

ORIGINAL PAGE IS
OF POOR QUALITY

N89-19846

PHYSIOLOGICAL ASSESSMENT OF TASK UNDERLOAD

J. Raymond Comstock, Jr.
Randall L. Harris, Sr.
Alan T. Pope

NASA Langley Research Center
Hampton, Virginia 23665-5225

Abstract

The ultimate goal of research efforts directed at underload, boredom, or complacency in high-technology work environments is to detect conditions or states of the operator that can be demonstrated to lead to performance degradation, and then to intervene in the environment to restore acceptable system performance. Physiological measures may provide indices of changes in condition or state of the operator that may be of value in high-technology work environments. The focus of the present study was on the use of physiological measures in the assessment of operator condition or state in a task underload scenario. A fault acknowledgment task characterized by simple repetitive responses with minimal novelty, complexity, and uncertainty was employed to place subjects in a task underload situation. Physiological measures (ECG, EEG, and pupil diameter) were monitored during task performance over a one-hour test session for 12 subjects. Each of the physiological measures exhibited changes over the test session indicative of decrements in subject arousal level. While high correlations between physiological measures were found across subjects, individual differences between subjects supports the use of profiling techniques to establish baselines unique to each subject.

Introduction

With the development and increasing number of high-technology work environments, the roles, functions, and tasks performed by people in these environments need to be considered carefully. Although not unique to high technology or automated environments, words such as boredom, underload, and complacency describe conditions of the human operator or system manager that may have major impact on the performance and safety of high-technology systems.

Increased automation can change the nature of the task that the operator is performing, and may increase the monitor-

ing or vigilance component of a task (Parasuraman, 1987). Many studies and reports concerning the effects of aircraft automation may be found in the literature (e.g. Bergeron & Hinton, 1985; Wiener, 1985; Wiener & Curry, 1980). Each report echoes concern, both about errors made in the use of automated systems and failure of the operator or crew to detect errors when using an automated system.

Reports concerning the effects of automation on commercial aircraft pilots note that too much automation may result in complacency and boredom (Chambers & Nagel, 1985), and pilot reports of just "being along for the ride." (Chambers & Nagel, 1985; Wiener, 1985). Chambers and Nagel (1985) discuss potential alternatives in the allocation of functions between automated systems and the human operator. In one system, which they called the "pilot's assistant concept," the pilot plays an active role as the system manager. It was suggested that the active role of the pilot may serve to reduce complacency and boredom. A second system, the "electronic copilot concept," was presented in which the pilot functions as a passive monitor, with concomitant risks of complacency and boredom. In reviewing the report by Chambers and Nagel, Parasuraman (1987) suggested "that the goal of human factors engineering must not necessarily be to reduce workload but to optimize it."

In research on vigilance, conditions described as boredom, monotony, and fatigue have been reported as having overwhelming effects compared to the many variables (e.g., temperature, noise, circadian effects, illumination, and vibration) that have received experimental attention (Mackie, 1987). Studies examining stress effects on sonar operators show that boredom, monotony, and fatigue were ranked the highest of all the stressors considered (Mackie, Wylie, & Smith, 1985).

The ultimate goal of research efforts directed at underload, boredom, or complacency is to detect conditions or states of the human operator or crew member that can be shown to lead to performance degrada-

ORIGINAL PAGE IS OF POOR QUALITY

tion, and then to intervene in the automated work environment to restore acceptable system performance. While monitoring of machines by human operators has a long history, monitoring of the human operator by the machine, or better yet, an interaction or "dialog" between the human and machine is only becoming feasible through the technology of today and the promises and potential of the technology of tomorrow.

In order to detect changes in the condition or state of the human operator or crew member, physiological measures may provide the indices required. Many studies have been conducted exploring the relationship between physiological parameters and "workload," "vigilance," or "fatigue" (See O'Donnell, 1979, for a review), and physiologically driven "alertness indicators" (O'Donnell, 1979, pp. 57 & 62) have been suggested and explored. However, in order for these measures to be employed as viable indices of operator condition or state, research efforts must take advantage of new and steadily improving sensor technology, and state-of-the-art analysis methodologies. Horst (1988) discusses some of the areas of application and problem areas associated with physiological measures, and notes that on-line or real-time applications of physiological indices may include: (1) assessment of "the general state of the operator," in order to determine whether the operator should be "in the loop" at all, (2) dynamic task allocation between human and machine, and (3) detection of important events that the operator did not attend to or detection of events with error responses (Horst, 1988, p. 30).

The focus of the present study was on the use of physiological measures in the assessment of operator condition or state in a task underload scenario. Research directed at detecting conditions or states of the operator in a task underload scenario that can be related to performance degradation can be conceptualized as having four stages:

(1) Stage 1 involves establishing the sensitivity of selected physiological measures to the condition or state of the operator in task underload scenarios. This stage may be accomplished, as in the present study, through physiological measurement of subjects in laboratory or simulation environments. One question to be answered at this stage is whether there are generic measures that may be used as indicators for all human operators or whether individual differences between people require the development of profiles unique to individuals or groups of individuals.

(2) Stage 2 is concerned with establishing the relationship between observed physiological changes and task performance. Obviously, there may be many attributes to tasks in the high technology environment, involving a variety of types of cognitive processes on the part of the operator. For example, if physiological measures are indicative of operator "boredom," on which types of tasks can one look for degraded performance (e.g. perceptual, memory, problem solving / decision making). This stage may be investigated most easily by studies conducted in settings in which a high degree of control over the task is possible, such as in laboratory or simulation environments.

(3) Stage 3 involves establishing the relationship between physiological indices and performance in real-world settings. In order to accomplish this, performance assessment and physiological measurement may be conducted in instrumented aircraft or instrumented ground transportation vehicles. In order to investigate underload or "boredom" in real-world settings, it may be necessary to monitor operators or crew members over extended periods of time.

(4) Stage 4 involves the development and testing of off-line and on-line interventions. Such interventions or remediation strategies would be designed to insure that conditions or states of the operator associated with performance degradation are handled in such a way as to restore acceptable system performance.

The present study was directed at Stage 1, described above, and was focused on establishing the sensitivity and inter-relationship of heart rate and heart rate variability, pupil diameter, and electroencephalographic data during performance of a task in which the subject was underloaded. The characteristics of such a task, simple repetitive responses with minimal novelty, complexity, and uncertainty (Cooper, 1968), were such that boredom would be expected.

Methodology and Design

Subjects. Physiological and task performance data were obtained from 12 subjects (10 male; 2 female). Subject ages ranged between 20 and 44 years, with a median age of 28.5 years. Subjects were volunteers from among the staff and contract personnel of the NASA Langley Research Center.

Task. A fault acknowledgment task (five alternative choice-reaction-time task) was selected as a laboratory benchmark task for inducing boredom. The fault acknowledgment task consisted of a desk-

ORIGINAL PAGE IS OF POOR QUALITY

top-computer-generated jet engine pictorial in which one of five areas could be highlighted in red. The task for the subject was to detect the fault, then to press the appropriate response key for the highlighted area, which turned off the highlighted area. The highlighted area remained lighted until the subject responded. If the subject did not respond, a time-out occurred after about 30 seconds, and a new fault was presented. The task was presented on a CGA computer monitor and responses were made on the desktop computer keyboard.

Subject response times and accuracy of responding were recorded for each fault presentation. Frequency of presentation of faults, or intertrial-interval (ITI), for half the subjects was 6 seconds (randomly ranging between 4.8 and 7.2 seconds) throughout the test session and for the remainder of the subjects alternated between 20 second ITIs (range: 16 to 24 seconds) and 2 second ITIs (range: 1.6 to 2.4 seconds) with each ITI maintained for blocks of 5 minute duration. The alternating ITI condition was incorporated in the study in order to examine contrasts in the physiological data during periods in which the task made differing demands on the subject. A complete test session lasted approximately one hour and consisted of 12 five minute blocks. A time synchronization signal was sent from the desktop computer to a Digital Equipment Corporation (DEC) Modular Instrument Computer (MINC) system on which the physiological data were recorded.

Procedure. Upon arrival at the laboratory, subjects initially completed voluntary consent forms, then electrode attachment began. After electrode installation, subjects were seated in front of the task display where instructions on performing the task were given and calibration of the non-contact oculometer system used to measure pupil diameter was completed. The experimental session was then started. Time keeping devices (e.g. clocks and wrist watches) were removed from the experimental setting during the test session to minimize the subject's awareness of the time remaining in the test session. At the conclusion of the test session information was obtained from each subject concerning food, beverages, medications, and sleep prior to the test session. Subjects were also queried about task strategies or mental "games" that they engaged in during the session.

Subject self-report measures. Each subject completed the Jenkins Activity Survey (JAS) (The Psychological Corporation) and the Eysenck Personality Inventory (EPI) (Educational and Industrial Testing Service). The JAS provided an index of the Type A behavior pattern.

Prior studies have explored the relationship between the Type A behavior pattern and physiological responses, and, in particular, heart rate and blood pressure (Perkins, 1984; Lake, Suarez, Schneiderman, & Tocci, 1985). The EPI provided an index of introversion-extraversion.

Physiological measures

Heart Rate (HR) and Heart Rate Variability (HRV). The electrocardiogram (ECG) signal, from which HR and HRV were derived, was obtained through active electrodes attached to the top of the sternum and to the lower left rib cage. A reference electrode was attached to the left ankle. The ECG signal was fed through an opto-isolated bioamplifier to a schmitt trigger which produced a digital output upon an ECG signal positive-going threshold crossing (R-wave). The time interval between successive threshold crossings, or inter-beat interval (IBI), was recorded on the DEC MINC system.

Pupil Diameter (PD). The PD of each subject was measured using a non-contact infrared-light bright-pupil oculometer system. The oculometer electro-optical head was located 60-80 cm from the subject and to the left of the CGA color monitor presenting the task. PD was sampled at one-second intervals throughout the test session and recorded on the DEC MINC system.

Electroencephalography (EEG). EEG recording sites were over the occipital-parietal cortex (locations O1 and P3) using a bipolar lead configuration, with the reference electrode at the ipsilateral mastoid bone. The power (integrated amplitude) of the EEG signal in the various frequency ranges was obtained by filtering the signal and integrating the filtered output. The EEG signal was filtered into five frequency range components: (1) delta (1-4 Hz), (2) theta (4-8 Hz), (3) alpha (8-13 Hz), (4) beta1 (13-20 Hz), and (5) beta2 (20-40 Hz). The output of each filter was fed to a contour following integrator with a 1 second time constant (Coulbourn Instruments filters and integrators). The integrated amplitude of each of the five EEG frequency ranges was sampled by the DEC MINC system at one-second intervals throughout the test session.

Results and Discussion

Task Performance. The primary purpose of the task in the present study was to place the subject in a situation of task underload, a situation in which boredom would be expected. As would be anticipated on a task of this type, few

error responses were made and typical initial response times were under one second. Two subjects showed decreased response times throughout the test session, and three others maintained relatively steady response times. The remaining seven subjects displayed either slow increases in response time or increased response times during several 5 minute blocks during the test session (typically near the end of the session). At no time did a subject completely fail to respond to the task, although response times exceeding 10 seconds were obtained from two subjects during periods when they appeared to be near sleep.

Heart Rate and Heart Rate Variability. As would be expected from subjects performing a task with low demands, HR (compared for each 5 minute block) showed a slight decrease after the initial block, and then remained steady during the rest of the test session. Despite the relative constancy of average HR, on a within-subjects basis, HRV was found to correlate positively with block number for each of the 12 subjects, indicating an increase in HRV with time on the task. Indices of HRV used in the present study were based on filtering and computing the power spectral density for the IBI data. Two measures of variability were examined: (1) total power, and (2) power in the .05 to .15 Hz bandwidth. This band has typically been associated with blood pressure regulation (Kitney, 1980). Figure 1 shows the plot of power (means for 11 subjects) in the the task. The figure shows an increase in average HRV with time on the task.

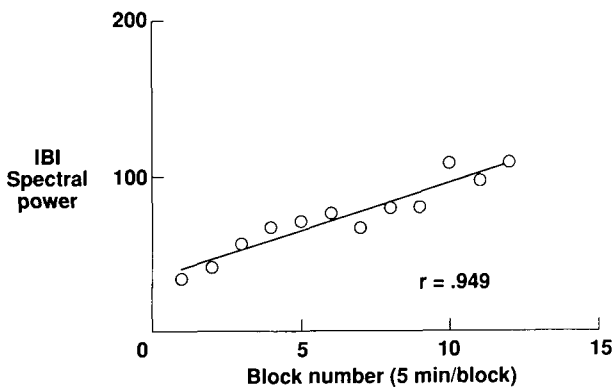


Figure 1. Averaged spectral power (.05 - .15 Hz) of IBI for 11 subjects.

Pupil Diameter. In general, PD has been found to decrease when variety and novelty in the field of view is minimized. Figure 2 presents the relationship between mean pupil diameter for 11 subjects over the course of the test session. As shown in the figure, mean pupil diameter decreased with time on the task.

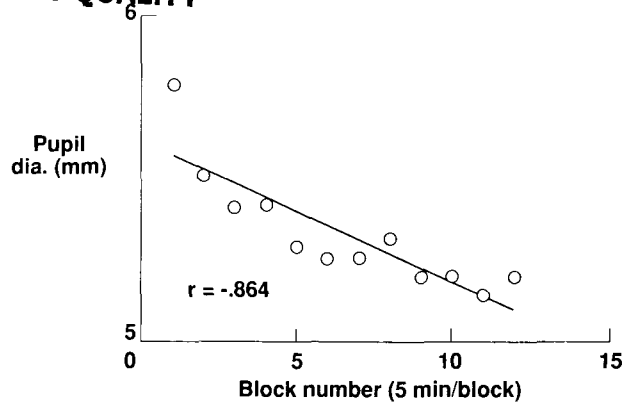


Figure 2. Averaged pupil diameter for 11 subjects.

EEG. As noted above, the EEG signal was filtered into five frequency range components: (1) delta (1-4 Hz), (2) theta (4-8 Hz), (3) alpha (8-13 Hz), (4) beta1 (13-20 Hz), and (5) beta2 (20-40 Hz). EEG data consisted of the level of the integrated amplitude of each of the frequency components (sampled at 1 Hz). An initial hypothesis under test was that continued performance of an underload task would lead to shifts in the EEG activity from faster activity (beta) to slower activity (delta and theta). The data did not support this hypothesis. Analogous to the relationship between HR and HRV, average levels of activity within each frequency band generally remained steady, while variability of the levels within each frequency band exhibited changes with time on the task. Although variability of each

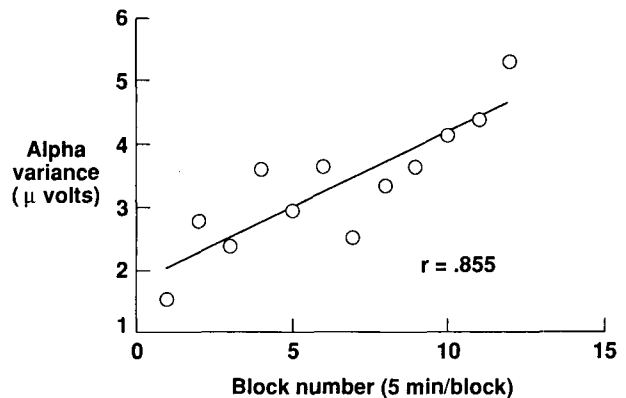


Figure 3. Averaged alpha variance for 11 subjects.

of the lower frequency EEG bands (delta, theta, and alpha) exhibited increases with time on the task, the measure having the highest correlation with HRV and PD was alpha variability. Figure 3 presents alpha variance (means of 11 subjects) versus block number on the task, and shows an increase in alpha variance with time on the task.

Interrelationship of measures. Simultaneous recording of ECG, PD, and EEG permitted examination of the interrelationship between these measures. Figure 4 shows an index of HRV plotted against PD. In this figure the spectral power of the IBI data in the .05 to .15 Hz band (means of 11 subjects) for each block is plotted versus mean pupil diameter for each block. The figure shows increased HRV concomitant with decreased PD. Figure 5 presents the relationship between the HRV index and alpha variance (means of 11 subjects) for each block, and illustrates increases in alpha variance accompanying increases in the HRV index. While the data presented in figures 4 and 5 show high correlation coefficients between measures, when means for the 11 subjects are used, it should be noted that large individual differences in each of the physiological measures are noted when examining the data at the single subject level. Therefore, figures 4 and 5 should not be interpreted simply as suggesting a high degree of convergent validity between measures. Illustrating the independence of the measures, it was found that some subjects displayed changes in one physiological index while showing little or no change in another. These results support the use of profiling techniques in order to establish baselines unique to each individual.

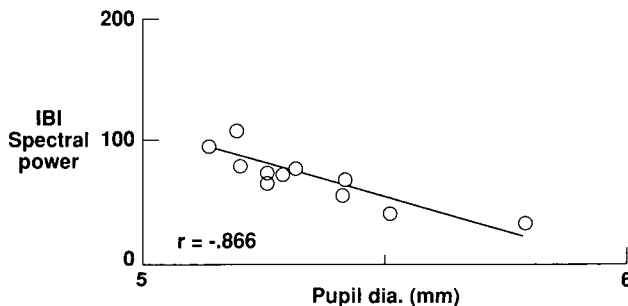


Figure 4. IBI spectral power (.05 - .15 Hz) versus pupil diameter (Mean for 11 subjects).

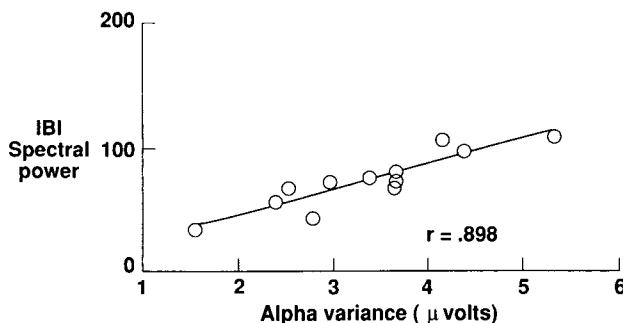


Figure 5. IBI spectral power (.05 - .15 Hz) versus alpha variance (Mean for 11 subjects).

In an underload scenario each of the physiological measures used in the present study are prone to a "ceiling" (or "floor") effect as there are limits to increases in HRV, alpha variability, or decreases in PD. Therefore, it would be expected that over sufficient time periods, these parameters would reach asymptotic levels. As illustrated by figures 1, 2, and 3, which show averages over the test session for HRV, PD, and alpha variance, asymptotic levels were not reached in the one hour test session using the fault acknowledgment task. This suggests that future studies may benefit from a longer period of testing, or perhaps an on-line assessment of physiological indices to determine the duration of an experimental session for a given subject.

Unlike changing the speed of operation or load on a mechanical device, placing a human operator in an underload situation does not guarantee performance deficits or that the operator will be bored. Many of the subjects in the present study reported engaging in a variety of mental tasks or "games" in order to maintain alertness. In effect, the subjects were making changes in load or "busyness" beyond the demands placed on them by the task. Despite engaging in these activities, the majority of subjects exhibited physiological changes during the test session indicative of decreases in arousal level.

The next step in exploring the physiological assessment of task underload is to establish the relationship between observed physiological changes and attributes of task performance. In order to accomplish this step, it is necessary to measure physiological indices while subjects perform (for extended time periods) tasks incorporating and sensitive to the types of activities that operators may encounter in operational settings.

References

- Bergeron, H. P., and Hinton, D. A. (1985). Aircraft automation: The problem of the pilot interface. Aviation, Space, and Environmental Medicine, 56(2), 144-148.
- Chambers, A. B., and Nagel, D. C. (1985). Pilots of the future: Human or computer? Communications of the Association for Computing Machinery, 28, 1187-1199.

ORIGINAL PAGE IS
OF POOR QUALITY

- Cooper, R. (1968). The psychology of boredom. Science Journal, 4(2), 38-42.
- Horst, R. L. (1988). An overview of current approaches and future challenges in physiological monitoring. Mental-State Estimation 1987, NASA CP-2504, 25-42.
- Kitney, R. I. (1980). An analysis of the thermoregulatory influences on heart-rate variability. In Kitney, R. I., and Rompelman, O. (Eds.), The study of heart-rate variability. (p. 81). New York: Oxford University Press.
- Lake, B. W., Suarez, E. C., Schneiderman, N., and Tocci, N. (1985). The Type A behavior pattern, physical fitness, and psychophysiological reactivity. Health Psychology, 4(2), 169-187.
- Mackie, R. R. (1987). Vigilance Research - Are we ready for countermeasures? Human Factors, 29(6), 707-723.
- Mackie, R. R., Wylie, C. D., and Smith, M. J. (1985). Comparative effect of 19 stressors on task performance: Critical literature review (what we appear to know, don't know, and should know). In Proceedings of the Human Factors Society 29th Annual Meeting (pp. 462-469). Santa Monica, CA: Human Factors Society.
- O'Donnell, R. D. (1979). Contributions of psychophysiological techniques to aircraft design and other operational problems. AGARD-AG-244.
- Parasuraman, R. (1987). Human-Computer Monitoring. Human Factors, 29(6), 695-706.
- Perkins, K. A. (1984). Heart rate change in Type A and Type B males as a function of response cost and task difficulty. Psychophysiology, 21(1), 14-21.
- Wiener, E. L. (1985). Beyond the sterile cockpit. Human Factors, 27(1), 75-90.
- Wiener, E. L. (1980). Flight-deck automation: Promises and Problems. NASA TM-81206.