regression analyses, and analyses of variance, tests were performed independently on the two data sets to identify main effects of gender on individual performance and self-report metrics (these are effects that are independent of fatigue-mediated circadian pacemaker and/or homeostatic drive for sleep). Analyses focused on interactions of gender status with the circadian rhythm in performance and mood; interactions of gender status with the homeostatic buildup in fatigue from consecutive hours of sleep loss; and interactions of gender status with functioning after recovery sleep. Gender-related differences in variance among metrics were also assessed. Appropriate transformations were used to normalize certain performance metrics (e.g., lapse frequency), and nonparametric tests were applied when appropriate. The results of planned data analyses from the three different experiments were compared among the studies to determine gender outcomes that are reliable (replicable).

More detailed descriptions of the analysis methods can be found in the individual reports from the Unit for Experimental Psychiatry that comprise the following Results and Discussion section.

Results and Discussion: UEP Study

This section is organized around the regular reports submitted by UEP in support of their part of the study. Note that the figure numbering is internally consistent within this section of this report but is different from the other sections.

Report 1: August 15, 1996. The purpose of this project is to analyze neurobehavioral data (i.e., performance, subjective activation, mood, and body temperature) from three sleep deprivation experiments carried out by the Unit for Experimental Psychiatry, with a special focus on gender differences during night work and sustained wakefulness. The findings have relevance for the selection and training of military women and men to engage in continuous operations in which alertness is a key factor.

Work efforts performed under this contract to date have focused on two principal objectives: (1) acquiring relevant data records (computer files, when available, as well as archived hardcopy study protocols) for 52 healthy young adult males and 30 healthy young adult females, studied in two major experiments on the effects of 54 h (with nap[s]) and 64 h of sleep deprivation, and one investigation of the effects of 7 days of partial sleep deprivation (5 h per night); and (2) planning analytical and meta-analytical approaches across studies (including graphic analyses, regression, and analyses of variance) to identify effects of gender on individual performance and self-report metrics.

Data from the study of the effects of 64 h of total sleep deprivation are currently being prepared for analyses. Dependent variables will include raw and transformed measures of psychomotor reaction-time (RT) performance (e.g., fastest 20% RTs, median RTs, lapse duration, vigilance decrement slope), cognitive throughput (e.g., short-term memory probes), mood states, and body temperature.

Data from the 7-day partial sleep deprivation study will be targeted for analysis in September. The relevant dependent measures from this investigation will include indices of mood, cognitive functioning, subjective sleepiness, and performance over the course of the restricted sleep regimen.

Because the remaining study (54-h total sleep deprivation with naps) was run 10-13 years ago, the majority of its data are archived. We are currently extracting these data and converting it to a form suitable for the analytical purposes of the current project. The bulk of the performance data was retrieved from computer files created in Digital Equipment Corporation's early OS-8 operating system for PDP minicomputers, which were available to us only on Linc Tape media. To accomplish the transfer of these data files to a format that could be read by our MicroSoft Windows-based statistical software (Systat), we transferred the data from the OS-8 format to generic ASCII files that can be read under DOS v. 6.11 and imported into Systat for Windows v. 6.0. Measures of cognitive task performance, sleepiness, and mood are in the process of being coded and transcribed from hardcopy experimental protocols for data entry. As of this report, these activities are approximately 30% complete. We project that these data will be fully available for analysis by early November.

Finally, the senior investigators have been planning an analytical strategy geared at the isolation of effects on performance and mood attributable to circadian variation from those engendered by the homeostatic buildup of fatigue associated with consecutive hours of sleep loss. Such a distinction will be important in assessing potential interactions of gender status with the various performance metrics utilized when modeling fatigue-related neurobehavioral decrements during sustained operations.

Report 2: September 15, 1996. Data from our investigation of the effects of 64 h of total sleep deprivation among 24 healthy young adult volunteers (17 males, 7 females) have been the basis of our most recent focus of analyses. The dependent variables have included raw and transformed measures of psychomotor vigilance testing (PVT) (e.g., fastest 10% RTs, median RTs, lapse duration, number of lapses, vigilance decrement slope), cognitive throughput (e.g., short-term memory probes, digit symbol substitution task), mood states (POMS), and body temperature (sublingual). Using repeated-measures analysis of variance, we examined each of these variables during a key period of the protocol, representing 48 consecutive hours of sustained wakefulness with intensive assessments of neurobehavioral functioning every two hours. This sampling window is ideal because it permits gender comparisons in response to (a) 2 days of total sleep deprivation, across (b) the circadian cycle, allowing (c) a determination of the impact of these factors on daytime vs. nighttime performance capability. Although these preliminary findings are based on an intensively studied sample of only 17 males and 7 females, and thus will require replication across the remaining data sets, they identify specific gender differences in three assessment domains: body temperature rhythms, cognitive throughput on the digit symbol substitution task (DSST), and self-reported fatigue.

Figure 1 depicts variation in body temperature as a function of time-of-day and cumulative sleep loss for male and female study participants. The usual circadian pattern, with a peak at midday and a trough in the early morning hours, is clearly discernible among both male and female subjects ($F(11,242) = 19.21$, $p < .001$). In addition, the daily average temperature for all subjects decreased by more than 0.1°F from the first to the second day of sleep deprivation ($F(1,22) = 23.48$, $p < .001$). Neither of these general main effects interacted with gender. However, follow-up analyses revealed a significant interaction in the single degree of freedom polynomial contrast (linear component) of the temperature rhythm ($F(1,22) = 4.575$, $p = .044$), involving gender, time-of-day and Day 1 vs. 2 of sleep deprivation. Examination of the linear regressions of temperature on time indicated that body temperature in males decreased on average by 0.023°F per h, while temperature in females decreased at a somewhat greater rate (0.031°F per h) during the first 24 h of sleep deprivation; during the second day of deprivation, however, this pattern was reversed, with the linear rate of temperature decrease being greater for males (0.031°F per h) than for females (0.015°F per h).

Figure 2 shows performance of males and females, over the same period of sleep deprivation, on the digit symbol substitution task, a subscale of the Wechsler Adult Intelligence Scale-Revised. As can be seen, performance deterioration occurs for both groups during the early morning hours of their first full day of sleep deprivation, and becomes even more pronounced by the second day, resulting in a significant interaction of days x trials ($F(11,242) = 4.617$, $p < .001$). Moreover, task performance is significantly poorer on Day 2 than on Day 1 ($F(1,22) = 45.25$, $p < .001$), which is indicative of the impact of sleep deprivation. The interaction of gender with these outcomes was not significant. Nevertheless, gender differences were revealed by a significant 3-way interaction in the third order polynomial contrast (i.e., cubic trend). This effect may be understood by reference to Figure 2. Female participants showed greater intertrial variability over the period from 1000 h to 2200 h of Day 2, including a greater capacity for short-lived recoveries to the performance levels of the preceding day, than did male subjects.

Mean scores on the Fatigue subscale of the Profile of Mood States (POMS) are portrayed graphically for male and female participants in Figure 3. The corresponding analyses of variance found evidence of circadian variation in self-reported fatigue over trials within days ($F(11,242) =$

11.69, $p < .001$), and an increase in fatigue across days of sleep deprivation ($F(1,22) = 59.83$, $p < .001$). Moreover, the trials x days x gender interaction was also significant ($F(11,242) = 2.15$, $p = .018$). This interaction is evident in Figure 3. Females reported greater levels of fatigue than males for the first 2 h of waking, but less fatigue during the hours of 0600 through 1000 on the second morning, after a night without sleep. By 2200 h of the second night of sleep deprivation, and continuing through 0400 h, however, the self-reported fatigue levels of female participants consistently exceeded those of male subjects.

Objectives for the next budget period include: (a) continued analyses of the data from the 64-h total sleep deprivation study that will encompass the final 48-h period of sleep deprivation as well as the recovery period; and (b) preliminary analyses of data from the study of partial sleep deprivation. The current findings will be discussed in conjunction with these subsequent analyses.

Figure 1

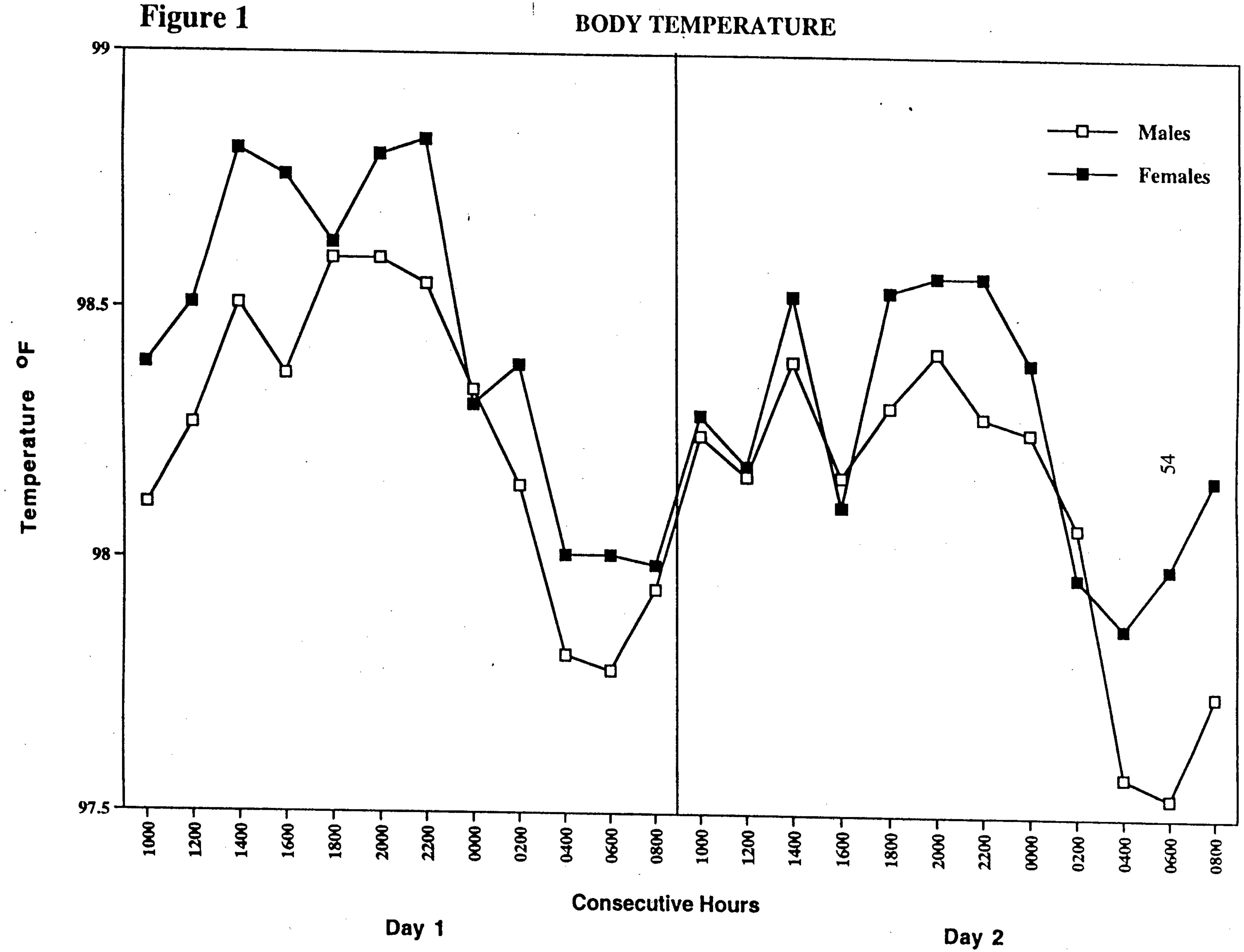

BODY TEMPERATURE
Males
Females
Temperature °F
99
98.5
98
97.5
1000
1200
1400
1600
1800
2000
2200
0000
0200
0400
0600
0800
1000
1200
1400
1600
1800
2000
2200
0000
0200
0400
0600
0800
Consecutive Hours
Day 1
Day 2

Figure 2

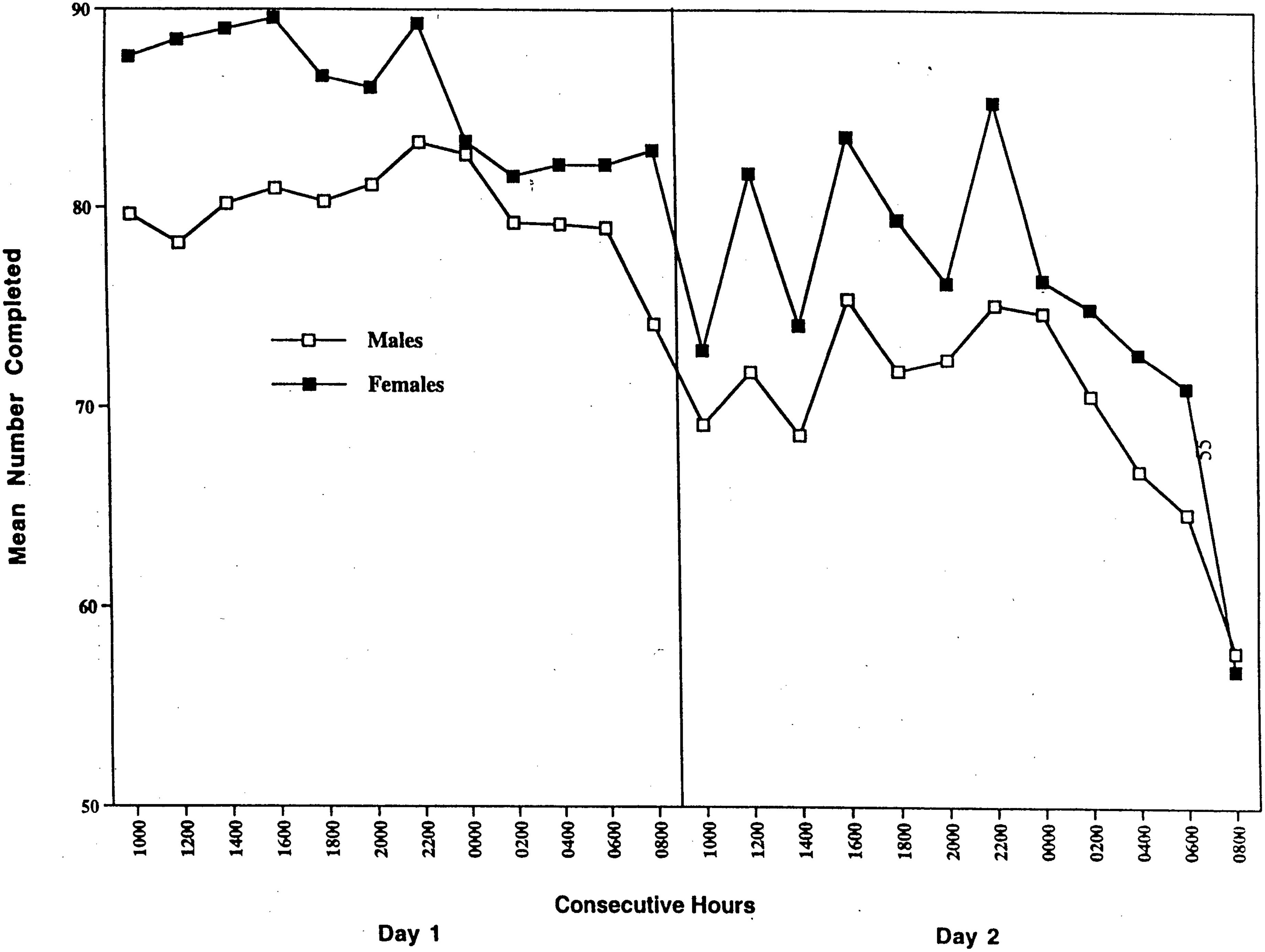
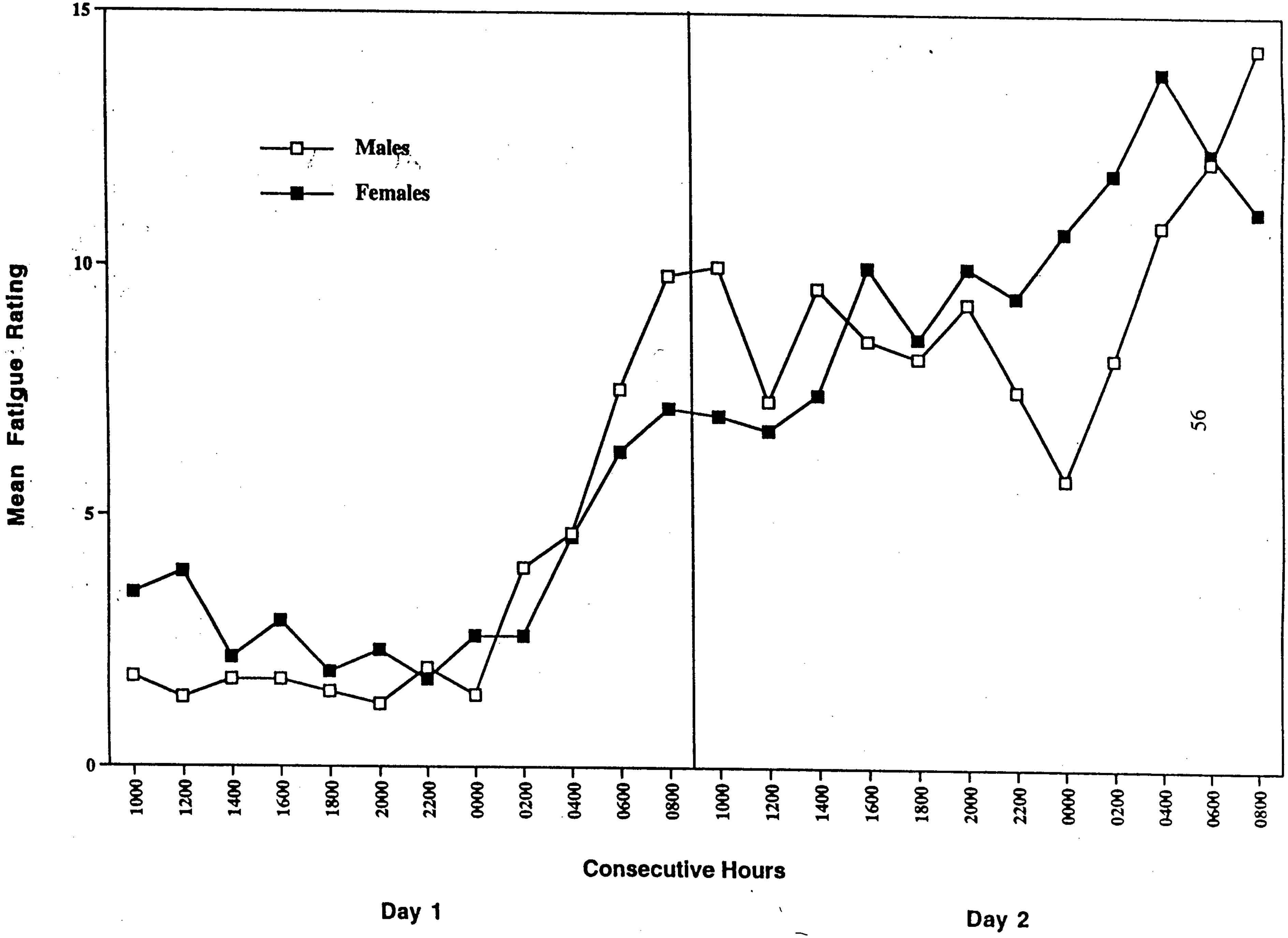

DSST
Mean Number Completed
90
80
70
60
50
Males
Females
1000
1200
1400
1600
1800
2000
2200
0000
0200
0400
0600
0800
1000
1200
1400
1600
1800
2000
2200
0000
0200
0400
0600
0800
Consecutive Hours
Day 1
Day 2

Figure 3


POMS FATIGUE
Males
Females
Mean Fatigue Rating
15
10
5
0
1000
1200
1400
1600
1800
2000
2200
0000
0200
0400
0600
0800
1000
1200
1400
1600
1800
2000
2200
0000
0200
0400
0600
0800
Consecutive Hours
Day 1
Day 2

Report 3: October 15, 1996. Analyses undertaken for the current report are based on two studies performed by our laboratory. One of these is an investigation of the effects of 64 h of total sleep deprivation among 24 healthy young adult volunteers (17 males, 7 females). Our previous progress report (September 15, 1996) detailed findings relevant to neurobehavioral functioning during the sleep deprivation phase of this study; the current report is concerned with recovery of functioning following restoration of access to normal sleep requirements. The second study examines gender differences in response to seven days of partial sleep deprivation (5 h per night) in a sample consisting of 8 male and 8 female participants.

During recovery from 64 h of total sleep deprivation, the following variables were monitored over a 48-h period: Raw and transformed measures of psychomotor vigilance testing (PVT) (e.g., fastest 10% RTs, median RTs, lapse duration, number of lapses, vigilance decrement slope), cognitive throughput (e.g., short-term memory probes, digit symbol substitution task), mood states (POMS), and body temperature (sublingual). Repeated-measures analyses of variance found that female volunteers generally had higher body temperatures than male subjects (Means = 98.77°F vs. 98.41°F, respectively) during the recovery phase. Gender differences were also detected for several mood states. Figure 1 shows the differential response of males and females along the POMS dimension of Anger-Hostility during the recovery period when normal nighttime sleep was permitted. A significant Days x Gender interaction was obtained ($\underline{F}(1,13) = 6.32$, $\underline{p} = .03$). As the figure illustrates, females exhibited a greater decrease in Anger-Hostility than males over the two recovery days, with scores at or near zero throughout the second recovery day. Figure 2 depicts gender differences in self-reported Fatigue during the recovery period. A repeated-measures ANOVA found a significant main effect of gender, such that female participants reported greater Fatigue than male participants throughout the entire recovery phase of the study ($\underline{F}(1,14) = 4.70$. $\underline{p} = .05$). Nevertheless, both groups appear to have benefitted from the recovery sleep opportunities, as fatigue scores decreased reliably across trials ($\underline{F}(4,56) = 9.59$, $\underline{p} < .05$). Finally, Figure 3 shows self-reported levels of Tension-Anxiety during the recovery period. Repeated-measures ANOVA identified a significant Days x Trials x Gender interaction ($\underline{F}(4,52) = 3.05$, $\underline{p} = .025$), which was due primarily to an increase in Tension-Anxiety among female subjects toward the end of the protocol. Males, on the other hand, reported relatively stable levels of Tension-Anxiety throughout recovery.

Analyses of the investigation of seven days of partial sleep deprivation were based on 3 sampling periods (i.e., 1000 h, 1600 h, 2200 h) for each of 8 days (1 baseline day and 7 consecutive partial sleep deprivation days). The dependent variables examined consisted of psychomotor vigilance testing (PVT) (e.g., fastest 10% RTs, median RTs, lapse duration, number of lapses, vigilance decrement slope), cognitive throughput (e.g., short-term memory probes), and mood states (POMS). A repeated-measures ANOVA found a reliable Days x Gender interaction for performance on the probed-recall task ($\underline{F}(7,98) = 2.1$, $\underline{p} = .05$). As Figure 4 indicates, male subjects had higher baseline scores than females, and exhibited a substantial decline in performance on the second day of partial sleep deprivation. Female participants maintained a more consistent level of performance throughout the study. Figure 5 documents differential changes in total mood disturbance (POMS) as a function of time-of-day for male and female subjects over the course of the 8-day protocol (F(2,28) = 3.96, p = .03). Whereas male subjects showed a steady increasing trend toward greater mood disturbance as the study progressed, females exhibited sharp variations in mood disturbance within days. Follow-up analyses determined that the source of this effect consisted of gender-specific variations in Fatigue (F(2,28) = 2.76, p = .09) and Vigor (F(2,28) = 4.41, p = .02) within days.

Objectives for the next budget period include: (a) continued analyses of the data from the 7-day partial sleep deprivation study; and (b) preliminary analyses of data from the study of 54 h of total sleep deprivation with naps. Efforts are underway to develop graphical and meta-analytic strategies to aggregate the major findings of the three sleep-deprivation investigations.

Figure 1

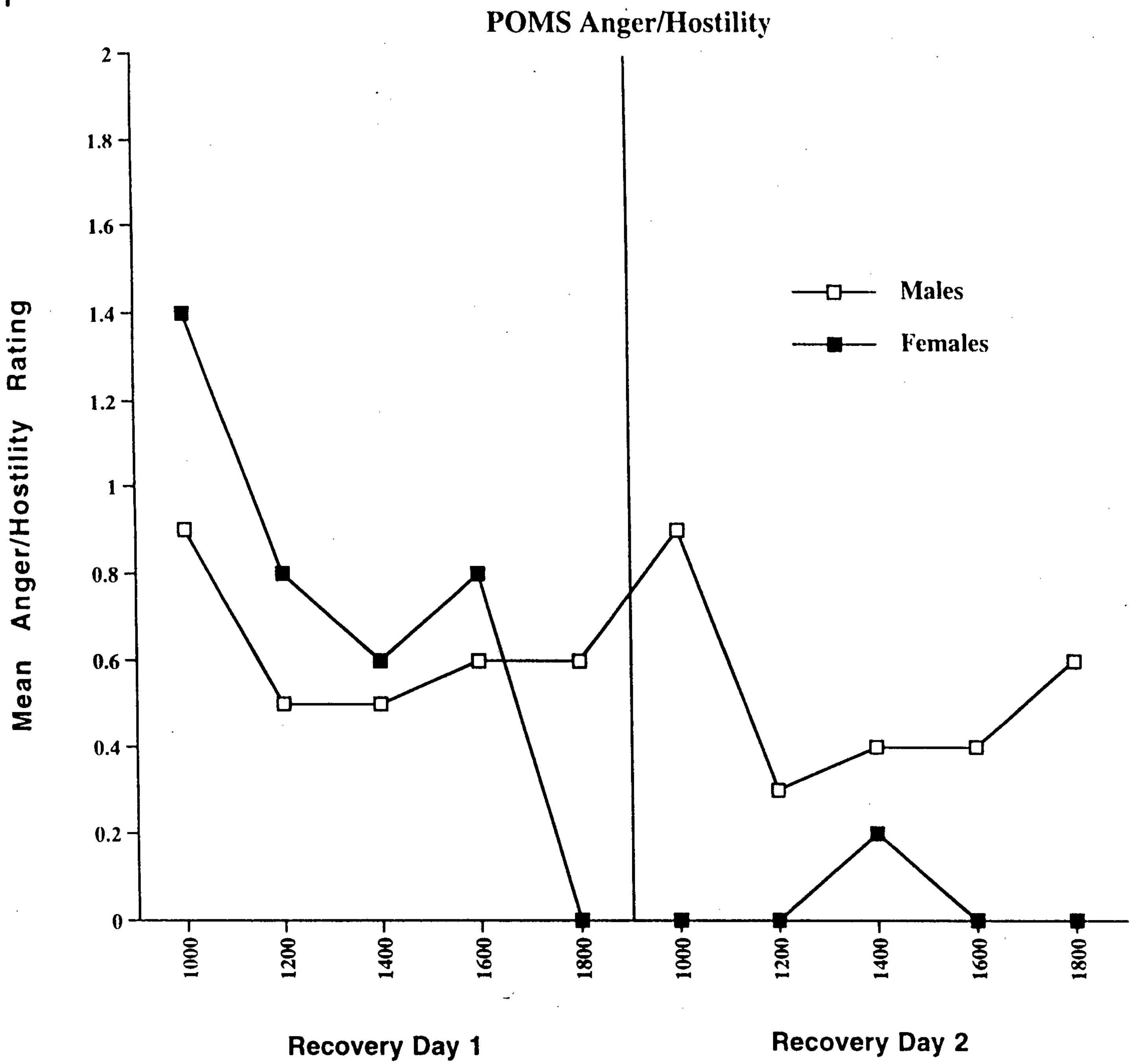

POMS Anger/Hostility
Mean Anger/Hostility Rating
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1000
1200
1400
1600
1800
1000
1200
1400
1600
1800
Recovery Day 1
Recovery Day 2
Males
Females

Figure 2

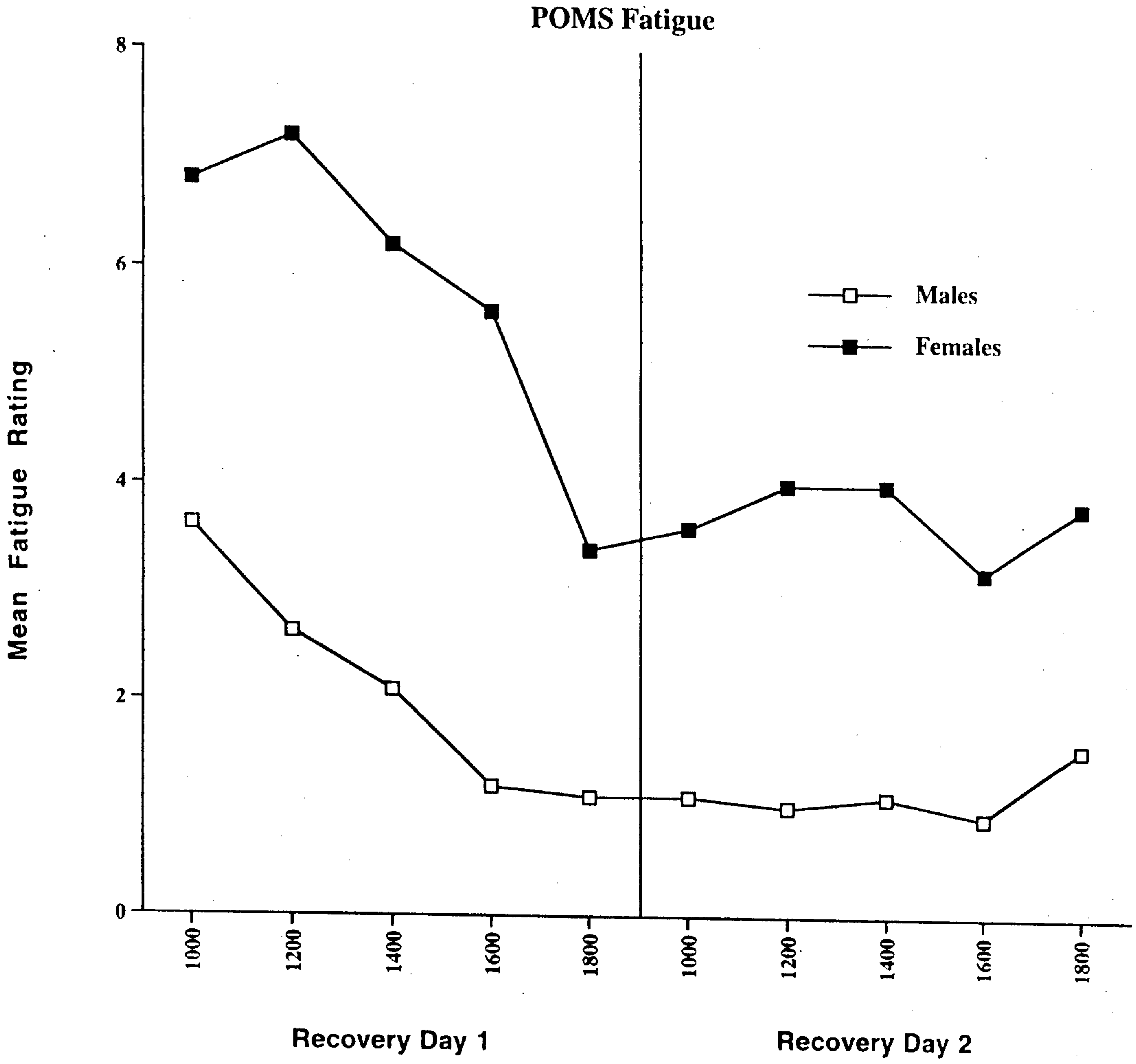

POMS Fatigue
Mean Fatigue Rating
0
2
4
6
8
Males
Females
1000
1200
1400
1600
1800
1000
1200
1400
1600
1800
Recovery Day 1
Recovery Day 2

Figure 3

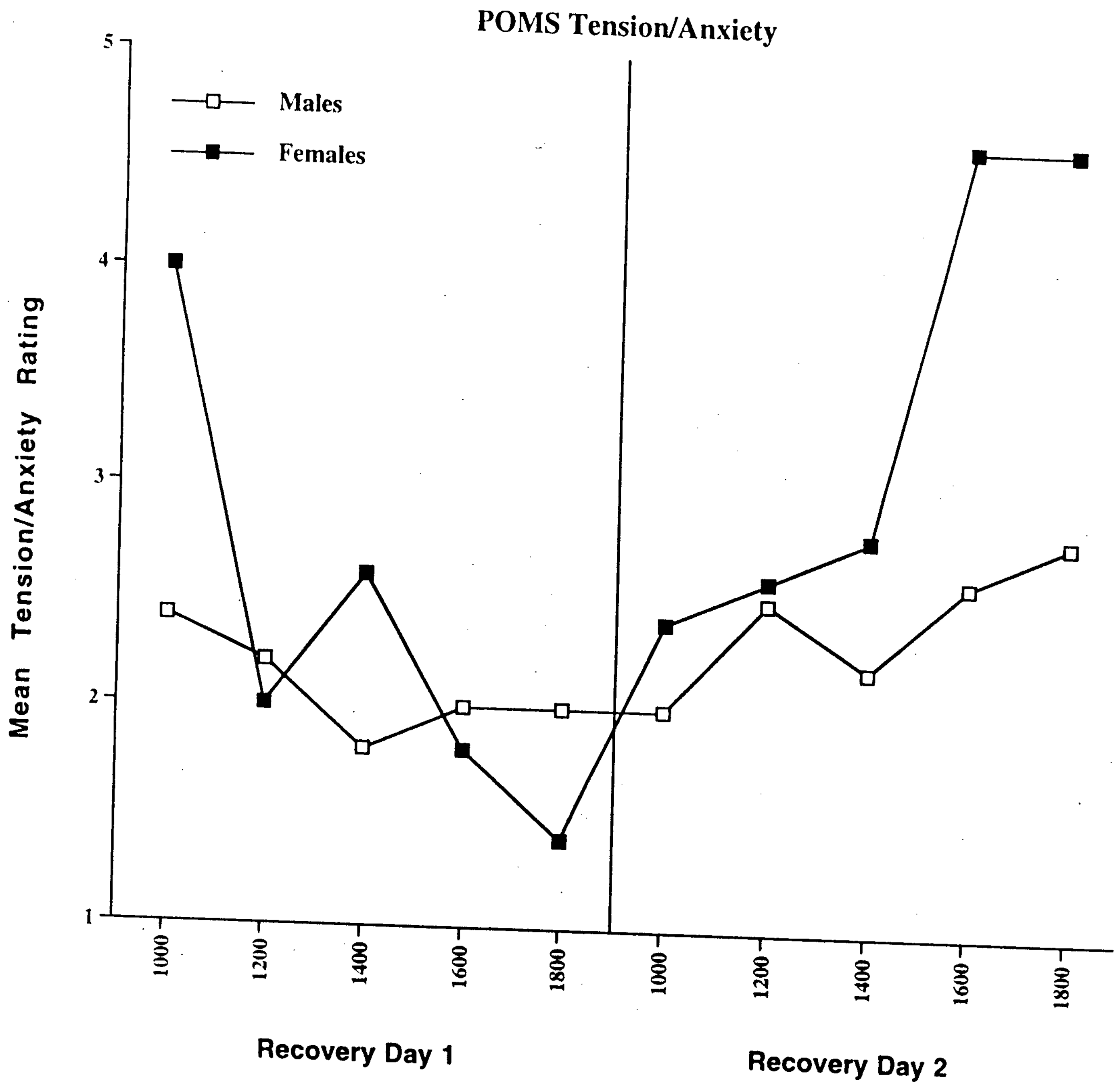

POMS Tension/Anxiety
Males
Females
Mean Tension/Anxiety Rating
1
2
3
4
5
1000
1200
1400
1600
1800
1000
1200
1400
1600
1800
Recovery Day 1
Recovery Day 2

Figure 4

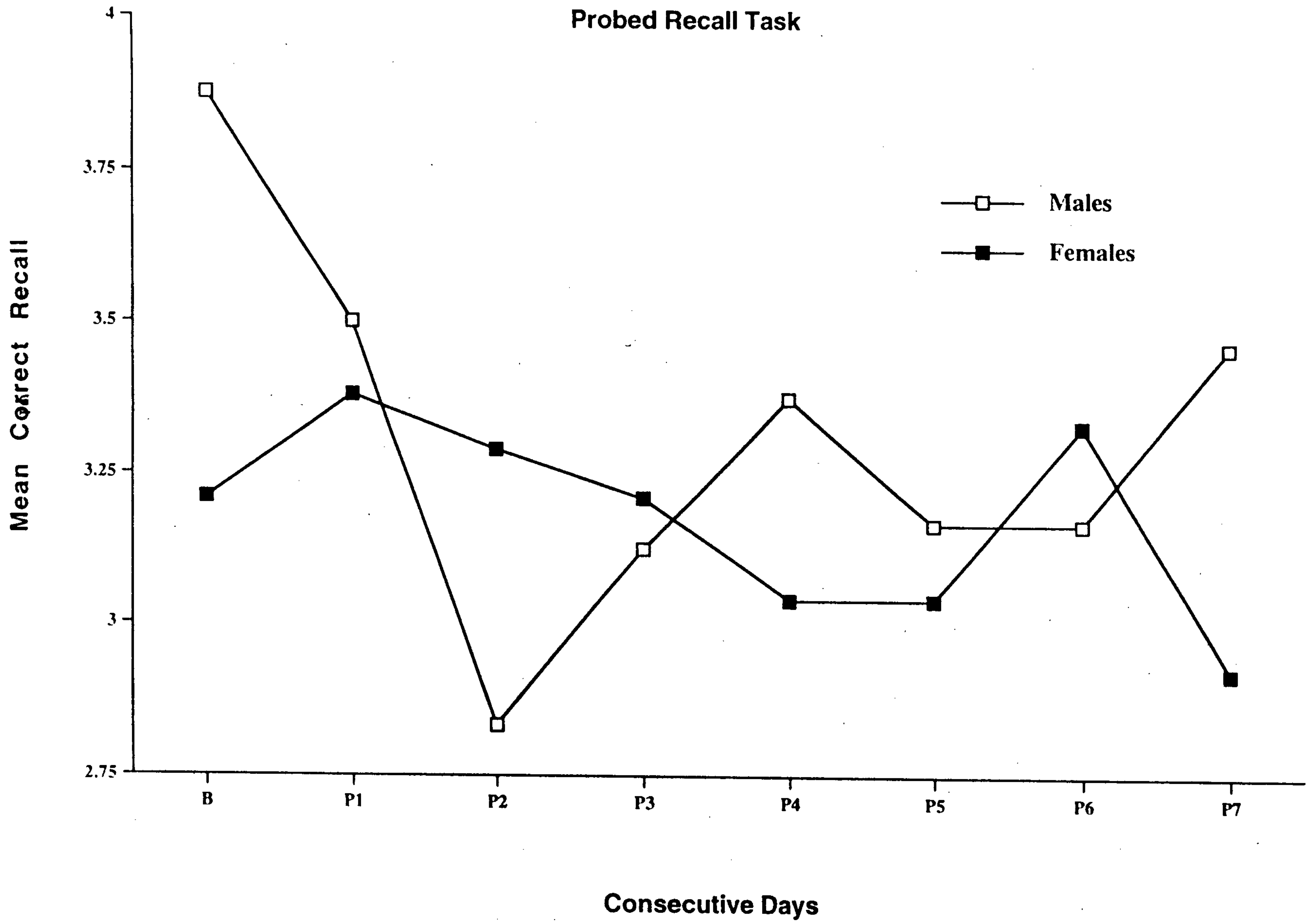
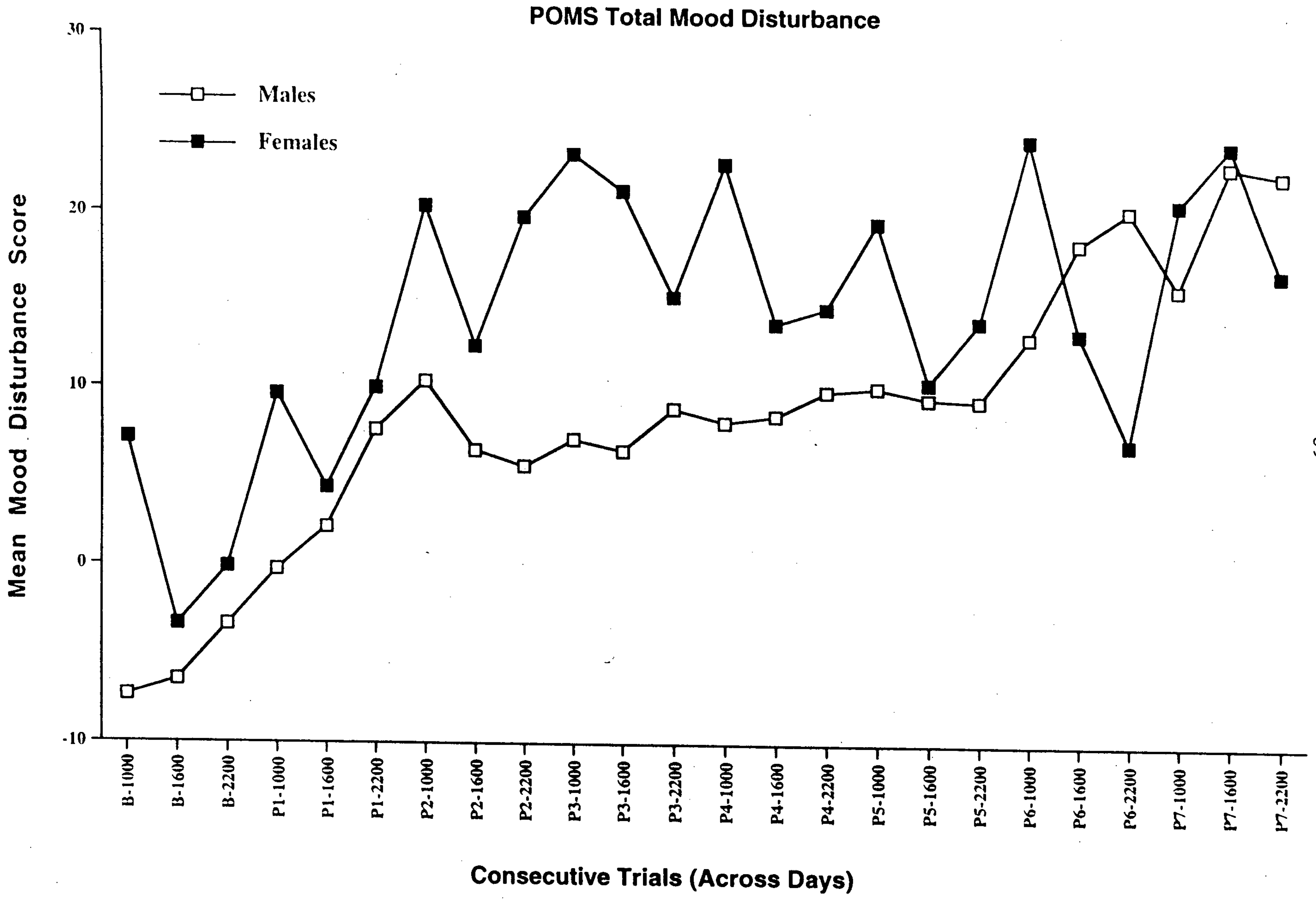

Probed Recall Task
Mean Correct Recall
4
3.75
3.5
3.25
3
2.75
B
P1
P2
P3
P4
P5
P6
P7
Consecutive Days
Males
Females

Figure 5


POMS Total Mood Disturbance
Males
Females
Mean Mood Disturbance Score
30
20
10
0
-10
B-1000
B-1600
B-2200
P1-1000
P1-1600
P1-2200
P2-1000
P2-1600
P2-2200
P3-1000
P3-1600
P3-2200
P4-1000
P4-1600
P4-2200
P5-1000
P5-1600
P5-2200
P6-1000
P6-1600
P6-2200
P7-1000
P7-1600
P7-2200
Consecutive Trials (Across Days)

Report 4: November 15, 1996. Analyses undertaken for the current report are based on an investigation of 54 h of total sleep deprivation in a volunteer, young adult sample consisting of 26 males and 15 females, with 2-h nap opportunities strategically placed following either 6, 18, 30, 42, or 54 h of total sleep loss. Naps occurring after 6, 30, and 54 h of participation in the sleep deprivation protocol took place at the peak of the circadian cycle of alertness, while those scheduled to occur following 18 and 42 h of continuous wakefulness took place at the circadian nadir. The principal objective of this investigation was to ascertain the optimal placement of a brief nap opportunity during a period of prolonged wakefulness, so as to ensure either prophylactic or restorative effects against the performance and alertness decrements associated with sleep loss. Major focuses of the current set of analyses were to continue our investigation of possible gender-related differences in response to sleep loss in the context of sustained operations, and to determine whether brief nap opportunities might differentially offset performance impairment for men as compared to women.

The following dependent measures were monitored over the course of the 54-h sleep deprivation protocol: Raw and transformed measures of psychomotor vigilance testing (PVT) (e.g., fastest 10% RTs, median RTs, lapse duration, number of lapses, vigilance decrement slope), cognitive throughput (e.g., digit span subtest of the Weschlser Adult Intelligence Scale (WAIS)), mood states (POMS), and body temperature (sublingual).

A repeated measures analysis of variance (ANOVA) carried out on the PVT data found a significant between-subjects main effect of gender ($\underline{F}(1,28) = 4.67$, $\underline{p} = .039$) as well as significant gender x trials ($\underline{F}(14,392) = 1.89$, $\underline{p} = .026$) and napping condition x trials ($\underline{F}(56,392) = 1.80$, $\underline{p} = .001$) interactions for median response latency. Figure 1 graphically depicts both the between-subject effect of gender and the gender x trials interaction. As the figure illustrates, female volunteers generally exhibited slower reaction times than their male counterparts throughout the study, but they were particularly impaired during the early morning hours (circadian trough) of the second night of sleep deprivation (0500 - 1000 h). Supplemental analyses of the PVT data found nonsignificant trends for effects of gender on lapses, as well as on fastest and slowest 10% of responses, which suggests that the gender differences in median response latency represent primarily a generalized response slowing among female volunteers that may have been exacerbated as a function of cumulative sleep loss. In addition, follow-up analyses of the napping condition x trials interaction revealed that naps taken early in the protocol (following 6 and 18 h) yielded longlasting performance benefits; nevertheless, they appeared to do so equally for male and female participants.

There was a significant napping condition x trials effect on oral temperature ($\underline{F}(68,510) = 2.17$, $\underline{p} < .001$). Follow-up analyses indicated that naps taken early in the protocol resulted in smaller differences in temperature derived from equivalent time points during Day 1 and Day 2 than did naps that occurred late in the protocol (after 42 or 54 h), which typically were characterized by lower temperatures during the same time periods of Day 2 relative to Day 1.

However, the effect of the timing of naps on oral temperature was not influenced by gender (F(68,510) = 1.22, p = .125).

The Digit Span subtest of the WAIS was used as an index of sleep loss effects on short-term memory. The task assesses an individual's ability to correctly repeat an increasing series of single-digit numbers, read by the examiner, in both forward and backward orders. Repeated-measures ANOVAs carried out over the full range of trials conducted during the 54-h sleep deprivation protocol yielded a between-subjects main effect of gender for the backward report condition that failed to achieve the conventional level of significance (F(1,31) = 3.69, p = .064). This effect represented an overall greater ability of women relative to men to reproduce a series of digits in backward order. Because the Digit Span task was administered before and after naps, we also examined the immediate effects of brief (2-h) sleep opportunities on short-term memory capability. Figure 2 shows the performance of male and female study participants in both the forward and backward Digit Span conditions prior to and following a 2-h nap, which occurred at one of five predetermined times during the sleep deprivation study. Women exhibited superior performance to men overall in the forward condition (F(1,31) = 7.91, p = .008), but both groups were adversely affected by being tested shortly after a nap opportunity (F(1,31) = 5.02, p = .032). In the backward condition, however, women showed no effect of sleep inertia, whereas men evidenced significant impairment (F(1,31) = 5.29, p = .028).

Variation in mood, as assessed by the POMS, was found to be statistically significant only in the case of the Anger-Hostility subscale. Figure 3 depicts self-reported anger-hostility over trials of sleep deprivation as a function of gender. A significant gender x trials interaction (F(14,434) = 2.04, p = .014) confirmed the pattern apparent in Figure 3 in which female volunteers reported higher levels of anger-hostility than males during the late morning and afternoon hours of Day 1 of sleep deprivation as well as during the early morning hours of Day 2. Additionally, the napping condition x gender x trials interaction was significant (F(56,434) = 1.68, p = .003). The locus of their 3-way interaction involved an exacerbation, in the 30-h nap onset condition, of the general pattern observed among females, coupled with a singular rise in anger-hostility among males in the 30-h nap onset condition at approximately 1000 h of the second day of total sleep deprivation.

Primary objectives for the next budget period include continued analyses of data from the 54-h total sleep deprivation study, as well as initial efforts to aggregate and integrate gender-relevant findings across the three sleep deprivation experiments carried out by our laboratory in the domains of psychomotor performance, body temperature, cognitive throughput, and mood.

Figure 1

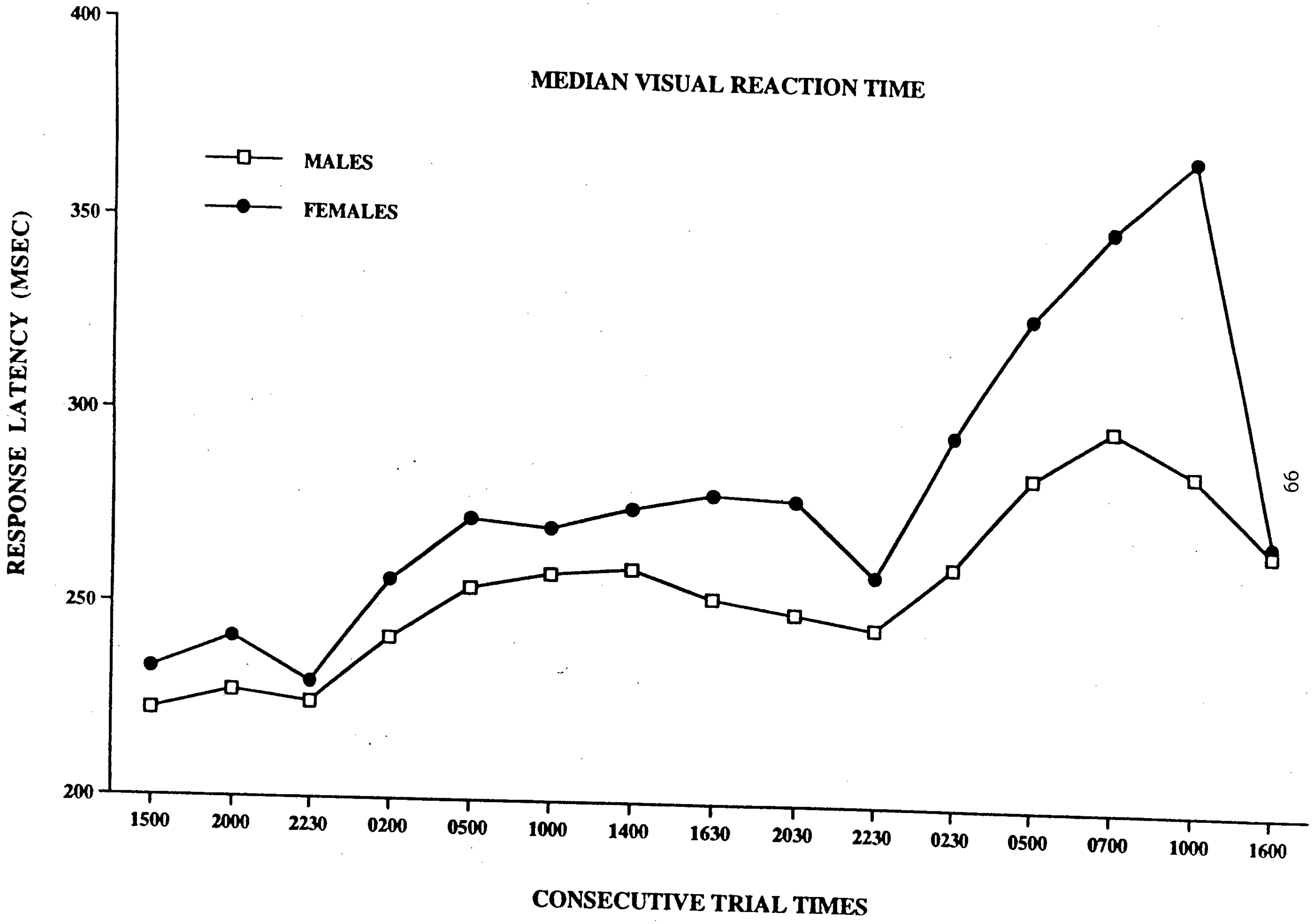

MEDIAN VISUAL REACTION TIME
MALES
FEMALES
RESPONSE LATENCY (MSEC)
400
350
300
250
200
1500
2000
2230
0200
0500
1000
1400
1630
2030
2230
0230
0500
0700
1000
1600
CONSECUTIVE TRIAL TIMES

Figure 2

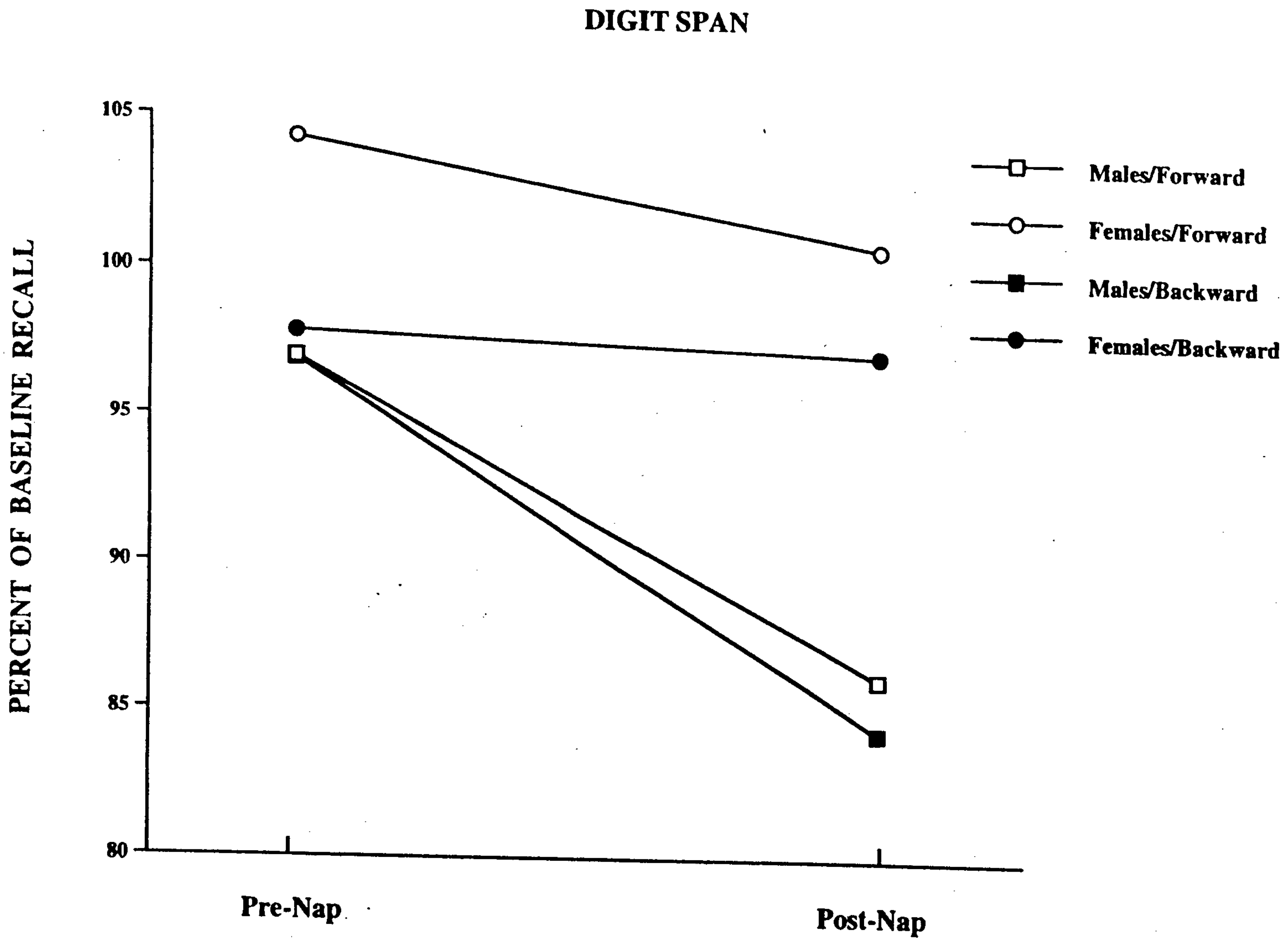
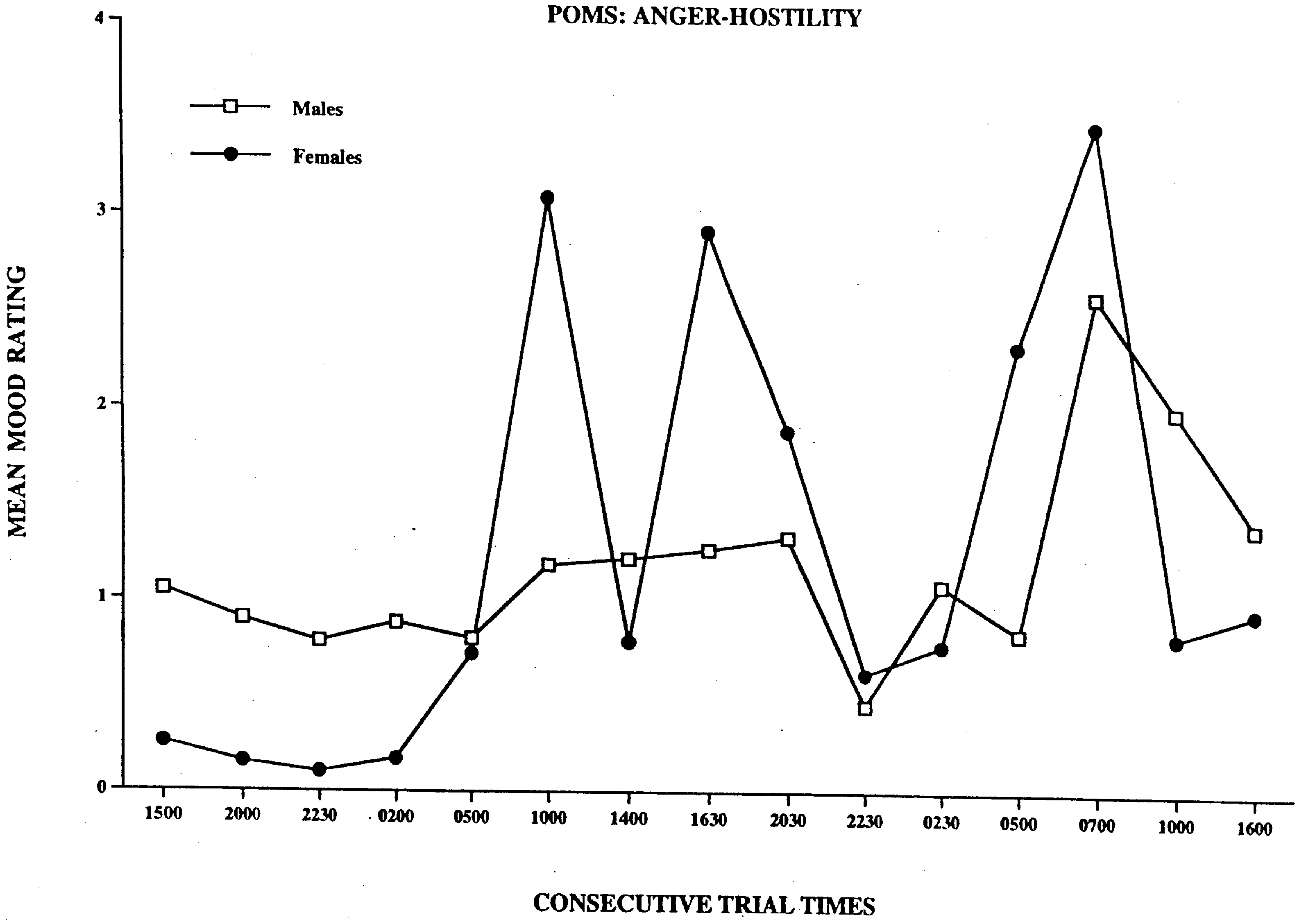

DIGIT SPAN
PERCENT OF BASELINE RECALL
105
100
95
90
85
80
Pre-Nap
Post-Nap
Males/Forward
Females/Forward
Males/Backward
Females/Backward

Figure 3


POMS: ANGER-HOSTILITY
Males
Females
MEAN MOOD RATING
0
1
2
3
4
1500
2000
2230
0200
0500
1000
1400
1630
2030
2230
0230
0500
0700
1000
1600
CONSECUTIVE TRIAL TIMES

Report 5: December 15, 1996. Analyses undertaken for the current report follow upon those of the previous budget period, and are based on an investigation of 54 h of total sleep deprivation in a volunteer, young adult sample consisting of 26 males and 15 females, with 2-h nap opportunities strategically placed following either 6, 18, 30, 42, or 54 h of total sleep loss. Naps occurring after 6, 30, and 54 h of participation in the sleep deprivation protocol took place near the peak of the circadian cycle of alertness, while those scheduled to occur following 18 and 42 h of continuous wakefulness took place near the circadian nadir. The principal objective of this investigation was to ascertain the optimal placement of a brief nap opportunity during a period of prolonged wakefulness, so as to ensure either prophylactic or restorative effects against the performance and alertness decrements associated with sleep loss. Major focuses of the current set of analyses were to continue our investigation of possible gender-related differences in response to sleep loss in the context of sustained operations, and to determine whether brief nap opportunities might differentially offset performance impairment for men as compared to women.

For the current report, the following dependent measures were monitored over the course of the 54-h sleep deprivation protocol: Descending Subtraction Task (DST), a brief-duration, short-term memory task that requires the subject to perform rapid serial subtractions in which the subtrahend decreases successively from 9 to 1 in repeating cycles (e.g., starting number = 876: 876 - 9 = 867, 867 - 8 = 859, 859 - 7 = 852, etc.; thus, the task places great demands on short-term memory in that subjects must keep track of both the current minuend and subtrahend while performing the operation quickly and accurately); and the Memory and Search Task (MAST), which requires subjects to rapidly search successive strings of printed characters and digits for the presence of predesignated target stimuli. The MAST provides a measure of both short-term memory and vigilance capacity and includes two levels of difficulty, involving searches for 2 target stimuli or for 6 target stimuli.

A repeated measures analysis of variance (ANOVA) carried out on the DST data found a significant trials main effect ($F(10,310) = 8.72$, $p < .001$) and a napping condition x trials interaction ($F(40,310) = 2.36$, $p < .001$). The interaction was due to the restorative effects of a 2-h nap opportunity having less impact when scheduled late in the protocol (i.e., at 42 h and 54 h than at 30 h or earlier). However, there were no interactions involving gender, which suggests that male and female participants benefitted equally from an early nap opportunity.

The MAST data revealed significant main effects for trials (i.e., increasing search times over trials) for both 2-target ($F(6,186) = 3.52$, $p = .003$) and 6-target ($F(6,186) = 4.23$, $p = .001$) conditions, as well as significant interactions of trials with napping occasions for both 2-target ($F(24,186) = 2.25$, $p = .001$) and 6-target ($F(24,186) = 1.63$, $p = .039$) stimulus conditions. The interactions reflected poorer performance during the latter trials of the sleep deprivation phase for subjects who received the earliest nap opportunity (i.e., at 6 h) and for those subjects whose naps occurred only at the end of the study (i.e., after 54 h), compared to subjects whose naps were placed at intermediate times (i.e., at 18 h, 30 h, and 42 h). As in the case of the DST task, however, task performance did not vary as a function of gender.

In an earlier report, we identified certain measures that revealed an influence of gender on indices of cognitive throughput during sleep deprivation (i.e., the Digit Span subtest of the WAIS during the 54-h total sleep deprivation protocol with 2-h nap opportunities and a probed recall test during 7 days of partial sleep deprivation); nevertheless, a similar pattern of findings did not obtain for the DST and MAST assessments administered during the 54-h total sleep deprivation study. These differences in outcome might reflect variation in the sensitivity of the memory tests themselves, or they might be due to differences in the timing of the tests (i.e., the MAST was administered on only 7 occasions, while the DST was administered on only 11 occasions throughout the protocol). We will attempt to determine if time of administration, which could reflect the impact of circadian influences or proximity to a napping opportunity, is a plausible explanation for the failure of the MAST and DST to support the potentially differential effects of sleep deprivation on memory processes in male and female study participants.

In a preliminary effort to summarize our findings with regard to gender influences across the three sleep deprivation studies, we have listed in Table 1 the specific dependent measures examined and the corresponding main and interactive effects that were identified. Although none of the effects were equally prominent in all three studies, our current objective is to equilibrate metrics across studies, so as to allow large-scale analyses to be conducted, which take advantage of the substantial increase in statistical power afforded by combining the three subject samples. With judicious selection of sampling points most representative of the extremes of sleep deprivation within each investigation, it should be possible to derive strong conclusions concerning gender effects despite differences in experimental protocols among the three studies.

TABLE 1
Summary of Gender Differences Across Sleep Deprivation Studies

Dependent Measure	54-h TSD with naps	64-h TSD	7 Days PSD
Visual Reaction Time (Median response time)	Gender (F(1,28) = 5.91, p = .022)		
	Gender x Group (F(4,28) = 2.86, p = .042)		
	Trials x Gender (F(14,392) = 2.49,p = .002)		
	Trials x Gender x Group (F(56,392) = 1.8, p = .001)		
Temperature		Days x Trials x Gender (F(1,22) = 3.46, p=.076)	
	Trials x Group x Gender (F(4,30) = 3.77, p=.013)	Days x Trials x Gender (F(1,22) = 4.58, p = .044)	
Digit Symbol		Days x Trials x Gender (F(1,22) = 5.12, p=.034)	
POMS fatigue		Trials x Gender (F(11,242) = 1.72, p = .07)	Time of day x Gender (F(2,28) = 2.76, p =.09)
		Trials x Days x Gender (F(11,242) = 2.15, p = .018)	
POMS Tension - Anxiety		Days x Gender (F(1,22) = 4.88, p = .095)	
POMS Anger - Hostility	Trials x Gender (F(14,434) = 2.04, p = .014)		
	Trials x Group x Gender (F(56,434) = 1.68, p = .003)		
POMS Depression - Dejection	Trials x Group x Gender (F(56,420) = 1.32, p =.068)		
POMS Total Mood Disturbance			Time of day x Gender (F(2,28) = 3.96, p = .03)
POMS Vigor			Time of day x Gender (F(2,28) = 4.41, p = .02)
Digit Span	Gender (F(1,31) = 3.69, p = .064)		
	Pre-Post Nap x Gender (F(1,31) = 5.29, p = .028)		
Probed Recall			Days x Gender (F(7,98) = 2.1, p=.05)

Report 6: January 15, 1997. Analyses undertaken for the current report continue to examine data collected in the course of an investigation of 54 h of total sleep deprivation in a volunteer, young adult sample consisting of 26 males and 15 females, with 2-h nap opportunities strategically placed following either 6, 18, 30, 42, or 54 h of total sleep loss. Naps occurring after 6, 30, and 54 h of participation in the sleep deprivation protocol took place near the peak of the circadian cycle of alertness, while those scheduled to occur following 18 and 42 h of continuous wakefulness took place near the circadian nadir. The principal objective of this investigation was to ascertain the optimal placement of a brief nap opportunity during a period of prolonged wakefulness, so as to ensure either prophylactic or restorative effects against the performance and alertness decrements associated with sleep loss. Major focuses of the current set of analyses were to continue our investigation of possible gender-related differences in response to sleep loss in the context of sustained operations, and to determine whether brief nap opportunities might differentially offset performance impairment for men as compared to women.

In the previous month's progress report (i.e., December 15, 1996), we found no statistically significant effects related to gender of the study participants for two tasks that are sensitive to decrements in primary memory capacity precipitated by sleep loss. One of these was a challenging mental arithmetic task, the Descending Subtraction Task (DST), while the other, the Memory and Search Task (MAST), examines subjects' ability to rapidly search long strings of alphanumeric characters for the presence of certain predesignated target stimuli. The level of difficulty of the task increases with the number of target stimuli, which can be either 2 or 6 stimuli. The failure to find gender effects with these measures was somewhat unexpected, given that gender differences had been previously reported for these very same subjects on another index of short-term memory performance (i.e., the Digit Span subtest of the Wechsler Adult Intelligence Scale), and a similar outcome was obtained using a probed-recall test in a sample of young adult volunteers who underwent 7 days of partial sleep deprivation. Accordingly, additional analyses of the MAST data were undertaken in an attempt to conduct a more comprehensive assessment of possible gender-related differences in memory performance following sleep loss.

These supplemental analyses of the MAST data examined rates of errors across sleep deprivation trials, rather than search times, which had been evaluated during the previous budget period. Both the 2-target ($\underline{F}(7,217) = 7.54$, $\underline{p} < .001$) and 6-target ($\underline{F}(7,217) = 7.04$, $\underline{p} < .001$) tasks were vulnerable to increasing error rates as sleep debt accumulated. Unlike the analyses based on search times, however, there were no significant main effects of napping condition nor napping condition x trials interactions. Nevertheless, for the 2-target task, a significant 3-way interaction (i.e., gender x napping condition x trials) was identified for the single degree-of-freedom polynomial contrast for linear trend ($\underline{F}(4,31) = 3.02$, $\underline{p} = .033$). Figures 1 and 2 show mean numbers of errors across trials as a function of napping condition for males and females, respectively, in the

2-target task condition. Because napping condition did not have an overall systematic effect on error rate, and performance was generally variable within a small range, the source of the interaction appears to be the difference between males and females in the P-6 napping condition. Thus, males in this condition exhibited a relatively low and stable rate of errors over trials, whereas their female counterparts showed a tendency to escalate errors during the latter trials.

A similar pattern was observed for the 6-target task condition, in which the repeated-measures ANOVA yielded a significant gender x napping condition x trials interaction ($F(28,217) = 1.54$, $p = .046$). Figures 3 (males) and 4 (females) graphically depict changes in mean number of errors over trials for each of the 5 nap-placement conditions. Variability within a modest range of error was again a characteristic feature of each of the napping conditions. The most salient exception, however, was the difference in performance between males and females in the P-6 condition. As in the case of the 2-target task, males in the P-6 group performed the more difficult 6-target MAST task with fewer errors over trials than did female subjects in the same napping condition, who showed a progressive increase in errors over each of the final 4 trials. In view of our previously reported findings that, irrespective of gender, subjects in the P-6 nap group exhibited longer search times as sleep loss increased, relative to most of the other napping conditions (with the exception of the P-54 group), these differences in error rates between males and females suggest that sleep-deprived males may attempt to conserve accuracy by adopting a slower, more deliberate, response strategy, whereas females may be dually challenged in both speed and accuracy outcomes with excessive sleep loss. Nevertheless, these findings were not particularly dramatic and did not replicate in many other tasks.

The final report for this project is in preparation. We are attempting to equilibrate metrics across the three major sleep-deprivation studies conducted by our laboratory, so as to permit large-scale analyses of those variables that appear to be sensitive to gender differences, as well as to take advantage of the substantial increase in statistical power afforded by combining these subject samples. With judicious selection of sampling points most representative of the extremes of sleep deprivation within each investigation, it should be possible to derive strong conclusions concerning gender effects and to relate these findings to the extant scientific literature relevant to human performance during scenarios of sustained wakefulness.

Figure 1

MAST 2-TARGET TASK/MALES

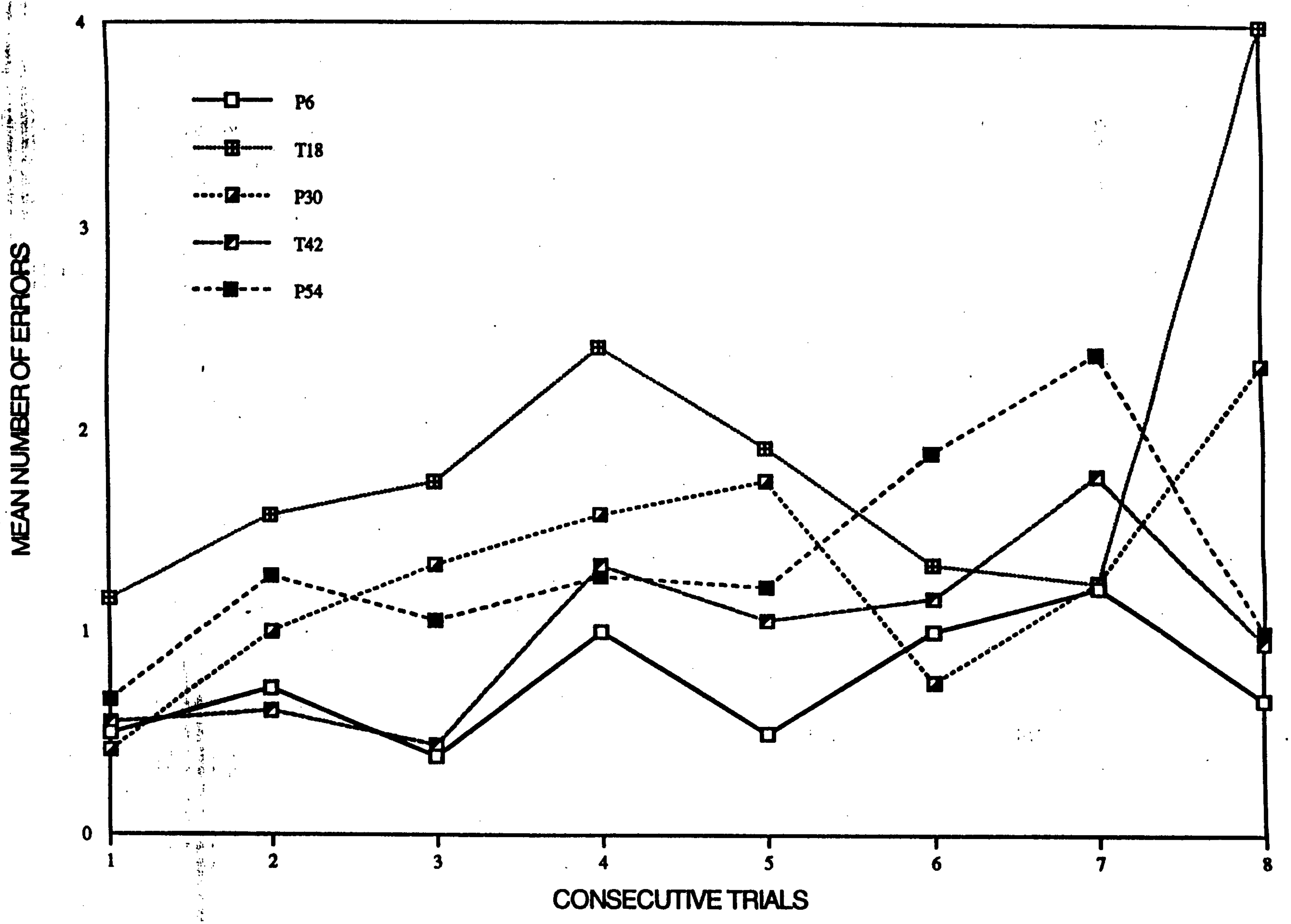


Figure 2

MAST 2-TARGET TASK/FEMALES

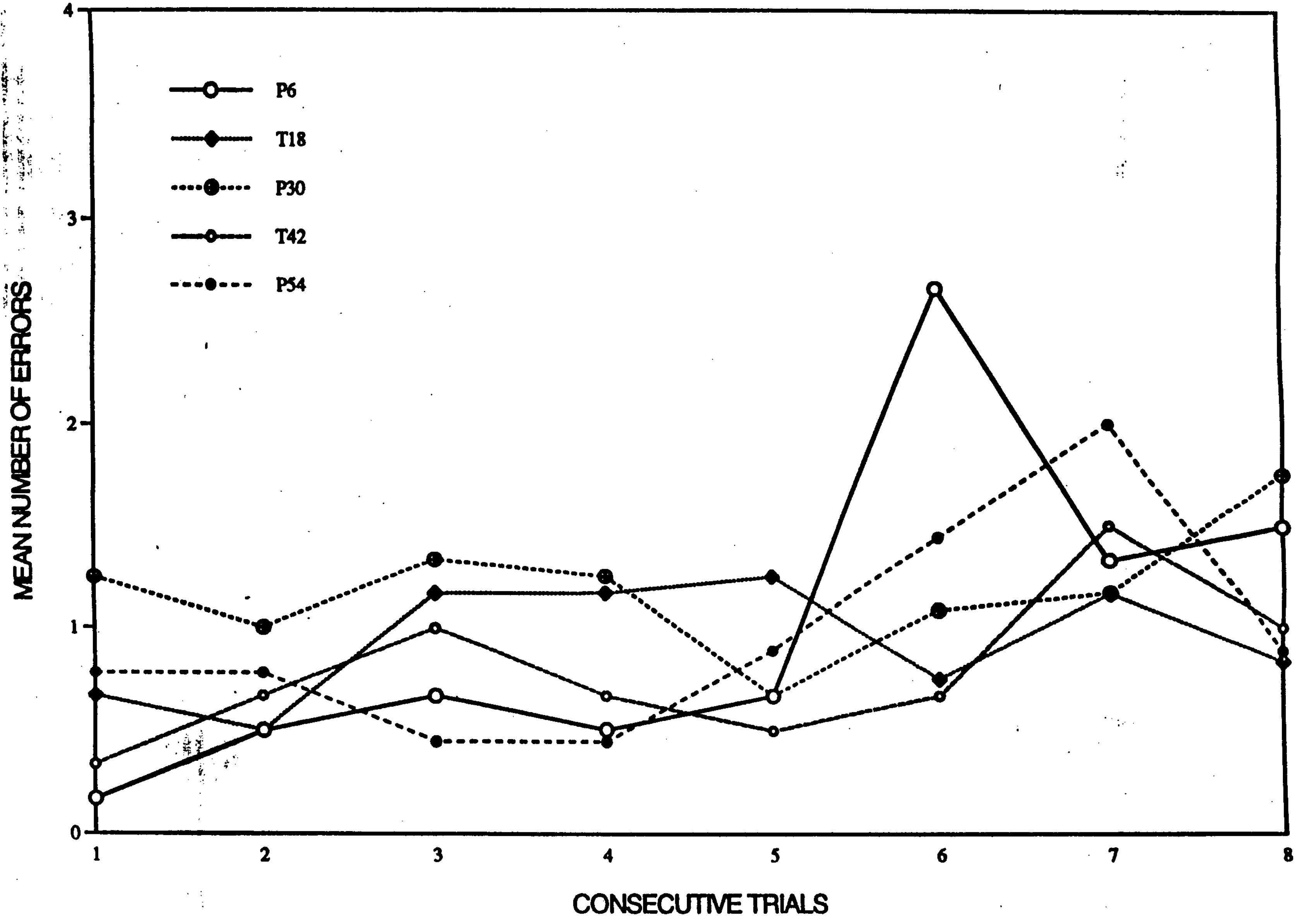


Figure 3

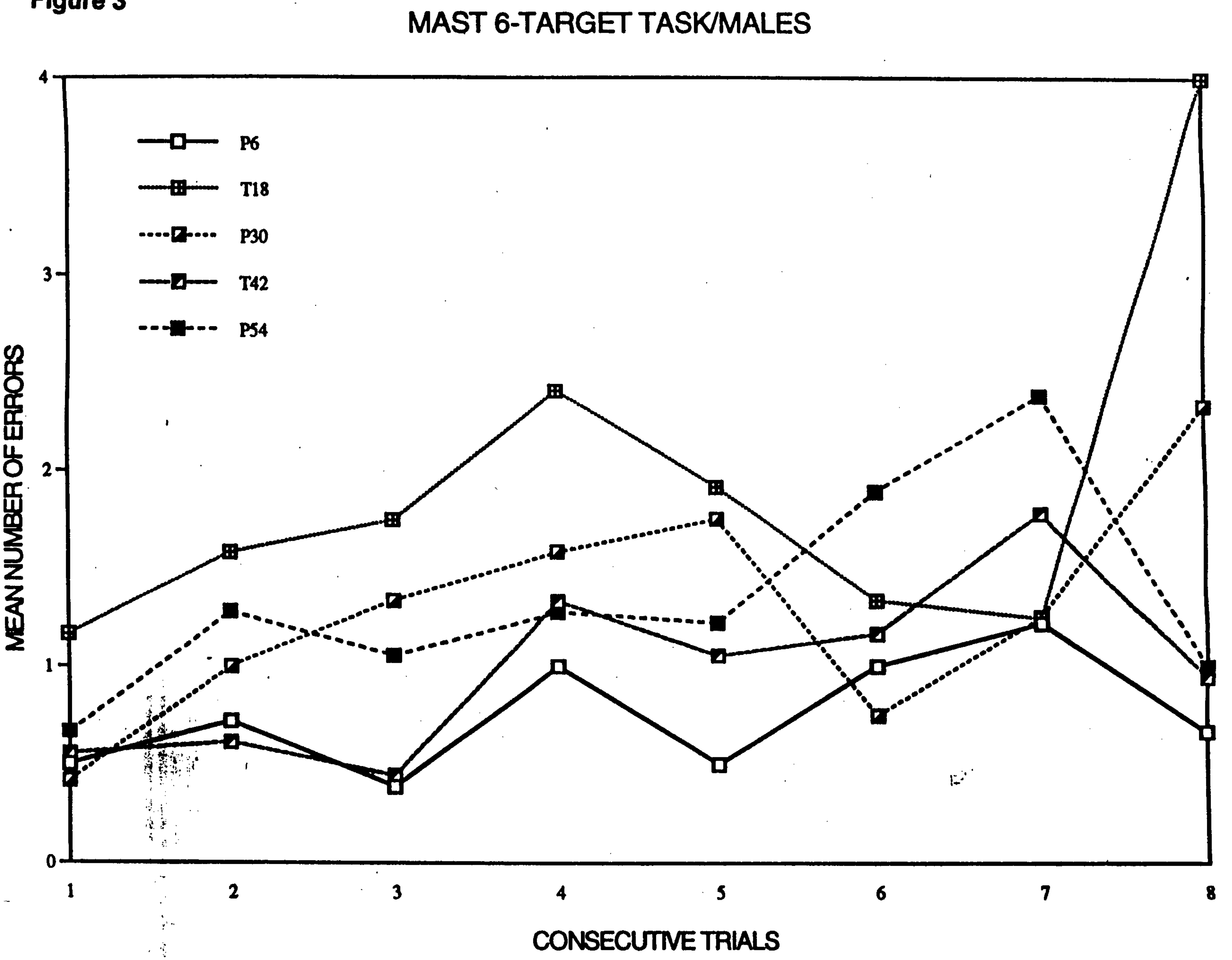
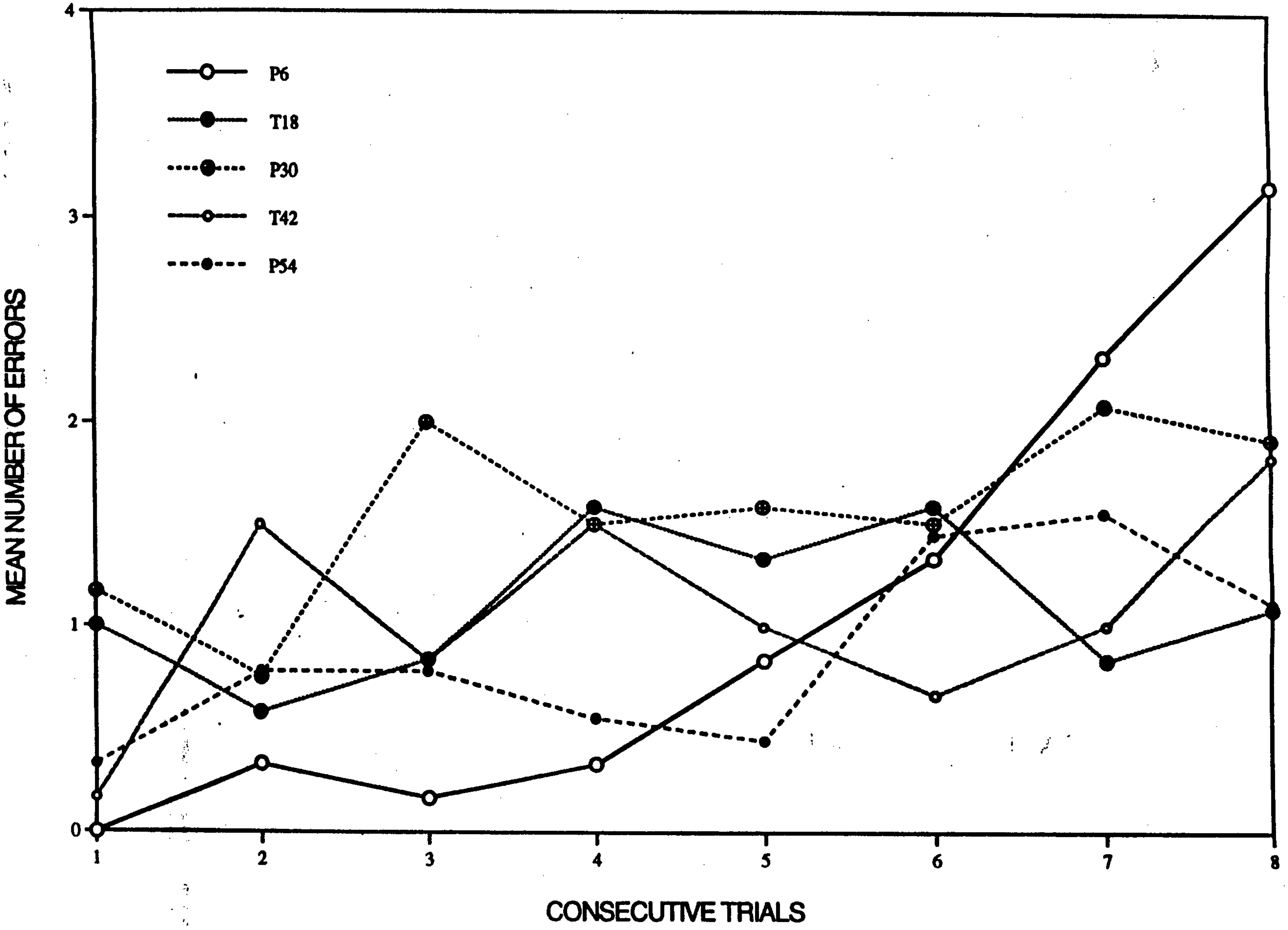

MAST 6-TARGET TASK/MALES
P6
T18
P30
T42
P54
MEAN NUMBER OF ERRORS
0
1
2
3
4
CONSECUTIVE TRIALS
1
2
3
4
5
6
7
8

Figure 4

MAST 6-TARGET TASK/FEMALES



Final Report: January 27, 1997.

BACKGROUND AND PROJECT OBJECTIVES

Military operations routinely require continuous manning around the clock. As a result, personnel must attempt to acclimate to unusual, and sometimes prolonged, duty schedules that can seriously compromise levels of alertness, performance, and safety. Although numerous laboratory investigations and field studies have been carried out, which document significant decrements in vigilance with sustained monitoring tasks, there is as yet no consensus on whether men and women are equally vulnerable to this effect (e.g., Davies & Parasuraman, 1982; Thackray, Touchstone, & Bailey, 1978; Waag, Halcomb, & Tyler, 1973). Moreover, there has been virtually no research dedicated to identifying potential gender differences in continuous performance scenarios that involve substantial sleep loss and the consequent disruption of endogenous biological rhythms.

The current project sought to address these issues by analyzing neurobehavioral data (i.e., performance, subjective activation, mood, and body temperature) from three sleep deprivation experiments, involving healthy young (age range = 21-37) adult males ($n = 51$) and females ($n = 30$), previously carried out by the Unit for Experimental Psychiatry, with a special focus on gender differences during night work and sustained wakefulness. The findings have relevance for the selection and training of military women and men to engage in continuous operations in which alertness is a key factor.

Two of the studies on which these analyses were based involved total sleep deprivation -- one being an investigation of the effects of 64 h of total sleep deprivation (17 males, 7 females; see Dinges, Douglas, Zaugg, Campbell, McMann, Whitehouse, E.C. Orne, Kapoor, Icaza, & M.T. Orne, 1994), while the other featured 2-h nap opportunities strategically placed at 6, 18, 30, 42, and 54 h into the 54-h sleep deprivation protocol (26 males, 15 females; see Dinges, M.T. Orne, Whitehouse, & E.C. Orne, 1987). A third study (8 males, 8 females; see Dinges, F. Pack, Williams, Gillen, Powell, Ott, Aptowicz, & A.I. Pack, in press) investigated the effects of 7 consecutive days of partial sleep deprivation (5 h per night).

In the Monthly Progress, Status and Management Reports for this project, we identified some evidence of gender differences within each of the 3 sleep deprivation studies. The major findings of these individual studies are reproduced in Table 1, which was originally included in the December 15, 1996 report.

In this final report we have standardized measures across the three sleep-deprivation studies, in order to undertake large-scale analyses of those variables that appear to be sensitive to gender differences, as well as to take advantage of the substantial increase in statistical power afforded by combining these subject samples. The dependent measures selected for analysis across studies were guided by the findings outlined in

Table 1. They consisted of various parameters of the psychomotor vigilance test (PVT), which is a visual reaction-time task administered in each of the 3 studies, cognitive performance (comprised either of probed recall, Digit Symbol Substitution Test, and Memory and Search Test accuracy scores), and the Profile of Mood States (POMS), which was also available for each study. Investigation of male-female body temperature differences across studies was not pursued because this variable was not measured in the 7-day partial sleep deprivation study and because increases in female body temperature may be due to a rise in progesterone levels during the luteal phase of the menstrual cycle (Kattapong, Fogg, & Eastman, 1995). Sampling points for the cross-study analysis were selected to represent conditions of baseline, midpoint of sleep deprivation, and peak of sleep deprivation for each study. For the 64-h total sleep deprivation study, the times selected were 1000 h, 1600 h, and 2200 h during baseline and each of the following two days of sleep deprivation; for the 54-h total sleep deprivation study with 2-h naps, testing times deviated slightly on an individual basis from these target sampling points, therefore sessions closest in time to the target points were selected; for the 7-day partial sleep deprivation study, each of the three target time periods were included for the baseline day, the third day, and the seventh day of partial sleep loss.

RESULTS

Repeated-measures multivariate analyses of variance (MANOVAs) applied to the various parameters of the PVT data revealed a significant between-subjects main effect of gender ($\underline{F}(1,74) = 4.54$, $\underline{p} = .036$), and an overall days x time interaction (Wilks' Lambda = 0.77, $\underline{F}(4,71) = 5.29$, $\underline{p} = .001$) for median response times, both of which can be observed in Figure 1. Thus, the median response time of males was significantly faster than that of females at each test session, although cumulative sleep loss had a comparably debilitating impact on the performance of both males and females.

Analysis of the fastest 10% of PVT responses per session also yielded a significant main effect of gender ($\underline{F}(1,74) = 6.30$, $\underline{p} = .014$), as well as a significant gender x days interaction (Wilks' Lambda = 0.90, $\underline{F}(2,73) = 3.91$, $\underline{p} = .024$). These data are presented in Figure 2, which reveals faster response times by males than by females overall, as well as a tendency for female subjects to exhibit a somewhat greater performance decrement in the late morning testing session that corresponds to the peak period of sleep deprivation.

Analyses of other aspects of PVT performance, such as the number of lapses and the slowest 10% of responses per session, failed to reveal any significant gender differences. In both cases, however, the days x time interaction was reliable (Wilks' Lambda = 0.78, $\underline{F}(4,71) = 5.15$, $\underline{p} = .001$ for lapses; Wilks' Lambda = 0.72, $\underline{F}(4,70) = 6.95$, $\underline{p} < .001$ for slowest 10% of responses), reflecting greater deterioration of performance during those testing sessions that were conducted near the peak of sleep deprivation.. Nevertheless, the failure to find gender differences in the effects of sleep deprivation on lapses and other components of PVT performance known to be sensitive to sleep loss suggests that those differences that were obtained for median and fastest response times are not associated with vulnerability to sleep deprivation. We believe these differences, which are apparent even during baseline, reflect the somewhat competitive manner in which males perform the task relative to females (i.e., males tended more than females to compete for the fastest response times in the experiments that make up the current database).

The cognitive performance measures also failed to differentiate male and female participants across the sampling time points represented in the current analyses. However, for the cognitive measures, the interaction of days with time was, once again, significant (Wilks' Lambda = 0.75, F(4,72) = 6.13, p < .001), revealing the expected deleterious effect of extreme sleep loss on cognitive tasks.

Gender differences were obtained on two subscales of the POMS. Analyses of the Anger-Hostility dimension yielded a significant three-way interaction involving gender, days, and time (Wilks' Lambda = 0.84, F(4,71) = 3.33, p = .015). The effect is evident in Figure 3, wherein the gender differences that emerged during repeated baseline testing were eliminated with additional time-in-study, being replaced by a general elevation of the negative mood state for both males and females during the peak of sleep deprivation. In addition, the three-way interaction of gender x days x time was also significant for the Depression-Dejection subscale of the POMS (Wilks' Lambda = 0.86, F(4,64) = 2.60, p = .044). Figure 4 portrays these data, which indicate greater variability and generally higher levels of depressive affect among female participants relative to males throughout these investigations. It is important to recognize, however, that although these effects were statistically reliable, they represent minor differences in mood states as a function of gender status. To elucidate this caveat further, note that in Figure 3, the range from 0 to 2 scale points along the ordinate manages to capture the extent of variation in mean Anger-Hostility scores for males versus females across the three sleep deprivation studies, whereas scores for this dimension have the potential to range from 0 to 48. Similarly, the ordinate in Figure 4 extends from 0.5 to 3, while the potential range for Depression-Dejection scores is from 0 to 60.

Gender differences were not evident on any of the other subscales of the POMS. With the exception of the Tension-Anxiety dimension, which showed no significant variation across sampling points, analyses of these remaining subscales revealed each to be sensitive to the effects of sleep deprivation as expressed in the form of a days x time interaction (Confusion-Bewilderment: Wilks' Lambda = 0.83, F(4,72) = 3.83, p = .007; Fatigue: Wilks' Lambda = 0.81, F(4,72) = 4.23, p = .004; Vigor: Wilks' Lambda = 0.79, F(4,72) = 4.69, p = .002). As in the case of previously discussed measures in which a days x time interaction was observed, these instances in self-reported mood and subjective activation reflect exacerbations (or a decrement for vigor) associated with the peak period of sleep deprivation. Given the demonstrated sensitivity of the POMS to sleep loss, coupled with an absence of gender effects for those subscales that characteristically change with sleep deprivation (i.e., Fatigue, Vigor, Tension-Anxiety, and Confusion-Bewilderment), it is all the more prudent to discount the small gender differences observed for Anger-Hostility and Depression-Dejection (i.e., they do not represent meaningful mood disturbances precipitated by sleep deprivation).

CONCLUSIONS

In an extensive database derived from our extant program of research on the effects of sleep deprivation and night operations in healthy young adults, few reliable differences between men and women were observed. The omnibus analyses reported here, which aggregated subjects across three independent studies, selectively assessed performance, cognitive throughput, and mood/subjective activation at three times per day on each of three days, representing periods of baseline, the midpoint of sleep deprivation, and the peak of sleep deprivation. It is possible that the specific dependent measures, or sampling time points, that were selected for the omnibus analyses were not sufficiently sensitive to the expression of gender differences in response to sleep deprivation. However, even the more extensive and detailed analyses, which were possible within each of the three sleep deprivation studies individually, did not reveal gender effects that replicated across studies (see Table 1).

Although results from the omnibus analyses revealed differences between males and females on the PVT in the domains of median and fastest response times -- differences that remained reliable even during the peak period of sleep deprivation -- the observations that these differences originated during the baseline period and were not paralleled by differences in lapses with heightened sleep loss, suggest that they are not important indices of vulnerability to sleep deprivation. Rather, they appear to reflect a more competitive response style among men relative to women in these laboratory studies, which persists despite increasing pressure to sleep.

Similarly, findings of gender differences in self-reported anger and depression with progressively greater sleep loss were small in magnitude and do not constitute evidence for mood disturbance among healthy young adults. Variation in levels of fatigue, vigor, and confusion corresponded to the expected pattern associated with cumulative sleep loss, however, no male-female differences were evident for any of these measures.

Thus, alertness appears to be comparable for men and women across the circadian cycle and as a function of the number of hours of sleep deprivation. When statistically reliable differences did emerge, they tended to be small and did not reflect a compelling superiority for either gender in alertness. Two notable exceptions involved gender differences detected following prophylactic naps and recovery sleep, which were reviewed by us in earlier reports. In the 54-h total sleep deprivation protocol with 2-h nap opportunities, changes in short-term memory performance from prenap to postnap suggested that females experienced less sleep inertia than did males. On the other hand, in the 64-h total sleep deprivation study, males reported significantly less fatigue following recovery sleep than their female counterparts. More research will be needed to determine if these post-sleep differences have any bearing on the extent to which either gender can tolerate protracted loss of sleep.

REFERENCES

Davies, D.R., Parasuraman, R.: The Psychology of Vigilance, London: Academic Press, 1982.

Dinges, D.F., Orne, M.T., Whitehouse, W.G., Orne, E.C.: Temporal placement of a nap for alertness: Contributions of circadian phase and prior wakefulness. Sleep 10:313-329, 1987.

Dinges, D.F., Douglas, S.D., Zaugg, L., Campbell, D.E., McMann, J.M., Whitehouse, W.G., Orne, E.C., Kapoor, S.C., Icaza, E., Orne, M.T.: Leukocytosis and natural killer cell function parallel neurobehavioral fatigue induced by 64 h of sleep deprivation. Journal of Clinical Investigation 93:1930-1939, 1994.

Dinges, D.F., Pack, F., Williams, K., Gillen, K.A., Powell, J.W., Ott, G.E., Aptowicz, C., Pack, A.I.: Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. Sleep, in press.

Kattapong, K.R., Fogg, L.F., Eastman, C.I.: Effect of sex, menstrual cycle phase, and oral contraceptive use on circadian temperature rhythms. Chronobiology International 12:257-266, 1995.

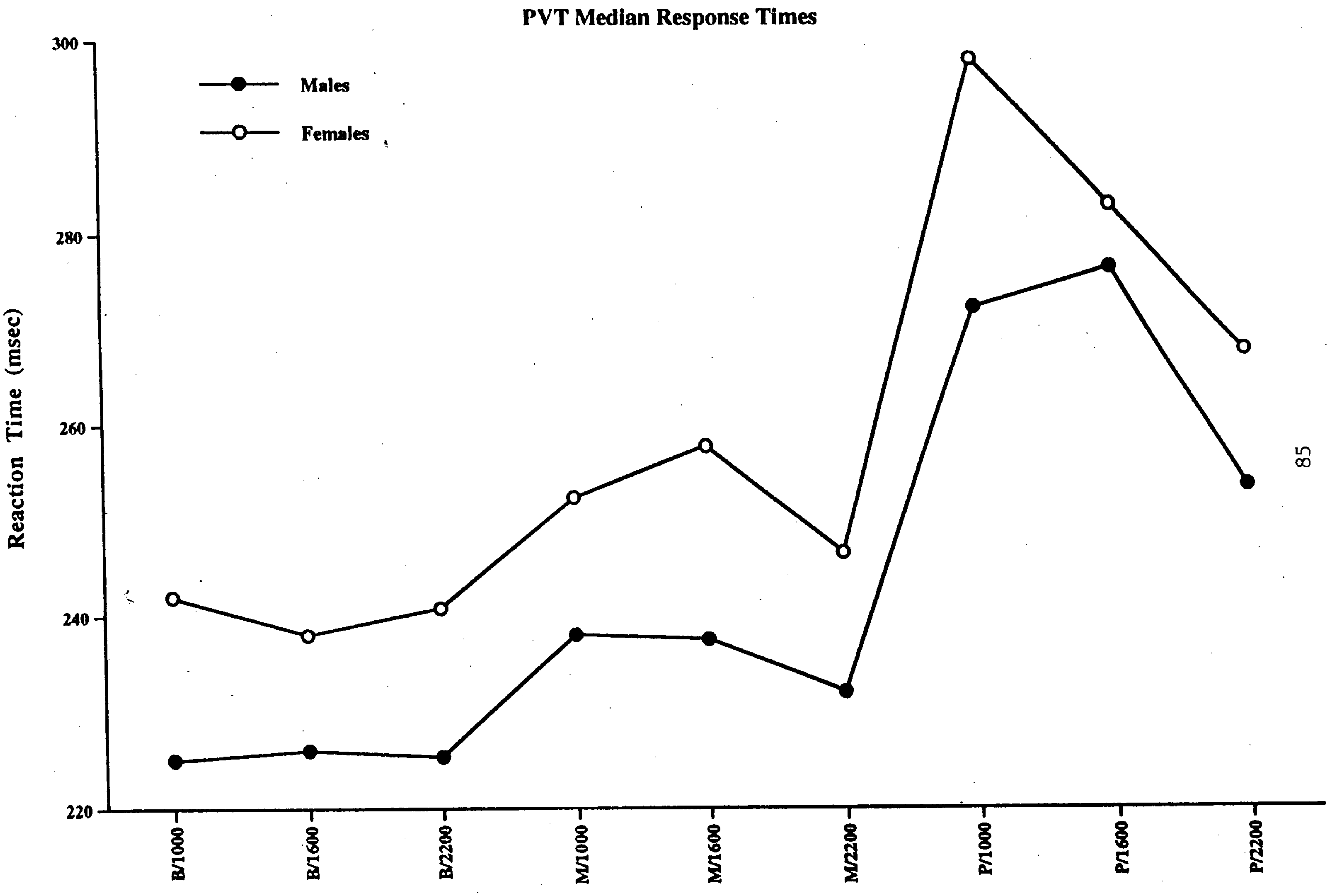
Thackray, R.I., Touchstone, R.M., Bailey, J.P.: Comparison of the vigilance performance of men and women using a simulated radar task. Aviation, Space & Environmental Medicine 49:1215-1218, 1978.

Waag, W.L., Halcomb, C.G., Tyler, D.M.: Sex differences in monitoring performance. Journal of Applied Psychology 58:272-274, 1973.

TABLE 1
Summary of Gender Differences Across Sleep Deprivation Studies

Dependent Measure	54-h TSD with naps	64-h TSD	7 Days PSD
Visual Reaction Time (Median response time)	Gender (F(1,28) = 5.91, p = .022)		
	Gender x Group (F(4,28) = 2.86, p = .042)		
	Trials x Gender (F(14,392) = 2.49,p = .002)		
	Trials x Gender x Group (F(56,392) = 1.8, p = .001)		
Temperature		Days x Trials x Gender (F(1,22) = 3.46, p=.076)	
	Trials x Group x Gender (F(4,30) = 3.77, p=.013)	Days x Trials x Gender (F(1,22) = 4.58, p = .044)	
Digit Symbol		Days x Trials x Gender (F(1,22) = 5.12, p=.034)	
POMS fatigue		Trials x Gender (F(11,242) = 1.72, p = .07)	Time of day x Gender (F(2,28) = 2.76, p =.09)
		Trials x Days x Gender (F(11,242) = 2.15, p = .018)	
POMS Tension - Anxiety		Days x Gender (F(1,22) = 4.88, p = .095)	
POMS Anger - Hostility	Trials x Gender (F(14,434) = 2.04, p = .014)		
	Trials x Group x Gender (F(56,434) = 1.68, p = .003)		
POMS Depression - Dejection	Trials x Group x Gender (F(56,420) = 1.32, p =.068)		
POMS Total Mood Disturbance			Time of day x Gender (F(2,28) = 3.96, p = .03)
POMS Vigor			Time of day x Gender (F(2,28) = 4.41, p = .02)
Digit Span	Gender (F(1,31) = 3.69, p = .064)		
	Pre-Post Nap x Gender (F(1,31) = 5.29, p = .028)		
Probed Recall			Days x Gender (F(7,98) = 2.1, p=.05)

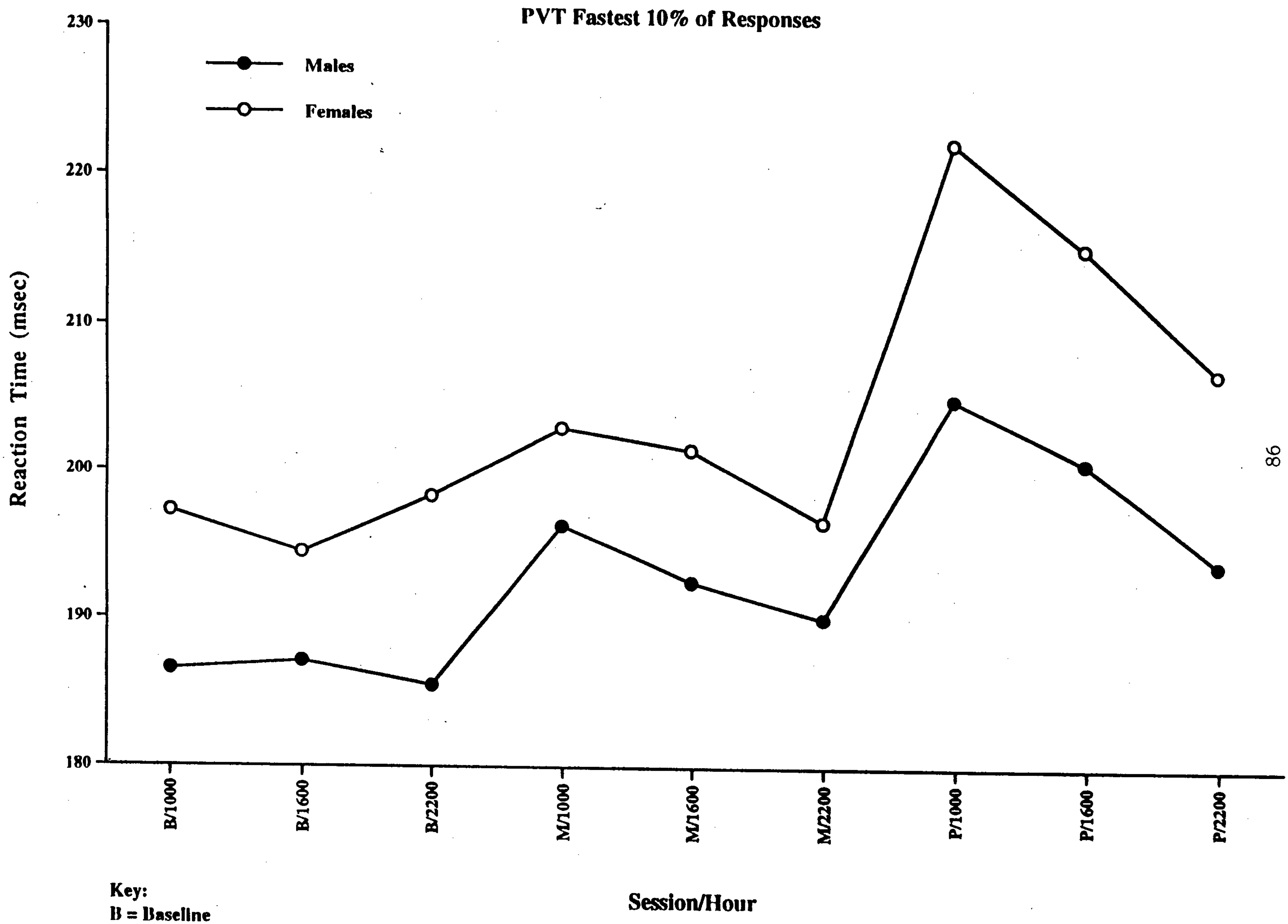
Figure 1



Session/Hour

Key:
B = Baseline
M = Midpoint of Sleep Deprivation
P = Peak of Sleep Deprivation

Figure 2



Key:
B = Baseline
M = Midpoint of Sleep Deprivation
P = Peak of Sleep Deprivation

Figure 3

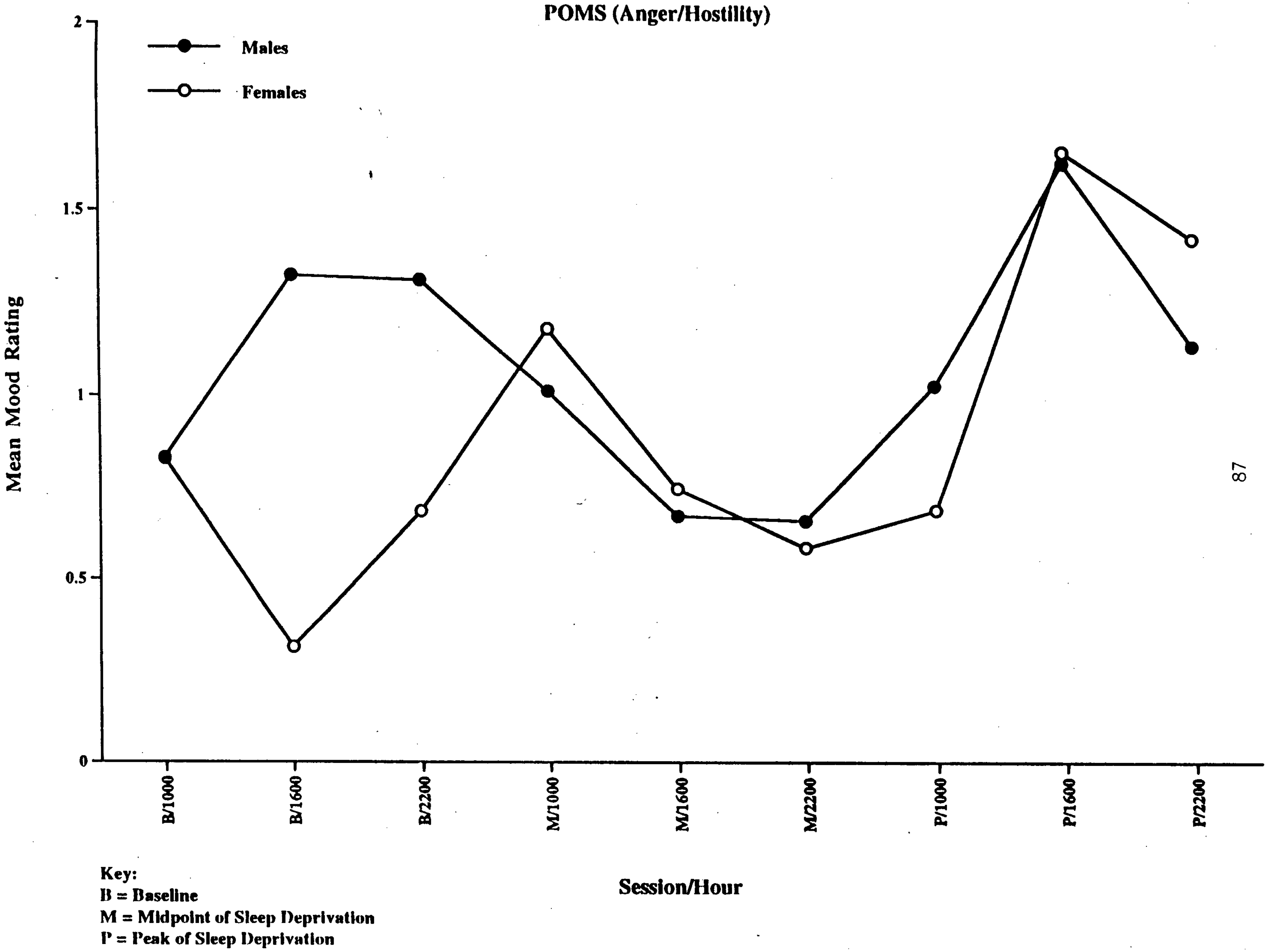
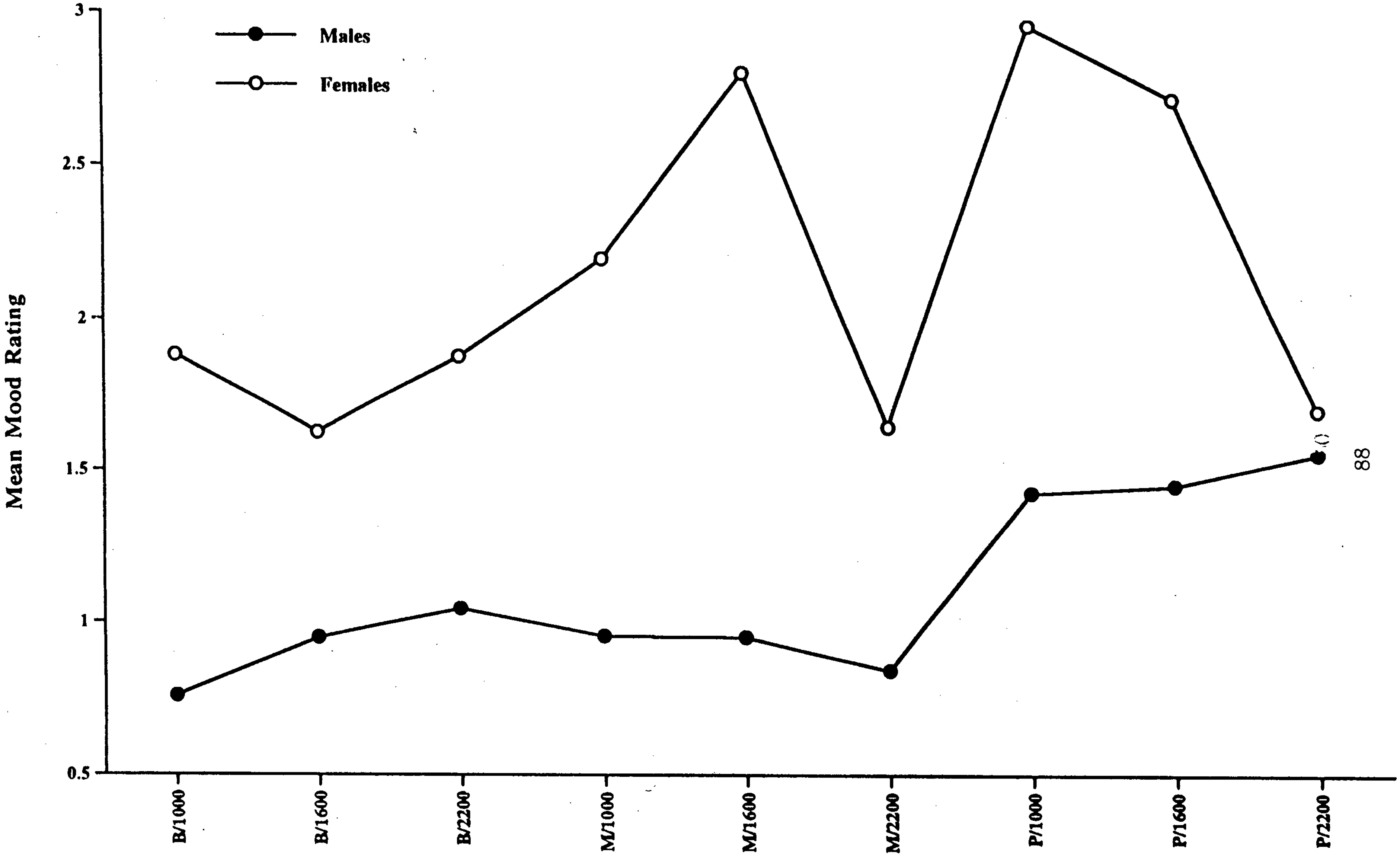

POMS (Anger/Hostility)
Males
Females
Mean Mood Rating
2
1.5
1
0.5
0
B/1000
B/1600
B/2200
M/1000
M/1600
M/2200
P/1000
P/1600
P/2200
Session/Hour
Key:
B = Baseline
M = Midpoint of Sleep Deprivation
P = Peak of Sleep Deprivation

Figure 4

POMS (Depression-Dejection)



Session/Hour

Key:
B = Baseline
M = Midpoint of Sleep Deprivation
P = Peak of Sleep Deprivation

CONCLUSIONS

Examined in their entirety, the results from the studies by the Naval Health Research Center (NHRC) and the Unit for Experimental Psychiatry (UEP) are similar and consistent. There were but a few cognitive performance effects reaching significance in the omnibus data analyses conducted by UEP (aggregating subjects from three separate studies) and in the single study conducted by NHRC. The differences consisted of faster responding by the men on the Psychomotor Vigilance Task (UEP) and a similar simple reaction time task (NHRC), as well as better performance on a tracking task (NHRC). In neither study was this gender difference accompanied by any significant differences in lapses or other response metrics strongly correlated with sleep loss. Furthermore, the differences seen during sleep deprivation frequently were evident during baseline measures (UEP) or during training and the first few sessions of the experiment (NHRC). Therefore, it is unlikely that the reported gender differences are associated with sleep deprivation. The conclusion from the UEP study is that the reaction time differences likely reflect a more competitive response style among the men. This style difference apparently persists even in the presence of substantial sleep deprivation in the laboratory testing environment. It is quite possible that, in the NHRC study, the reaction time differences and the difference in the long and monotonous tracking measure represent this same response style difference. Finally, differences seen in the subjective sleepiness and mood measures from both NHRC and UEP represent either minor differences in mood state (UEP) or differences that become smaller with increasing sleep deprivation (NHRC).

To conclude, the extensive data from the NHRC and UEP studies are consistent in failing to detect any reliable differences in cognitive performance and alertness between men and women as a function of sleep loss. Although there may be some gender-related response style differences, the present data do not support distinguishing between men and women in their ability to withstand the effects of sleep deprivation.

REFERENCES

Akerstedt T, & Gillberg M. Subjective and objective sleepiness in the active individual. *Inter-national Journal of Neuroscience*, 1990, 52:29-37.

Beatty J, Greenberg A, Deibler WP, & O'Hanlon JF. Operant control of occipital theta rhythm affects performance in a radar monitoring task. *Science*, 1974, 183:871.

Belyavin A, & Wright NA. Changes in electrical activity of the brain with vigilance. *Electroencephalography and Clinical Neurophysiology*, 1987, 66:137-144.

Blough PM, & Slavin LK. Reaction time assessments of gender differences in visual-spatial performance. *Perception & Psychophysics*, 1987, 41(3):276-281.

Churchland, P. and Sejnowski, T. Perspectives on cognitive neuroscience. *Science*, 1988, 241:741-745.

Cooper LA. Mental rotation of random two-dimensional shapes. *Cognitive Psychology*, 1975, 7:20-43.

Davies DR, & Parasuraman R. *The Psychology of Vigilance*. London: Academic Press, 1982.

Dember WN, Galinsky TL, & Warm JS. The role of choice in vigilance performance. *Bulletin of the Psychonomic Society*, 1992, 30(3):201-204.

Dinges DF. Probing the limits of functional capability: The effects of sleep loss on short-duration tasks. In RJ Broughton & RD Ogilvie (Eds.), *Sleep, Arousal, and Performance*, Boston: Birkhauser, 1992, 176-188.

Dinges DF, Gillen K, Powell J, Carlin M, Ott G, Orne EC, & Orne M. Discriminating sleepiness by fatigueability on a psychomotor vigilance task [Abstract]. *Sleep Research*, 1994, 23:129.

Dinges DF, & Kribbs NB. Performing while sleepy: effects of experimentally-induced sleepiness. In TH Monk (Ed.), *Sleep, Sleepiness and Performance*, Chichester, UK: Wiley, 1991, 97-128.

Dinges DF, & Powell JW. Sleepiness is more than lapsing [Abstract]. *Sleep Research*, 1988, 17:84.

Fruhstorfer H, & Bergstrom. Human vigilance and auditory evoked responses. *Electroencephalography and Clinical Neurophysiology*, 1969, 27:346-355.

Geheb R, Whitfield KE, & Brannon L. Effect of visual complexity in identification of tachistoscopic images. *Perceptual and Motor Skills*, 1994, 78:971-978.

Holding DH, & Baker MA. Toward meaningful noise research. *Journal of General Psychology*, 1987, 114:395-410.

Institute of Medicine, Committee on Defense Women's Health Research. *Recommendations for Research on the Health of Military Women.* Washington, DC: National Academy Press, 1995.

Jokeit H & Makeig S. Different event-related patterns of gamma-band power in brain waves of fast- and slow-reacting subjects. *Proceedings of the National Academy of Sciences USA*, 1994, 91:6339-6343.

Klein KE, Wegmann HM, Athanassenas G, Hohlweck H, & Kuklinski P. Air operations and circadian performance rhythms. *Aviation, Space, and Environmental Medicine,* 1976, 47:221-230.

Kirk RE. *Experimental Design: Procedures for the Behavioral Sciences* (2nd ed.), Pacific Grove, CA: Brooks/Cole, 1982.

Kribbs NB, & Dinges DF. Vigilance decrement and sleepiness. In: JR Harsh & RD Ogilvie (Eds.), *Sleep Onset Mechanisms*, Washington, DC: American Psychological Association, 1994, 113-125.

Loomis AL, Harvey E, & Hobart GA. Cerebral states during human sleep as studied by human brain potentials. *Journal of Experimental Psychology*, 1937, 21:127-144.

Mackworth NH. The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1948, 1:6-21.

Makeig S, Elliott FS, Inlow M, & Kobus D. *Predicting Lapses in Vigilance Using Brain Evoked Responses to Irrelevant Auditory Probes* (NHRC Technical Report 90-39). San Diego, CA: Naval Health Research Center, 1990.

Makeig S, Elliott FS, & Postal M. *First Demonstration of an Alertness Monitoring/Management System* (NHRC Technical Report No. 93-36). San Diego, CA: Naval Health Research Center, 1993.

Makeig S, & Inlow M. Lapses in alertness: coherence of fluctuation in performance and the EEG spectrum. *Electroencephalography and Clinical Neurophysiology*, 1993, 86:23-35.

Makeig S, & Jolley K. COMPTRACK. *A Compensatory Tracking Task for Monitoring Alertness* (NHRC Technical Report). San Diego, CA: Naval Health Research Center (in press).

Makeig S, & Jung T. Tonic and phasic EEG correlates of auditory awareness in drowsiness. *Cognitive Brain Research*, in press.

Mann VA, Sasanuma, S, Sakuma, & Masaki S. Sex differences in cognitive abilities: a cross-cultural perspective. *Neuropsychologia*, 1990, 28:1063-1077.

Matousek M & Peterson I. A method for assessing alertness fluctuations from EEG spectra. *Electroencephalography and Clinical Neurophysiology*, 1983, 55:108-113.

Ogilvie RD, Simons IA, Kuderian RH, MacDonald T, & Rustenburg J. Behavioral event-related potential, and EEG/FFT changes at sleep onset. *Psychophysiology*, 1991, 28:54-64.

Penzel T & Petzold J. A new method for the classification of subvigil stages, using the Fourier transform, and its application to sleep apnea. *Journal of Computational Biology and Medicine*, 1989, 19:7-34.

Rechtschaffen A & Kales A (Eds.) *Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects.* Bethesda, MD: Public Health Service, 1968.

Rosvold HE, Mirsky AF, Sarason I, Bransome ED, & Beck LH. A continuous performance test of brain damage. *Journal of Consulting and Clinical Psychology*, 1956, 20:343-350.

Santamaria J, & Chiappa KH. The EEG of drowsiness in normal adults. *Journal of Clinical Neurophysiology*, 1987, 4:327-382.

Seashore CE, & Kent GH. Periodicity and progressive change in continuous mental work. *Journal of Psychology Reviews*, 1905, 6:47-101.

Shaywitz BA, Shaywitz SE, Pugh KR, Constable RT, Skudlarski P, Fulbright RK, Bronen RA, Fletcher JM, Shankweller DP, Katz L, & Gore JC. Sex differences in the functional organization of the brain for language. *Nature*, 1995, 373:607-609.

Smith A. Consistent sex differences in a specific (decoding) test performance. *Educational and Psychological Measurement,* 1967, 27:1077-1083.

Smith CS, Reilly C, & Midkiff K. Evaluation of three circadian rhythm questionnaires with suggestions for an improved measure of morningness. *Journal of Applied Psychology*, 1989, 74:728-738.

Stebel J, & Sinz, R. On central nervous minute-periodicity and its coordination. *Journal of Interdisciplinary Cycle Research*, 1971, 2:63-72.

Stroud JM. The fine structure of psychological time. *Annals of the New York Academy of Sciences*, 1966, 138:623-631.

Thackray RI, Touchstone RM, & Bailey JP. Comparison of the vigilance performance of men and women using a simulated radar task. *Aviation, Space, and Environmental Medicine*, 1978, 49:1215-1218.

Tong JE, Leigh, F, Campbell J, & Smith D. Tobacco smoking, personality and sex factors in auditory vigilance performance. *British Journal of Psychology*, 1977, 68:365-370.

Torsvall L, & Akerstedt T. Extreme sleepiness: Quantification of EOG and EEG parameters. *International Journal of Neuroscience*, 1988, 38:435-441.

Townsend RE, & Johnson LC. Relation of frequency-analyzed EEG to monitoring behavior. *Electroencephalography and Clinical Neurophysiology*, 1979, 47:272-279.

Treisman, M. Temporal rhythms and cerebral rhythms. In: J Gibbon and L Allen (Eds.), *Timing and Time Perception*, New York Academy of Sciences, 1984, 423:542-565.

U.S. Congress, Office of Technology Assessment. *Biological rhythms: implications for the worker* (OTA-BA-463). Washington, DC: U.S. Government Printing Office, 1991.

Waag WL, Halcomb CG, & Tyler DM. Sex differences in monitoring performance. *Journal of Applied Psychology*, 1973, 58:272-274.

Warner RM. Periodic rhythm in conversational speech. *Language and Speech*, 1979, 22:381-396.

Wertheimer M. An investigation of the "randomness" of threshold measurements. *Journal of Experimental Psychology*, 1953, 45:294-301.

Weschlser D. *The measurement and appraisal of adult intelligence*. (4th ed.). Baltimore, MD: Williams and Wilkins, 1958.

Williams HL, Granda AM, Jones RC, Lubin A, & Armington JC. EEG frequency and finger pulse volume as predictors of reaction time during sleep loss. *Electroencephalography and Clinical Neurophysiology*, 1962, 14:64-70.

APPENDIX: LIST OF PERSONNEL RECEIVING PAY FROM THIS CONTRACT

Naval Health Research Center
Scott Makeig, Ph.D.
Tamsin L. Kelly, M.D.
David Ryman
Matthew Sinclair
Keith Jolley
Shawn Wing
Andrew Spydell

Unit for Experimental Psychiatry Research Foundation
David F. Dinges, Ph.D.
W. Whitehouse, Ph.D.
Janet Mullington, Ph.D.
John Powell
Emily Orne
B. R. Barras
Edward Mooney