

# Dynamic changes in work/rest duty cycles in a study of sleep deprivation

HARVEY BABKOFF

*Bar-Ilan University, Ramat Gan, Israel*

and

DAVID R. THORNE, HELEN C. SING, SANDER G. GENSER,  
STEVEN L. TAUBE, and FREDERICK W. HEGGE

*Walter Reed Army Institute of Research, Washington, D. C.*

The effects of moderate workload and 72 h of sleep deprivation were studied using a modified continuous-performance paradigm. Ten subjects were tested hourly on a number of perceptual and cognitive tasks designed to require approximately 30 min to complete, with the remainder of each hour free. As sleep deprivation continued, the average time on task increased at an accelerating rate. The rate of increase differed among tasks, with longer tasks showing greater absolute and relative increases than shorter ones. Such increases confound sleep deprivation and workload effects. In this paper, we compare the advantages and disadvantages of several experimental paradigms; describe details of the present design; and discuss methodological problems associated with separating the interactions of sleep deprivation, workload, and circadian variation with performance.

During the past 40 years, laboratory and field researchers have regularly reported decrements in performance and changes in psychological variables during and following periods of complete sleep deprivation of one or more days. Although the effects are neither trivial nor self-evident, there is a lack of agreement as to their extent and meaning. The decrements are dependent upon the interaction of stimulus with contextual, subjective, and individual environmental factors. For example, Wilkinson (1964, 1969) and Wilkinson, Edwards, and Haines (1966) have reported that, although subjects' performances on a simple vigilance task requiring sustained attention may be impaired following partial sleep deprivation, performances on an interesting complex task may show no deficit for subjects lacking sleep for up to 42 h. Furthermore, performance is dependent upon circadian and other rhythmic influences, so there is greater relative deficit between the hours of 2:00 and 6:00 a.m. than

at other times during the day (Johnson & Naitoh, 1974; Meddis, 1982).

Various paradigms have been used to study sleep loss and performance (see discussion by Angus & Heslegrave, 1985). At one end of the continuum are studies which mainly manipulate sleep deprivation, but which give little emphasis to performance tasks (e.g., Alluisi, 1972; Johnson, 1969; Loveland & Williams, 1963; Williams & Lubin, 1967). The cognitive workload and demands on the subjects are light or few. Selected tasks of either long or short duration are presented at specified times (often hours apart) during the experimental period. On the other end of the continuum are studies using a continuous-performance schedule, in which workload is maximal and performance is measured continuously (e.g., Angus & Heslegrave, 1985; Mullaney, Kripke, Fleck, & Johnson, 1983). Both types of experimental protocol pose difficulties in evaluating performance decrement. Because the findings of performance degradation are dependent upon the types and lengths of the tasks, it is clearly preferable to assess behavior by a variety of tests. In a minimal workload paradigm, testing at widely spaced intervals does not usually provide for continual testing of a variety of tasks. Data in a minimal workload paradigm are usually sufficient for assessing the inherent rhythmic variations that tend to influence and interact with the performance decrements reported in sleep loss. On the other hand, with a sustained operation maximal workload paradigm, one may be studying fatigue rather than sleep deprivation, because subjects may not suffer appreciable sleep loss before performance fails. For example, when subjects were exposed to a continuous three-component vigilance task (breaks only to attend to basic needs) in a planned 48-h study,

Acknowledgement is hereby made of the superior technical support provided in the running of the experiment by Jeffrey Clemans, Darwin Darr, Debra Friedman, Joseph Fritz, Timothy Gawne, Catherine Lovelless, Mary Pat May, Lisa Michaels, Mozelle Silvers Parker, Helena Prince, Joel Robertson, Barbara Ross, and Marjorie Taylor. Thanks again are due to Joseph Fritz who, as usual, is responsible for the fine quality of the graphic illustrations. Our deep gratitude is expressed to Teresa Benitez and Sylvia Sachs who with much patience and competence transformed this document from handwriting to legible form. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense (para 4-3, AR 360-5). H. Babkoff was supported under a contract awarded by the USA Medical R&D Command (Contract No. DAMD 17-84-M-9502). The authors' mailing address is: Department of Behavioral Biology, Walter Reed Army Institute of Research, Washington, DC 20307-5100.

no subject completed the study, and only 2 of 11 subjects completed 44 h of work (Mullaney et al., 1983; Orr, Hoffman, & Hegge, 1974).

The study reported here was designed to test the effects of sleep loss on a variety of psychological, physiological, and performance measures using a workload schedule intermediate between the two extremes described above. Several different performance measures were used to test perceptual-cognitive abilities. The subjects were tested quite rigorously and frequently, over periods of 48 to 72 h, in effect undergoing a sleep-loss, modified continuous-performance experimental schedule. The design of such a schedule allows for the assessment of the rhythmic variations inherent in performance, along with the expected monotonically increasing deficits due to continued sleep loss.

One of the factors seldom noted in previous studies is the possible interactive aspect of this type of experimental schedule. The effects of sleep loss and fatigue are assumed to be measured by changes in performance over the entire course of the sleep-deprivation experiment. However, the performance demands on the subject, which can be assessed in part by the time taken to complete the required tasks, may not remain constant over the course of the experiment: Not only does sleep deprivation increase, but performance demands on the subject may also increase systematically over the course of the experiment.

In this report, we describe the experimental paradigm in detail, present data describing the changes in work/rest duty cycles over the course of the experiment, and discuss methodological issues associated with such experimental paradigms. Work/rest duty cycle is illustrated in the inset of Figure 1, where the hour in minutes is divided into "work" and "rest" segments. Work/rest duty cycle is defined as the ratio of time in minutes necessary for completion of tasks to the time in minutes spent in rest each hour. For example, at the beginning of the experiment, the average time to complete all tasks was 33.5 min. Work/rest duty cycle, therefore, was 0.56.

## METHOD

### Test Instruments

#### Performance Assessment Battery

A computer-generated, -controlled, and -scored multi-task performance assessment battery (PAB) developed at the Walter Reed Army Institute of Research was used. Test items and visual stimuli were presented to subjects on video monitors, and the subjects responded by pressing one or more keys on a conventional keyboard. Details of the battery and its component tasks are described elsewhere (Thorne, Genser, Sing, & Hegge, 1985) and thus are described only briefly here. The approximate task durations cited below represent those obtained from well trained, rested subjects, and not those measured during the sleep-deprivation experiment. Tasks that ran for a specified nominal duration still required a terminating response.

The total PAB duration under normal conditions was approximately 20 min. The subset of the battery used in this experiment included eight tasks, presented in the following order:

**Two-letter search:**—a visual search and recognition task. Two target letters were presented at the top of the screen, followed by a string of 20 letters in the middle of the screen. The subject determined as quickly as possible whether both target letters were or were not present in the string. If both were present, in any order, the "S" key was pressed for "same." If one or more letters was (were) missing, the "D" key was pressed for "different." This task was adapted from the Memory and Search Tasks of Folkard, Knauth, Monk, and Rutenfranz (1976), and took approximately 2 min.

**Six-letter search:**—as above, but with six target letters instead of two and also taking approximately 2 min.

**Two-column addition:**—a subject-paced mental arithmetic task. Five two-digit numbers were presented simultaneously in column format in the center of the screen. The subject determined their sum as rapidly as possible and entered it from the keyboard, most significant digit first. The column of digits disappeared with the first key entry, and no aids for the intermediate sum and carry operation were allowed. This task took approximately 3 min.

**Logical reasoning:**—an exercise in transformational grammar. The letter pair "AB" or "BA" was presented along with a statement that correctly or incorrectly described the order of the letters within the pair (e.g., "B follows A," or "A is not preceded by B"). The subject decided whether the statement was true (same) or false (different) and pressed the "S" or "D" key accordingly. The 32 possible sentence/pair combinations were presented once each or until 4 min elapsed, whichever occurred first. The task was adapted from Baddeley (1968).

**Digit recall:**—a test of short-term memory capacity. Nine random digits were displayed simultaneously in a row across the center of the screen for 1 sec. After a 3-sec blank retention interval, eight of the original nine digits were redisplayed in a different random order, and the subject was required to enter the missing digit. The task took approximately 3 min.

**Serial add/subtract:**—a machine-paced mental arithmetic task requiring sustained attention. Two randomly selected digits and either a plus or minus sign were displayed sequentially in the same screen location, followed by a prompt symbol. The subject performed the indicated addition or subtraction and entered the least significant digit of the result. However, if the result was negative, he added 10 to it and entered the positive single digit remainder (e.g.,  $39 -$  equaled  $-6$ , so 4 was entered). The digits and signs were presented for approximately 250 msec, separated by approximately 200 msec, with the next trial beginning immediately after the key entry. The task was adapted from Pauli as used by Wever (1979), and terminated after 50 trials or 4 min, whichever occurred first.

**Pattern recognition I:**—a spatial memory task. In each of 10 trials, a random pattern of dots (asterisks) was displayed for 1.5 sec, and then followed after a 3.5-sec retention interval by a second pattern that may be the same or different. The subject pressed the "S" or "D" key accordingly. The pattern consisted of 14 dots, of which either 3 or no dots changed locations.

**Pattern recognition II:**—a more difficult version of the previous task. The pattern consisted of 16 dots, of which either 2 or no dots changed. The task was devised by Genser as a nonlinguistic right-hemisphere task. (Ten trials.)

### Lexical Decision Task

The subject wore an eye patch on the left eye, sat with head fixed in position by a forehead-and-chin rest, and fixated at the center of a CRT screen located directly in front at a distance of 1 m. Strings of three to five letters were briefly displayed horizontally on the screen, either to the left or to the right visual field (LVF, RVF). The subject was instructed to press one of two buttons to indicate whether the string was a word or a nonword (Babkoff & Ben-Uriah, 1983; Babkoff, Genser, & Hegge, in press).

The stimulus pool consisted of 264 stimuli, 44 each of three-, four-, and five-letter high frequency words and 44 each of three-, four-, and five-letter nonwords (constructed by randomly rearranging the letters of the words). A Digital Equipment Corporation (DEC) PDP-8/e computer controlled both stimulus presentation and response recording. The angle of eccentricity, measured from the center of the graphics monitor to the midpoint of the stimulus was  $2^{\circ}28'$ . The stimuli were randomized by trial. Stimulus duration was 150 msec followed by a 100-msec blanking flash to eliminate phosphor afterglow and to overcome the effect of the slowly decaying image.

A trial began with the appearance in the center of the screen of a 400-msec focus signal, which signaled the subject to focus on the screen's center and not to move his/her eyes until after the response was made. The position of the subject's eyes was monitored by closed-circuit TV throughout the experiment. The test stimulus was presented at the offset of the focus signal. Responses were made with the right hand only. The subject positioned his/her right index finger on the middle of three electronic contact keys and responded as to whether the stimulus was a word or nonword by moving the finger from the center to one of the lateral keys as quickly as possible as soon as the stimulus was presented. If the finger was removed from the center key while the focus signal was still present, the test stimulus was not presented. The duration of this task was approximately 20 min.

### Mood-Activation Scale

Subjects were presented with 65 adjectives and were asked to assess the extent to which the adjectives reflected their current feelings using a 5-point scale. The adjectives were selected to represent either positive affect, negative affect, positive activation, or negative activation.

This scale consisted of the merger of adjectives from Thayer's Activation-Deactivation Check List (1967) with a Multiple Affect Adjective Check List (Zuckerman, Lubin, & Robins, 1965; Zuckerman, Lubin, Vogel, & Valerious, 1964). Three repeat items were inserted for the purpose of consistency check and three additional adjectives, unlikely to be associated with changes in mood or subjective activation, were added to provide control stimuli. The adjectives were presented singly on a video monitor and responded to by manual keyboard entry. The items also were presented as a list on a printed page once every 2 h and responded to verbally by dictating into a tape recorder. The task continued until all 65 items were completed and took approximately 1-3 min.

### Vigilance Discrimination Task

A series of single digits, randomly selected from "1" through "4," was rapidly presented either visually on a video monitor, vocally with a speech synthesizer, or both. The subject pressed a button as quickly as possible every time the digit "3" occurred. The rate of stimulus presentation was adjusted to the subject's reaction time and error rate. The task was developed in house and required approximately 5 min.

### Psychiatric Symptoms Scale

Subjects were presented with 52 questions concerning sensory/perceptual illusions, distortions, and hallucinations, along with self-assessments of motivation and performance to be scored on a 5-point scale. The task was developed in house and was included because subjects in earlier experiments reported experiencing visual hallucinations. This task continued until all items were completed and took approximately 3 min.

### Fatigue Checklist

Subjects were presented with a list of 30 questions based on the rating scale developed by Yoshitake (1971, 1978). The questions dealt mostly with possible somatic complaints, and the responses were made by checking "yes" or "no" on a printed sheet. This task took approximately 1 min.

### Physiological Measures

Throughout the experiment, subjects wore a belt-mounted, battery-powered solid-state digital recorder (Vitalog Corporation, Model PMS-8) that was about the size and weight of a Sony Walkman. This instrument was used for continuously recording heart rate and general activity levels. Heart rate was derived from a three-lead EKG signal. Activity measures were obtained from mercury-switch actigraphs worn on the subject's two wrists.

### Subjects

The subjects were 2 female and 8 male volunteers ranging in age from 18 to 29 years. All were high school

graduates with zero to four years of college (median of one year) and no history of diagnosable psychiatric disorders. All were screened by (1) a standard U.S. Army medical history questionnaire and (2) an examination by their family physicians for absence of acute physical illness, chronic illness requiring medication, seizures, current alcohol and drug abuse, and abnormal rest-activity cycles.

### Procedures

Subjects were tested in pairs. The initial experimental run tested 2 subjects under the conditions described above for 48 h. Subsequent experimental runs testing 8 additional subjects were increased to 72 h. Consequently all of the data presented in this paper were based on 10 subjects for the first 48 h and on only 8 subjects for the last 24 h of sleep deprivation.

Each pair of subjects spent approximately 3 days of the week prior to the sleep-deprivation procedure in initial training on the experimental tasks. The week of the experiment began with one additional day of training followed by the experimental sleep-deprivation procedure. Civilian subjects were paid on a scale that increased from \$4/h for the training period and first 24 h, to \$8/h for the last 24 h, allowing them to earn up to \$568 if they reached the full 72-h goal of the study. Military subjects received compensatory leave, and earned free time on a sliding scale ranging from 2 to 3 h per hour of participation in the study, for a possible total of 21 8-h shifts of time off. All subjects read and signed a consent form that explained the details and risks of the experiment and assured subjects that they could withdraw freely from the study at any time.

The subjects were tested hourly for a duration of 30 to 45 min. The remainder of each hour was free for any reasonable activity except sleeping. Books, magazines, cards, television, board and video games, video movies, music tapes, exercise equipment, a shower, and comfortable chairs were provided, and access was granted to the rest of the building and grounds. Subjects were given a variety of snacks, beverages, and fresh or frozen foods which they could prepare themselves. Use of tobacco was allowed, but consumption was restricted to individual norms. Aspirin and Tylenol were available upon request but no other drugs were permitted, and coffee was permitted in moderation. Eating, drinking, or smoking was not permitted during or immediately preceding the tests.

Testing was conducted individually but simultaneously inside acoustically isolated room-sized chambers and was monitored by staff members via intercom and closed-circuit TV. At least two staff members were always in attendance. Testing was initiated every hour on the hour. The tests were grouped into two "packages": one subject of each pair was tested with one package on odd-numbered hours, then with the other package on even-numbered hours; the reverse order was used with the other subject. The tests were grouped and ordered as follows:

Package A: Performance assessment battery plus mood-activation scale, vigilance task, verbal response mood-activation scale.

Package B: Mood-activation scale, lexical decision task, mood-activation scale, psychiatric symptoms scale, vigilance task, fatigue checklist.

Oral temperatures were taken at least once each hour near the end of the test period, using a Diatek model 400 electronic digital thermometer. Oral intake was restricted for a period of at least 30 min preceding temperature measurement.

At the end of the 72-h period, the subjects were requested to sleep for approximately 4 h to provide restorative sleep before being allowed to return home. Six of the subjects agreed to remain for a postnap 2-h test cycle to assess the degree of recovery.

### RESULTS

The data reported here cover only the times necessary to complete the various performance tasks, and not the performance scores or measures, which will be reported in detail elsewhere. The following analysis of changes in work/rest duty cycles involves the entire group data based on 10 subjects for the first 48 h and 8 subjects for the last 24 h of the experiment, as noted above.

Figure 1 shows the average time needed to complete all required tasks, as a function of hour and day into the experiment. Subject variance is indicated by  $\pm 1$  SD around each plotted point. These data were calculated for each 2-h test block and then normalized to minutes per hour. Average work/rest duty cycles appear relatively constant for the first 12–20 h of the experiment but increase rapidly thereafter, showing increases in variability. Over the course of the experiment, duty cycle increased from 0.56 (33.9 min per hour) initially, to a maximum of 0.71 (42.4 min per hour) in the early morning hours nearing the end of the 72-h period. The mean duty cycles for the first and last 12 h were 33.5 and 39.4 min per hour respectively, an increase of 17.8%. The peaks and troughs seen in the second half of the data, and their temporal locations suggest the presence of a 24-h cyclical component superimposed upon a more general trend, but this data makes it difficult to determine whether the trend is linear or curvilinear. The value for the final data point in Figure 1 is known to be contaminated by an anticipatory end-of-experiment effect independent of sleep-deprivation duration (Thorne, Genser, Sing, & Hegge, 1983), though it may be confounded as well with circadian influences.

An illustration of the changes in average work/rest duty cycles from the first 12 h to the last 12 h is provided in the inset of Figure 1. Note the increased average block of time assigned to the work segment during the last 12 h, relative to the time assigned to the work segment during the first 12 h.

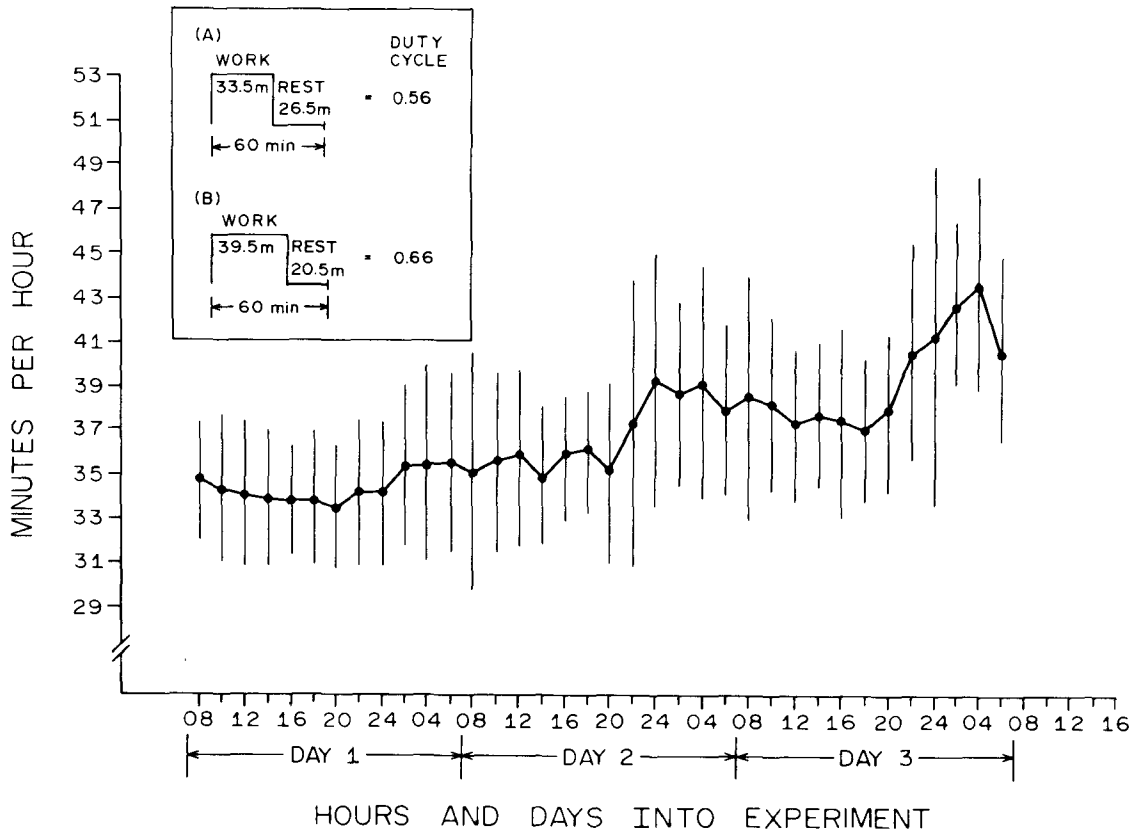


Figure 1. Average time ( $\pm 1$  SD) required to complete all tasks as a function of increasing sleep deprivation. Inset: Illustration of duty cycle change from beginning of experiment (A) to end of experiment (B). See text for explanation.

From the description of the tasks, it is evident that the major tasks, in terms of length of time and amount of data generated, were the PAB and the lexical decision task. The other tasks (i.e., vigilance discrimination, mood-activation scale, psychiatric symptoms scale, and fatigue checklist) were quite short by comparison. For the sake of clarity and manageability, the combined data here consist of three major components: (1) PAB, (2) lexical decision, and (3) all the other shorter tasks combined. The following analysis of the changes in work/rest duty cycles over the course of the experiment, therefore, relates to these three components.

Figure 2 shows the averaged work/rest duty cycles for the three components, plotted separately. For all three curves, the work/rest duty cycle increases as a function of time into the experiment. The two longer tasks are roughly comparable in magnitude throughout the study, with PAB being slightly shorter than the lexical decision task at the beginning, but longer near the end.

After 4 h of recovery sleep, there is a marked decrease in the work/rest duty cycles for the shorter tasks (i.e., a return to the level observed at the beginning of the experiment). The decrease for the two longer duration tasks is less and returns only to the level of Day 2.

Work/rest duty cycle data were analyzed by a two-way ANOVA with repeated measurements on two variables:

type of task (3)  $\times$  days of sleep deprivation (3). The results indicate that the main effects and the interactions are significant ( $p \leq .0001$ ) (see Table 1). Duncan's multiple range test indicates that duty cycle for PAB and lexical decision differ from the combined shorter tests ( $p \leq .0001$ ) and increases across each of the three days of sleep deprivation ( $p \leq .0001$ ). The interaction term indicates that the increases in duty cycle for PAB and lexical decision from Day 1 to Day 3 are greater than the increase in duty cycle for the combined shorter tests ( $p \leq .0001$ ).

These results are emphasized by a direct comparison of the mean duty cycle for the first and last 12 h of the experiment for each of the three tasks (Table 2). The two longer tasks show larger absolute and relative increases than do the combined shorter tasks, and PAB shows greater absolute and relative increases than does the lexical decision task ( $p \leq .0001$ ).

The data shown in Figure 2 were subjected to regression analyses using three different models: linear, exponential, and modified exponential ( $y = a + e^{bx}$ ). The results are shown in the first column of Table 3. The linear and exponential models both provide good fits for the two longer tasks, accounting for approximately 80% of the variance for PAB and 70% for the lexical decision task. The difference between the two models is only 2% in both

## DUTY CYCLE VS TIME

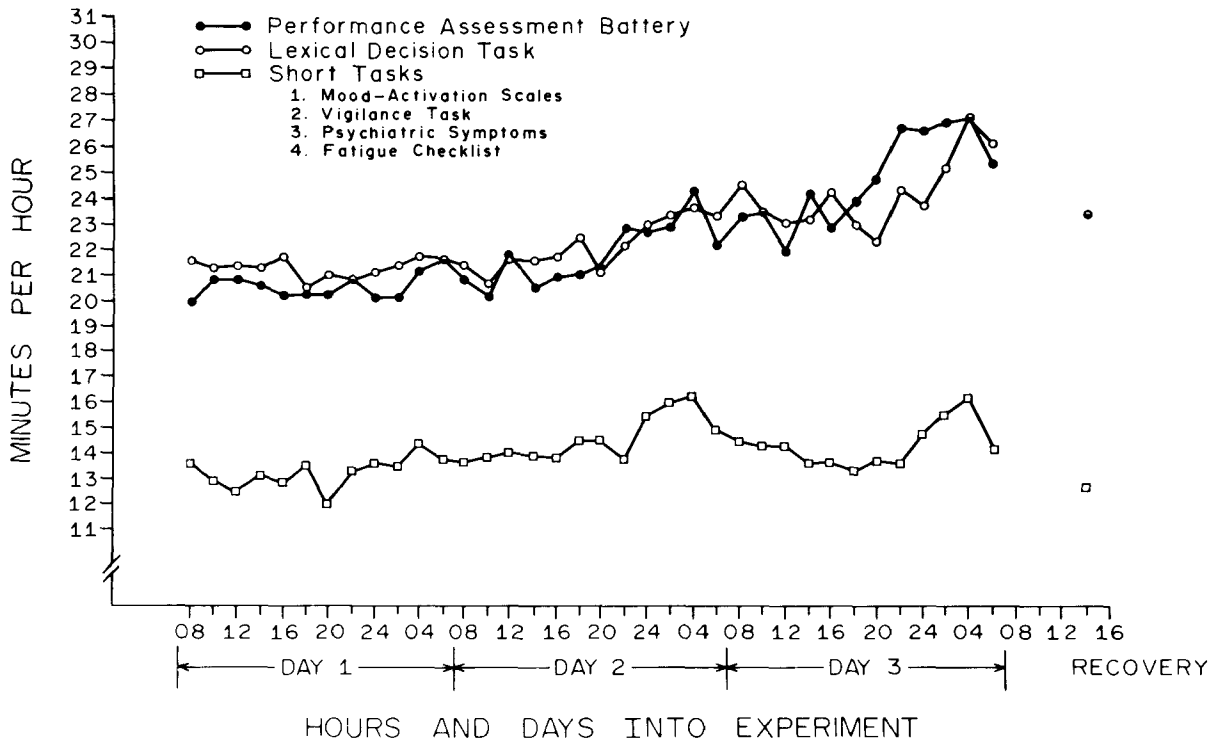


Figure 2. Average time required to complete the performance assessment battery, lexical decision task, and combined short tasks, as a function of increasing sleep deprivation. See text for explanation.

cases, which is probably too small a difference to justify choosing one model over the other. However, the fact that the exponential fit is slightly better in both cases indicates the possibility of a curvilinear relationship which neither model can accurately match. The modified exponential model yields significantly better fits for both tasks (accounting for 86% of the variance for PAB and 77% for the lexical decision task).

The duty cycles for the combined shorter tasks differ from the duty cycles for PAB and the lexical decision tasks. The results of the regression analyses shown in Table 3 indicate that the linear and exponential models accounted for 32% of the variance, and the modified exponential model accounted for even less (28%). A visual inspection of the data plotted in Figure 2 shows that time spent in task increases only gradually for the first half of the experiment, then exhibits a sharp increase followed by a long decline and another sharp increase. Such an erratic function might be fitted by a higher order polynomial, although its meaning would not be clear.

An assessment of the data leads to the conclusion that both monotonic and nonmonotonic components are present in the data and that the variance not accounted for by the regression analyses was most likely due to rhythmic components. Therefore, the three data sets were analyzed by multiple complex demodulation (MCD) (Red-

mond, Sing, & Hegge, 1982; Sing, Genser, Babkoff, Thorne, & Hegge, 1984; Sing, Redmond, & Hegge, 1980). This technique allows extraction of: (1) the circadian component, (2) other higher frequency and noise components, and (3) a function which we term the "smoothed trend." Regression analysis of the raw data, with the rhythmic components derived by MCD, shows that the circadian component accounts for 8% of the total variance in the case of PAB, 14% in lexical decision, and 41% in the combined short tasks. The remaining higher frequency and noise components account for about 5% of the variance in the PAB, 7% in the lexical decision task, and 22% in the combined short tasks.

The circadian components were subtracted from the original data for each of the three tasks, and the residuals

Table 1  
Analysis of Variance (Repeated Measures) of Work/Rest  
Duty Cycles for Days of Sleep Deprivation  
Versus Types of Tasks

Source	df	SS	MS	F
A. Days of Sleep Deprivation	2	1132.19	566.09	81.63*
B. Types of Tasks	2	14435.86	7217.93	1040.88*
C. Interaction (A × B)	4	289.30	72.32	10.43*
Error	974	6754.17	6.93	

\* $p \leq .0001$ .

**Table 2**  
Average Duty Cycle in Minutes per Hour for the First and Last 12 Hours of the Experiment

First 12-h Period	Last 12-h Period	Difference	t Value	Percent Increase
Performance Assessment Battery				
20.1	25.7	+5.6	138.2*	27.8
Lexical Decision Task				
20.9	24.3	+3.4	118.1*	16.2
Combined Shorter Tasks				
13.0	14.4	+1.5	80.9*	11.2

\* $p \leq .0001$ .

were subjected to regression analysis using the same three models as before. The fourth column of Table 3 lists the resulting  $r^2$  values. Removing the circadian variation improved the fit for all models for all three tasks. The amount of improvement was smallest for PAB and largest for the combined short tasks, as would be expected from the magnitude of the circadian components associated with each. In summary, comparing the results of the linear and exponential (monotonic) trends with those of MCD (rhythmic) trends indicated that the monotonic components accounted for more of the PAB and lexical decision data than did the rhythmic components; for the shorter tasks, the rhythmic components accounted for more of the data variance than did the monotonic components.

An additional regression analysis was performed for each of the tasks, using the same four models against the MCD-derived smoothed trend. This trend is obtained by a low-pass-filtering operation performed in the frequency domain, which is analogous to performing a weighted moving average in the time domain with weights determined from information in the entire data set. The result is a smooth curve characterizing the empirical data (not a theoretical equation, and not a simply transformed plot as was the case for data in Columns 4-6). Results of the analysis are shown in Columns 7-9 of Table 3. For the two longer tasks, the smoothed trend is fitted very closely by the modified exponential equations, indicating that the underlying function is quite definitely curvilinear and nearly parabolic, particularly in the case of PAB.

To aid visual interpretation of the regression analysis, the data from Figure 2 are replotted in Figure 3 with the MCD trend and trend-plus-circadian components. The scales have been expanded uniformly for each task but are offset from one another for clarity.

The trend-plus-circadian component tracks the original data points rather faithfully, even for the combined shorter tasks. Furthermore, the peaks and troughs of the three circadian components, which were calculated independently, are not arbitrarily placed but are aligned with one another at the times of day we would expect. That is, the peak increases in duty cycle are found during the hours of midnight to 8:00 a.m., and the relative decreases (troughs) are found between 4:00 and 10:00 p.m.

The smoothed trend for the PAB data, in fact, does appear relatively parabolic over most of its range but then turns downward near the end. Such a reversal is probably an artifact arising from the final data point and the end-of-experiment effect mentioned above. Most smoothing techniques, including MCD, cause the effects of any discontinuity to be spread out over several adjacent points. The same end-effect influences the circadian component as well, pulling down the last few points somewhat more rapidly than is appropriate for a 24-h sine wave.

The smoothed trend for the lexical decision task resembles that found for PAB both in its general form and in the reversal near the end, though the rate of increase is somewhat less. There is evidence of a small inflection in the second half of the record, but this probably has no functional significance. The minimum values for the two trends do not coincide with the first data point, but occur 9 h into the experiment in the case of PAB and 14 h into the lexical decision task. We call attention to this without speculating on its meaning.

The smoothed trend for the combined short tasks is multimodal and resembles neither the trend for the longer tasks nor any simple function we would expect from sleep deprivation and workload effects. Presumably the trend is a result of summing six different tasks having different individual functions, but we raise the question of why the same trend did not hold for the eight different tasks comprising PAB. Apparently other unidentified factors are at play.

**Table 3**  
Regression Analysis of Averaged Duty Cycles with Sequential Hours into the Experiment

Model	Original Data			Original minus Circadian			Smoothed Trend		
	$r^2$	F*	p	$r^2$	F*	p	$r^2$	F*	p
Performance Assessment Battery									
Linear	.78	119.79	.0001	.80	138.86	.0001	.90	305.97	.0001
Exponential	.80	67.33	.0001	.82	79.23	.0001	.92	195.35	.0001
Modified Exponential†	.86	109.31	.0001	.89	140.87	.0001	.98	827.82	.0001
Lexical Decision Task									
Linear	.69	76.90	.0001	.74	98.51	.0001	.90	312.42	.0001
Exponential	.71	40.88	.0001	.76	52.29	.0001	.92	182.61	.0001
Modified Exponential†	.77	57.94	.0001	.82	79.14	.0001	.97	498.06	.0001
Combined Shorter Tasks									
Linear	.32	16.17	.0003	.37	19.77	.0001	.68	72.08	.0001
Exponential	.32	7.93	.01	.36	9.66	.001	.67	34.13	.001
Modified Exponential†	.28	6.79	.01	.32	8.05	.01	.58	23.25	.001

\* $n = 36$ . † $y = a + e^{bx}$ .

DUTY CYCLE VS TIME

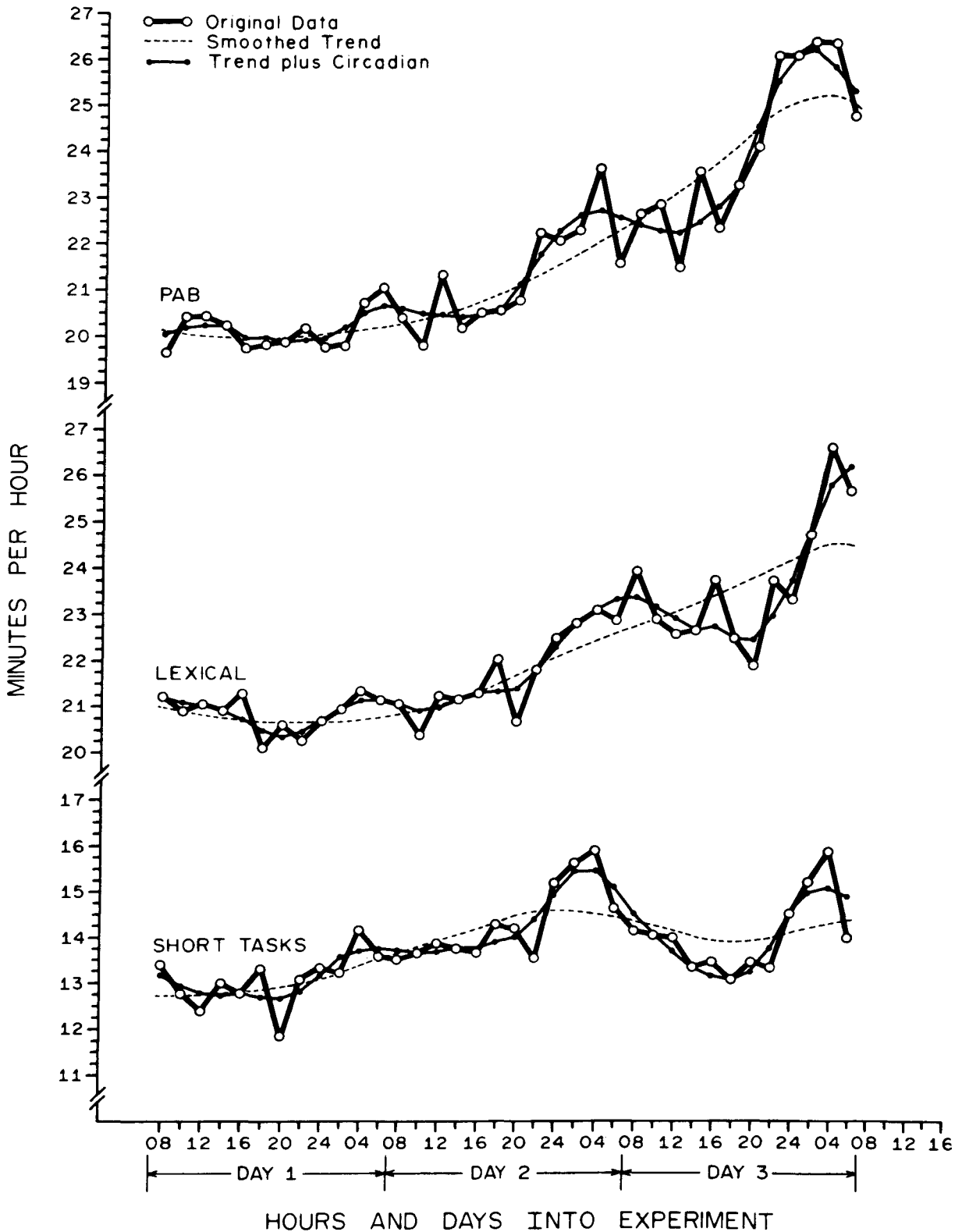


Figure 3. Trend and circadian components for the data in Figure 2, replotted on an expanded scale.



The MCD analysis indicates that there is a positive, accelerating increase in the work period for both PAB and the lexical decision task. For the combined short tasks, however, the work period is driven more by the circadian and higher frequency components and less by any systematic increase.

## DISCUSSION

We define workload as the time and effort required of subjects for task performance during the experiment. Workload, therefore, contains a temporal component as well as intensive components (e.g., the number and difficulty of tasks, the type and complexity of response to them, etc.). The work/rest duty cycle measures the time that subjects spend in task performance each hour, hence reflecting the temporal aspect of workload.

In this experiment, work/rest duty cycle was initially determined by the selection of tasks, and by adjusting their duration and number of trials so that the subject was engaged in work approximately 34 min/h. This duty cycle (0.56 or slightly more than 50% per hour) reflected the performance of subjects both well trained and well rested. The increase in duty cycle should not be viewed as a change only in a dependent variable describing the effects of fatigue and sleep deprivation (i.e., that it takes longer to perform cognitive tasks). In fact, this increase implies that the performance demanded of subjects during the course of sleep deprivation has changed dynamically. Subjects were engaged in "work" rather than in "rest" for longer periods toward the end of the experiment than at the beginning of the experiment.

We suggest the following interpretation. The changes in duty cycle reflect interaction of the independent variables (sleep deprivation and workload) with the dependent variable (duty cycle). Sleep deprivation and workload contribute to produce a relatively slow increase in duty cycle at the beginning of the experiment. However, as time progresses, the subjects spend more and more time in completing the required tasks. Because the defined temporal unit for testing is fixed (one test package per hour), the time available for rest and leisure activities decreases progressively and concomitantly. As the work/rest duty cycle increases, workload necessarily increases as well, even though the intensive task components may not be changing. The dependent variable duty cycle begins to contribute as an additional controlling variable.

From a sleep-deprivation, moderate workload paradigm at the beginning of the experiment, the manipulation became less moderate and more stressful as the experiment progressed. The acceleration in duty cycle was especially manifest toward the latter portion of the 72-h period. Such a dynamic change in the experimental conditions must have an impact on the nature of the results and their interpretation. This acceleration of duty cycle further enhances the fatiguing effects of the workload. The appropriate model to describe these effects seems to be a self-amplifying feedback loop.

The observed changes in work/rest duty cycles are not

equivalent across tasks, however, with the combined shorter tasks being least affected, and the longer tasks being most affected. Thus, additional support is lent to the general finding regarding sleep deprivation studies—the longer the task, the more sensitive it is to sleep loss (e.g., Johnson, 1982; Meddis, 1982)—but with a slightly different emphasis. Not only does performance show greater deficit with a lengthier task, but the time spent in performing the task increases as sleep deprivation increases, and the rate of increase is greater for lengthier tasks.

In light of the above interpretation of these data, there are several implications for further experiments. Can those factors that induce or support large changes in the work/rest duty cycles, especially in the latter segment of the experiment, be isolated and studied? If a self-amplifying feedback loop is operating in the experimental paradigm, then isolating and mitigating those factors could radically alter the form of the duty-cycle function. We must understand what factors tend to increase work time for one type of performance task differently than for another. For example, although length of task had an effect on rate of change in the duty cycle, PAB performance was more severely affected than performance in the lexical decision task, even though the tasks' lengths at the beginning of the experiment were quite similar. It seemed apparent that additional factors, other than length of task, are responsible for these differences, but we are presently unable to identify these variables.

The experimental paradigm used here allowed a number of different types of performance to be measured in a single 72-h study. Repetitive testing at equal intervals of 1 and 2 h allowed the normally confounded influences of circadian and ultradian modulations to be separated from the main effects of sleep deprivation and workload. However, the effects of sleep deprivation and workload also were partially confounded. The experimenter could hold constant the qualitative components of workload, and could set the minimum value of its duty-cycle component, but duty cycle would then be free to vary upward from that minimum.

In future experiments, it may be desirable to develop paradigms that remove or reduce the sleep-deprivation and workload interactions so that their separate effects on various performances can be studied. Several procedures for doing this immediately suggest themselves. Among these are the use of fixed rest periods, proportional rest periods, and fixed-duration tasks. Each procedure has its problems.

Fixing the rest period does not hold duty cycle constant, but it would reduce the rate at which duty cycle will increase with increasing work time. This procedure, however, would lead to variable and increasing intertest intervals and therefore would prevent using most available techniques for measuring the effects of circadian and ultradian rhythmicities.

Making the rest period equal or proportional to the duration of the preceding work period would maintain average duty cycle constant, but would also lead to variable and increasing intertest intervals, as above. In addition, this procedure contains unusual contingencies that could

lead subjects to lengthen the test period in the least effortful way possible, for example, by pausing, by napping, by not responding, or by other clever means less easily detected. Building in counter-contingencies does not guarantee the recovery of the originally intended behavior, and in any case changes the nature of the task.

The use of fixed-duration tasks would indeed hold duty cycle constant, but it would also restrict the number and types of performance tasks that could be used. Terminating certain tasks before all test items have been completed could make them unscorable or meaningless; prevent equalizing stimulus and choice frequencies; prevent assessment of chance accuracy; reduce reliability; and prevent comparisons across subjects or conditions. In addition, simple fixed-duration tasks almost always generate low response rates that gradually decline with repeated administrations, yielding fewer responses per session, less reliable means, and increased variability. This effect is magnified by tasks that are unrewarding, demanding, or aversive, so that the use of fixed-duration tasks in sleep-deprivation studies could lead to dynamic changes in performance as large as or larger than the effects on duty cycle reported in this paper.

Finally, although the confounding of sleep deprivation and workload may be undesirable from the standpoint of classical experimental design, it should be pointed out that the same dynamic interactions are probably an integral part of the real-world situations we are attempting to model.

Clearly, the attempt to unconfound sleep deprivation and workload effects is not a simple one. The above approaches and others should be weighed carefully. By way of perspective, we should point out that our major concern is with performance, not with task duration or duty cycle per se. If modifying tasks and procedures in order to separate confounded variables jeopardizes the accuracy, sensitivity, reliability, or meaningfulness of the performance measures the tasks were designed for, then other alternatives must be sought.

Gaining an understanding of the dynamic interactions between sleep deprivation, workload, and performance has significant practical importance for improving our ability to maintain effective work capacity in military, medical, and industrial settings.

#### REFERENCES

- ALLUISTI, E. A. (1972). Influence of work-rest scheduling and sleep loss on sustained performance. In W. P. Colquhoun (Ed.), *Aspects of human efficiency*. London: English Universities Press.
- ANGUS, R. G., & HESLEGRAVE, R. J. (1985). Effects of sleep loss on sustained cognitive performance during a command and control simulation. *Behavior Research Methods, Instruments, & Computers*, **17**, 55-67.
- BABKOFF, H., & BEN-URIAH, Y. (1983). Lexical decision time as a function of visual field and stimulus probability. *Cortex*, **19**, 13-30.
- BABKOFF, H., GENSER, S., & HEGGE, F. W. (in press). Lexical decision, parafoveal eccentricity, and visual hemifield. *Cortex*.
- BADDELEY, A. D. (1968). A 3 min reasoning test based on grammatical transformation. *Psychonomic Science*, **10**, 341-342.
- FOLKARD, S., KNAUTH, P., MONK, T. H., & RUTENFRANZ, J. (1976). The effect of memory load on the circadian variation in performance efficiency under a rapidly rotating shift system. *Ergonomics*, **19**, 479-488.
- JOHNSON, L. C. (1969). Psychological and physiological changes following total sleep deprivation. In A. Kales (Ed.), *Sleep: Physiology and pathology*. Philadelphia: Lippincott.
- JOHNSON, L. C. (1982). Sleep deprivation and performance. In W. B. Webb (Ed.), *Biological rhythms, sleep and performance*. New York: Wiley.
- JOHNSON, L. C., & NAITOH, P. (1974). *The operational consequences of sleep deprivation and sleep deficit* (AGARDograph-193). London: Technical Editing & Reproduction.
- LOVELAND, N. T., & WILLIAMS, H. L. (1963). Adding, sleep loss, and body temperature. *Perceptual Motor Skills*, **16**, 923-929.
- MEDDIS, R. (1982). Cognitive dysfunction following loss of sleep. In A. Burton (Ed.), *The pathology and psychology of cognition*. London: Methuen.
- MULLANEY, D. J., KRIPKE, D. F., FLECK, P. A., & JOHNSON, L. C. (1983). Sleep and nap effects on continuous sustained performance. *Psychophysiology*, **20**, 643-651.
- ORR, W. C., HOFFMAN, H. J., & HEGGE, F. W. (1974). Ultradian rhythms in extended performance. *Aerospace Medicine*, **45**, 995-1000.
- REDMOND, D. P., SING, H. C., & HEGGE, F. W. (1982). Biological time series analysis using complex demodulation. In F. W. Brown & R. C. Graeber (Eds.), *Rhythmic aspects of behavior*. Hillsdale, NJ: Erlbaum.
- SING, H. C., GENSER, S. G., BABKOFF, H., THORNE, D. R., & HEGGE, F. W. (1984). Complex demodulation—a technique for assessing periodic components in sequentially sampled data. *Proceedings of the 29th Conference on Design of Experiments in Army Research Development and Testing* (ARO Report 84-2). Bethesda, MD: Uniformed Services University of the Health Sciences.
- SING, H. C., REDMOND, D. P., & HEGGE, F. W. (1980). Multiple complex demodulation: A method for rhythmic analysis of physiological and biological data. *Proceedings of the Fourth Annual Symposium on Computer Applications in Medical Care* (pp. 151-158). New York: Institute of Electrical and Electronic Engineers.
- THAYER, R. E. (1967). Measurement of activation through self report. *Psychological Reports*, **20**, 663-678.
- THORNE, D. R., GENSER, S. G., SING, H. C., & HEGGE, F. W. (1983). Plumbing human performance limits during 72 hours of high task load. *Proceedings of the 24th DRG Seminar on the Human as a Limiting Element in Military Systems*. Toronto: NATO Defence Research Group.
- THORNE, D. R., GENSER, S. G., SING, H. C., & HEGGE, F. W. (1985). The Walter Reed performance assessment battery. *Neurobehavioral Toxicology & Teratology*, **7**, 415-418.
- WEVER, R. A. (1979). *The circadian system of man*. New York, Heidelberg-Berlin: Springer Verlag.
- WILKINSON, R. T. (1964). Effects of up to 60 hours sleep deprivation on different types of work. *Ergonomics*, **7**, 175-186.
- WILKINSON, R. T. (1969). Sleep deprivation: Performance tests for partial and reduced sleep deprivation. In L. E. Abt & B. F. Reise (Eds.), *Progress in clinical psychology* (Vol. 8, pp. 28-43). New York: Grune and Stratton.
- WILKINSON, R. T., EDWARDS, R. S., & HAINES, E. (1966). Performance following a night of reduced sleep. *Psychonomic Science*, **5**, 471-472.
- WILLIAMS, H. L., & LUBIN, A. (1967). Speeded addition and sleep loss. *Journal of Experimental Psychology*, **73**, 313-317.
- YOSHITAKE, H. (1971). Relationship between the symptoms and the feeling of fatigue. *Ergonomics*, **14**, 175-186.
- YOSHITAKE, H. (1978). Three characteristic patterns of subjective fatigue symptoms. *Ergonomics*, **21**, 231-233.
- ZUCKERMAN, M., LUBIN, B., & ROBINS, S. (1965). Validation of the multiple affect adjective check list in clinical situations. *Journal of Consulting Psychology*, **29**, 594.
- ZUCKERMAN, M., LUBIN, B., VOGEL, L., & VALERIOUS, E. (1964). Measurement of experimentally induced effects. *Consulting Psychology*, **28**, 418-425.