NAVAL HEALTH RESEARCH CENTER

AD-A280 858

FIRST DEMONSTATION OF AN ALERTNESS MONITORING MANAGEMENT SYSTEM

S. Makeig

F. S. Elliott

M. Postal

Report No. 93-36

NAVAL HEALTH RESEARCH CENTER
P. O. BOX 85122
SAN DIEGO, CALIFORNIA 92186 – 5122

NAVAL MEDICAL RESEARCH AND DEVELOPMENT COMMAND
BETHESDA, MARYLAND

FIRST DEMONSTRATION OF AN ALERTNESS MONITORING/MANAGEMENT SYSTEM

S. Makeig[1]
F. S. Elliott[1]
M. Postal[2]

This Report 93-36 was supported by a grant (RS34H21) to the Naval Health Research Center by the Office of Naval Technology. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the U. S. Government. Approved for public release; distribution unlimited.

1. Naval Health Research Center. P.O. Box 85122, San Diego, CA 92186-5122
2. Geo-Centers, Inc., 10903 Indian Head Highway, Fort Washington, MD 20744

Summary

A first laboratory version of an Alertness Monitoring/Management (AMM) system has been designed and implemented. The system continually estimates the level of alertness of a human subject using EEG spectral information recorded from the subject's scalp, and delivers auditory feedback to assist the subject in managing his or her own level of alertness in work environments requiring constant vigilance. The system allows experimenters to monitor its input and output via real-time color graphics displays.

As a first demonstration and evaluation of the system, six subjects participated in five half-hour sessions (three training and two feedback sessions), which involved dual detection tasks simulating the passive sonar environment. Auditory targets, 300-ms noisebursts presented at 6 dB above a noise background, were presented at a mean rate of 10 targets per minute. A continuous visual waterfall display presented illuminated vertical line targets at a mean rate of one per minute. Subjects pressed one response button to report noisebursts and another to report visual targets.

Neural net estimation algorithms were trained for each subject to estimate the current probability of detecting auditory targets using electroencephalogram (EEG) and performance data collected during one or more of the initial training sessions. During feedback sessions, real-time signal processing and individualized neural network analysis of EEG recorded from a central scalp electrode were used to estimate continuously, in near real-time, the current probability-of-detection of auditory targets. Whenever this probability-of-detection measure declined below a preset threshold (e.g., when it predicted more than a 40% chance of failure to detect the auditory targets), the system sounded an alarm in the subject's headphones. When training sessions comprising a relatively wide range of detection rates were used to train the estimation algorithms, the alertness estimates followed changes in observed detection probability relatively accurately.

Four of the six subjects reported that the alertness feedback helped them to maintain detection performance. A fifth subject did not produce enough detection lapses to fairly evaluate the system. Review of data from the sixth subject suggested that future versions of the system may be able to provide useful feedback to this subject as well. Review of results of the demonstration experiment have suggested several improvements to signal processing and training procedures used in the system. Effects of these enhancements on system performance are being evaluated.

Introduction

The increasing speed and complexity of current information technology make it possible for computerized information systems to monitor the physical environment with increasing acuity. However, the capacity of the human operator or decision maker to maintain alertness to this information is not increasing, and relatively little effort has been invested in designing technology to enhance operators' abilities to maintain alertness. Despite common assumptions that while we are awake our alertness to the environment is secure and unbroken, many years of vigilance research attests that for most or all monitoring equipment operators (e.g., air traffic control, sonar, radar), maintaining a constant level of alertness is rare, if not impossible (Mackworth, 1948). Yet, no system currently exists that can monitor an operator's level of alertness directly, deliver timely information about lapses in alertness to the operator and/or supervisor, and initiate appropriate countermeasures.

A system for accurately monitoring operator alertness would provide several benefits that could be realized in a variety of operational environments. The direct benefit of an alertness monitoring system would be the ability to deliver early warning of lapses in alertness to the operator and/or supervisor, prompting interventions or countermeasures leading to a direct improvement in total man/machine system performance. This in turn could reduce or eliminate the need for operator redundancy, resulting in more efficient operation and substantial cost savings through manpower reduction. Currently, little exact information is available on the dynamics of alertness on a minute-by-minute basis for sonar and radar operators, air traffic and engineering plant controllers, pilots, and workers in other job categories requiring constant alertness. An alertness monitoring system would also make possible direct estimates of risk of system failure due to loss of vigilance. This information would allow more productive and cost-effective human engineering and staffing designs for complex systems, which could result in additional overall cost savings.

Since the landmark investigations of Mackworth (1948), studies of alertness have confirmed that, on average, detection rates in laboratory tests begin to degrade after 2 to 3 min of task performance, eventually reaching a plateau during which 70% to 80% of targets are detected (Makeig, Elliott, Inlow, & Kobus, 1990). While most vigilance research in the past has focused on measuring mean trends in performance, performance actually tends to fluctuate irregularly within sessions, with periods of several minutes of complete unresponsiveness not uncommonly observed (Makeig et al., 1990). During the last century, scattered studies have presented evidence for the presence of regular fluctuations 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 (Stroud, 1966). However, most previous vigilance research has focused on simulating actual work environments using relatively low target presentation rates (2/min or less) too low to record minute scale fluctuations.

Methods of real-time alertness monitoring are needed which can detect the actual time course of performance lapses. The only direct, noninvasive measure of brain state available for continuous monitoring in an operational environment is the recording of the brain's electromagnetic activity through changes in potential difference between points on the scalp. This is known as the electroencephalogram (EEG). Since the first development of amplifiers capable of resolving brain electrical activity, it has been known that EEG dynamics change dramatically during and near periods of sleep or drowsiness (Beatty, Greenberg, Deibler, & O'Hara, 1974). However, in the past there have been several obstacles to using this information to estimate effectively and in near-real time an operator's state of alertness. Before the era of high-speed, portable computing and appropriate signal processing algorithms, the complexity of EEG dynamics exceeded the ability of technology to extract useful information in real-time for use in controlling system feedback or control.

The first laboratory version of an Alertness Monitoring/Management (AMM) system. described here has several advantages over previously proposed methods of objective alertness estimation:

1. Many proposals for performing cognitive monitoring using EEG data have relied exclusively on timelocked averaging of responses to unreliable system signals (auditory or visual events) or to signals injected into the operator's environment as probes of the operator's ability to respond. These methods require the operator to frequently shift his or her attention from the primary monitoring task. It would be preferable to monitor operator state without introducing distracting tasks or signals into the work environment.
2. Previous psychophysiological studies of alertness have often used the same estimator for all subjects. However, the relatively large individual variability in EEG dynamics accompanying loss of alertness means that for many operators, group statistics cannot be used to accurately predict changes in alertness and performance (Makeig et al., 1990; Makeig & Inlow, 1993), and estimators customized to individual operators would be more accurate in many cases.
3. Previous research on the relationship between the EEG spectrum and alertness used predefined frequency bands, and therefore were insensitive to individual subject frequency-band differences (Beatty et al., 1974: Matousek & Peterson, 1983; Belyavin & Wright, 1987). The system discussed here incorporates a more flexible signal processing and estimation procedure based on observations of changes in the individual EEG spectrum.

Phase I - System Design and Realization

The demonstration project consisted of two phases. In the first phase, a working laboratory version of a real-time AMM system was designed and implemented. This involved several types of real-time programming of three Unix systems linked by Ethernet, and a DOS-based personal computer with graphics display capabilities. A schematic diagram of this system is shown in Figure 1.

The signal processing performed by the system has several stages:

A. *Simultaneous data collection and dual-task stimulation.* The system records data in parallel on two Concurrent (Real Time Unix) workstations, one dedicated to auditory stimulus synthesis and data recording, the other to real-time spectral and neural net analysis, graphic display, and producing auditory feedback. Another program, originally developed in the laboratory of Dr. Steven Hillyard at the University of California at San Diego (Hillyard &

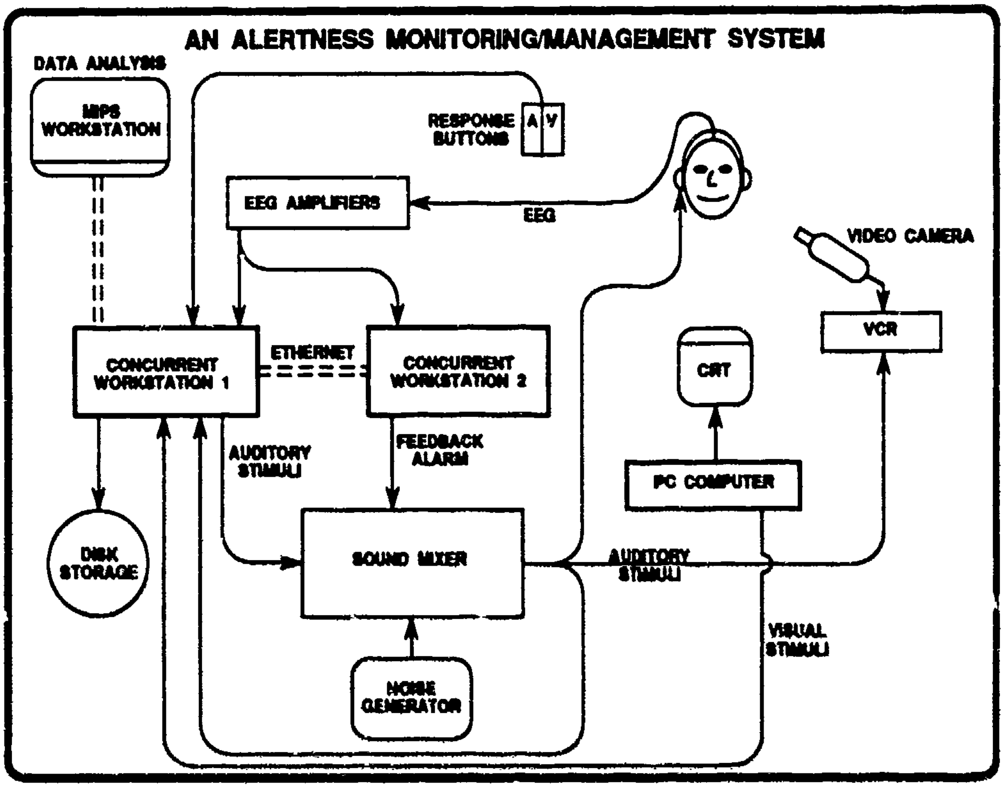

Figure 1. Schematic diagram of the first laboratory version of an Alertness Monitoring/Management (AMM) system constructed during the project. The subject (*upper right*) views a CRT waterfall display while hearing sounds through headphones coming from a sound mixer, and presses auditory and visual response buttons (*upper middle*) whenever he or she detects target signals in the respective modalities. Auditory signals include a noise background, tones, and noiseburst targets synthesized by a Concurrent workstation (*center left*), and occasional auditory alertness feedback signals (brief buzzes) synthesized by a second Concurrent workstation (*center*). The first Concurrent machine also records EEG and performance information to disk. The second Concurrent machine processes EEG information in near-real time, creates an estimate of the subject's current probability of failing to respond to the noiseburst signals, and produces auditory feedback signals whenever the estimated error probability rises above a preset threshold. The Concurrent machines are linked to a higher-speed MIPS workstation (*upper left*) for off-line data analysis. A video camera and recorder (*upper right*) record video and auditory stimulus information during the experiments.

Johnston, submitted), runs on a dedicated PC. It presents visual stimuli against a random visual waterfall background, and delivers stimulus timing information to the Concurrent computers.

B. *Data selection, scaling, and artifact rejection.* Real-time artifact rejection and scaling routines preprocess the EEG data using coefficients read from a file created by the neural net training software. Data from epochs contaminated by artifacts are replaced by the most recent accepted values.

C. *Robust real-time spectral estimation.* Real-time spectral analysis of the data is accomplished using Hanning-windowed data windows of 256 points with 50% overlap. After rejecting windows in which data exceed +/-50 uV, median filtering, a robust statistical smoothing technique, is used on the estimates, to further minimize the presence of data artifacts, principally muscle noise, in the EEG record. The moving median filter uses a moving square window of 11 overlapping power estimates for each frequency.

D. *Off-line neural net training.* An automated search procedure selects optimum EEG frequencies and time-smoothing parameters for each subject. A two-layer perceptron network is then trained to use smoothed power estimates at these frequencies to optimally predict local error rate, defined as the percentage of targets not responded to (i.e., lapses) in a moving time window. A 95-s square window is advanced through the performance record in steps of 1.64 s to create the local error rate measure from the sequence of behavioral responses (hits and lapses). Intertarget intervals are randomized in order to prevent perseveration in subject responses, and therefore smoothing requires use of a specially designed smoothing program that produces smoothed, regularly spaced estimates from irregularly spaced data.

   Inputs to the neural network estimation algorithm consist of smoothed and median-filtered EEG power estimates at a set of five EEG frequencies. These frequencies are chosen for each subject on the basis of previously published coherence results (Makeig & Inlow, 1993). A sigmoidal activation function is used at the single second-layer node. The networks are trained using a standard backpropagation method to predict the smoothed local error rate measure. A training script directs a commercial neural net simulation package (SN2) to train the perceptron network. We have recently shown that such networks give better estimation performance for this problem than linear regression (Venturini, Lytton, Sejnowski, Inlow, Elliott, & Makeig, submitted). For the demonstration experiments, this network was trained on data from the first sessions of each subject, producing an individualized alertness estimation network for each subject.

E. *Real-time neural network estimation.* During subsequent feedback sessions, the weights from the previously trained network for each subject were read into a real-time feed-forward neural network module that computes a local error rate estimate from the smoothed real-time spectral estimates every 1.64 s.

F. *Auditory feedback.* During feedback sessions, whenever estimated error rate rises above a preset level (e.g., 40%), the system synthesizes a 1-s burst of square-wave sound using the clock board of the Concurrent computer. An analog sound mixing board then feeds this signal into the subject's headphones. The alarm repeats once each 1.64-s epoch during which the error rate estimate remains above the alarm threshold. To maximize alarm impact on the subject, if the alarm sounds for more than 8 epochs in a row it is disabled for another 16 epochs, unless the error rate estimate first dips below, then rises above the alarm threshold.

F. *Graphic display of the estimation process:* Real-time graphic displays help the experimenters monitor the experiments, and provide a dynamic means of demonstrating the operation of the system. A set of multiwindow real-time graphics modules display the raw incoming EEG and the resulting power spectral values input to the estimation network. These spectral values are overplotted on a plot of the correlations between power at each EEG frequency and performance, during the training experiment used to train the network.

Phase II - Demonstration and Evaluation

As a first demonstration and evaluation of the system, 30 experimental sessions were conducted. In some of the sessions, real-time auditory feedback was delivered to the subject. These experiments were the first of an continuing series whose goal is to enhance and evaluate the operation of the prototype AMM system.

Methods

Subjects

Six volunteers, three male and three female, participated in dual-task auditory and visual simulations of a passive sonar target detection environment. Two of the subjects were prospective students in a Navy sonar course. The remaining subjects were members of the laboratory staff. Ages ranged from 18 to 36 years (M = 24.5 years). None of the subjects reported any known hearing loss; the three males had all passed standard Navy hearing tests.

Stimuli

A Concurrent Real Time Unix computer system synthesized the auditory stimuli using a 16-bit D/A converter sampling at 50 kHz. The auditory stimulus streams used in the experiment are shown schematically in Figure 2. The stimuli were presented binaurally, in a white noise background at 63 dB (A-weighting), through headphones mounted in isolating cuffs. For continuity with the existing experimental design, task-irrelevant probe tones of two frequencies (568 Hz and 1098 Hz) were presented in random order with interstimulus intervals of 2 s to 4 s at 72 dB. Probe tones were 50 ms in duration, with rise and fall times of 10 ms. High tones were presented more frequently (probability of occurrence 80%) than were low tones (20%).

Auditory targets were 300-ms noisebursts with rise and fall times of 150 ms and 110 ms, respectively. Noiseburst targets occurred in 50% of the 2-s to 4-s interstimulus intervals (i.e., at a mean rate of 10/min) and were presented at approximately 6 dB above the noise background. The auditory feedback alarm consisted of a 200-Hz square wave tone, 800 ms in duration, presented at approximately 75 dB. The auditory stimuli were similar to those used in earlier alertness monitoring experiments (Makeig & Inlow, 1993), except that in the present experiments, auditory steady-state response stimulation was not used.

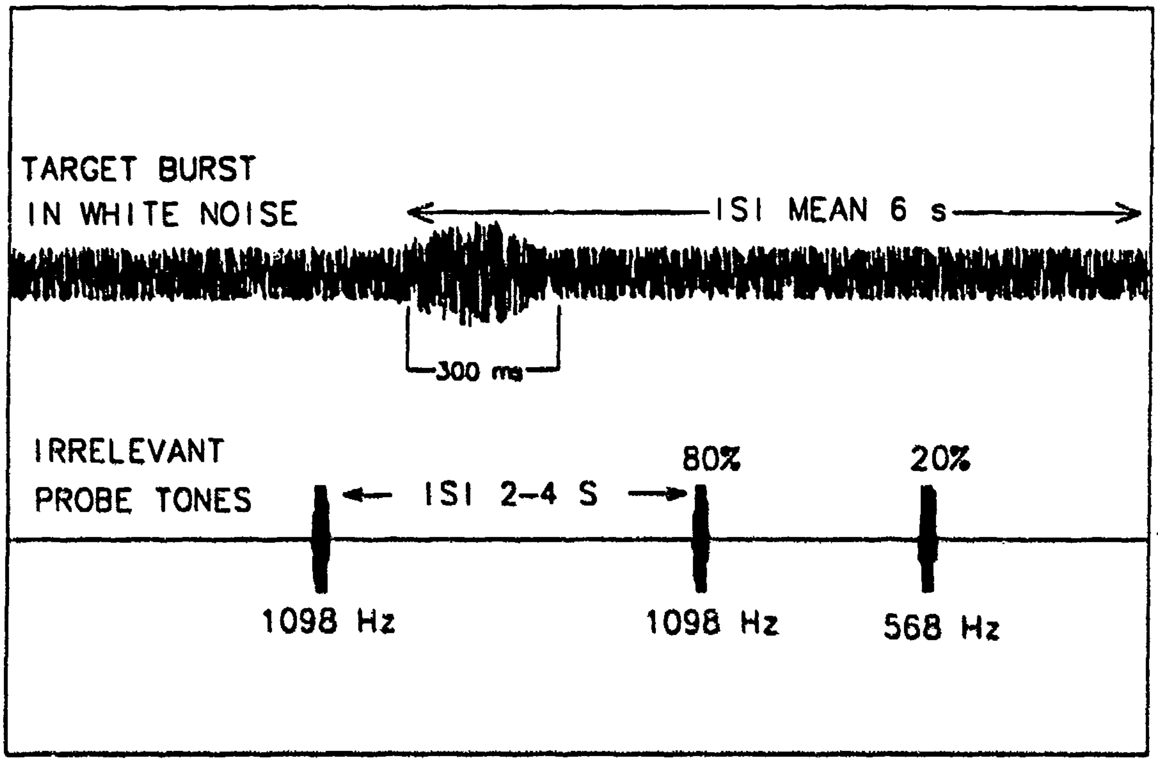

Figure 2. Schematic depiction of the auditory stimuli used in the experiments. Frequent (80%) and Rare (20%) tones were presented at random interstimulus intervals of 2 s to 4 s. In half the intertone intervals, noiseburst targets were presented 6 dB above their threshold in the 63 dB white noise background, which was continuously presented during the experiments. Auditory alarm signals consisted of 1-s, 75-dB, 200-Hz square-wave stimuli. During feedback sessions, these were presented every 1.64 s that the estimated error rate for the auditory detection task rose above a preset threshold (normally 40%).

Visual stimuli are produced using a PC (Everex EX-3000R) with a CTX 14-inch, VGA Color Monitor (CVP-5468). The display was 13-cm wide by 9-cm high. The display background was composed of 1-mm squares filled with randomly assigned greyscale values, the result resembling visual television noise ("snow"). This display was updated twice each second by adding new lines of squares to the top of the screen and scrolling the existing display down one line, creating a moving "waterfall" effect. Visual targets, introduced at a mean rate of 1/min, consisted of 20 consecutive white squares forming a vertical line. The target line was embedded in the clutter of the display background, and scrolled down the display from top to bottom with the rest of the display.

Procedure

Each subject participated in five 28-min experimental sessions on separate days. Subjects sat in a chair with their right index and middle fingers resting on the visual and auditory target response buttons, respectively. The chair, which was located in a sound-resistant booth, was positioned so that subjects faced a 39-by-48-cm window in the booth's wall. The CRT monitor was positioned approximately 110 cm from the subjects' eyes, so the display background subtended approximately 7 degrees of visual angle.

At the beginning of each session, the experimenter read a set of standardized instructions. Subjects were instructed to press one of two response buttons (auditory or visual) whenever they detected an auditory or visual target. The subjects then were given from 2 min to 5 min of training on the tasks in three stages: auditory alone, visual alone, and auditory plus visual.

The first three sessions for each subject were training sessions. The estimation network was then trained on EEG spectral and performance data from the training session that contained the most failures to respond (lapses). The training procedure generated sets of coefficients for an error estimation algorithm matching individual subjects' EEG dynamics. In subsequent feedback sessions, this individualized estimation algorithm was applied to spontaneous EEG to estimate changes in mean current error rate in a 95-s moving exponential time window. When the error estimate exceeded a predetermined threshold level (most often 40%), auditory alarm signals were delivered to the subject via the subject's headphones.

Data Collection

Data from all sessions were continuously recorded to disk for off-line analysis. Standard Grass EEG amplifiers (bandpass 0.1- to 100-Hz) amplified EEG, electrooculogram (EOG), and electrocardiogram (ECG) signals 50k, 10k, and 3k times, respectively. Biopotential data were multiplexed with response button press information and converted to 16-bit digital format at a sampling rate of 312.5 Hz. EEG data were recorded from two midline sites, one frontal (Fz) and the other midway between parietal and occipital sites (Pz/Oz), using 10-mm gold-plated electrodes referenced to the right earlobe. Periocular electrodes were used to record electrical potentials generated by eye movements. One electrode was placed at the outer canthus of the right eye and another about 0.5 cm above the left eyebrow. The ECG data were recorded from an electrode placed approximately 4 cm below the left clavicle. Electrical impedance at all electrode sites was adjusted to less than 5 kOhms by cleaning the skin under the electrodes with abrasive gel. Before each experiment, the EEG amplifiers were calibrated to compensate for any differences between actual and nominal gain values.

Analysis

Auditory targets were classified as Hits if the subject pressed the Auditory response button within 120 ms to 3000 ms of target onset, or were classified as Lapses if they did not. Responses to visual targets were considered Hits if the subject pressed the visual target button while the target was on the screen (i.e., within 20 s after it began to appear). Auditory performance was quantified by smoothing with a 95-s exponential window whose gain varied smoothly from 1.0 at the leading edge to 0.1 at the trailing edge. Visual performance was smoothed with a 155-s bell-shaped (Papoulis) window which was treated as non-causal (i.e., "now" was considered to be the middle of the visual performance window). Smoothed estimates of error rates and EEG power at 81 EEG frequencies were produced at intervals of 1.64s.

Moving-averaged EEG spectral power was calculated as follows: Hanning-windowed data windows of 256 points with 50% overlap were converted to the frequency domain using 512-point Fast Fourier Transforms (FFTs). Prior to transforming to spectral power, a simple out-of-bounds test rejected data epochs containing gross muscle artifacts. Median filtering was then performed on the accepted epochs using a moving window of 11 overlapping power estimates at each frequency. When a data window contained one or more out-of-bounds value, the latest accepted values were substituted prior to median filtering

For each session, power at each EEG frequency was correlated with changes in local error rate on the auditory task. To optimize this correlation for the purpose of real-time alertness estimation, the correlation was calculated separately for spectral power values smoothed with a set of 95-s exponential windows with different decay rates. The decay rate giving the best correlation was selected, along with five EEG frequencies for use in the estimation network. These were the most highly correlated single frequencies in each of five frequency bands whose power was shown in earlier experiments (Makeig & Inlow, 1993) to be most sensitive to changes in detection performance.

To train the neural network for each subject, EEG power from a training session (without feedback) was used to train a two-layer perceptron neural network. Data from each session consisted of 1024 spectral and error rate estimates at 1.64-s intervals. Of these, 260 were selected for testing the generalization of the network training results to determine a stopping point for the net training. EEG power values were normalized for import to the network by linearly transforming the 20th and 80th percentiles of the power distribution in the training session to range from -1.0 to 1.0.

In subsequent sessions, to compute the continuous estimate of local error rate, power at the same five frequencies used to train the network was first smoothed with the same causal 95-s exponential windows selected to train the network, normalized using the same scaling parameters, and delivered to the feedforward neural network software.

Feedback

In experiments employing auditory feedback, the local error rate estimate was compared against the alarm threshold value. In most of the feedback sessions this threshold was set to 40% estimated errors. If the estimate exceeded the threshold, a 0.8-s, 200-Hz square wave burst at 75 dB was synthesized by the computer's clock board, mixed with the auditory stimuli, and delivered to the subject. To avoid monotony and subsequent loss of arousal value, a rule was implemented that once the alarm had sounded for 8 consecutive epochs, it was silenced for 16 more epochs unless the alertness estimate fell below and then returned above its threshold value. Alarm history was saved for later review.

Results

Subject Performance

Local error rate, the performance measure used in these experiments, is not a uniquely defined measure. In the summary plots below, a 155-s (95-epoch) noncausal bell-shaped Papoulis window was used to compute local error rate. Total percentage of correct responses to auditory targets in the 20 experimental sessions ranged from 99.6% to 45.4%. Mean auditory error rates for the 20 sessions, and visual error rates for 18 sessions (shown in Figures 3a and b), followed the same trend as earlier experiments, with group mean error rate increasing for the first 10 min and remaining stable thereafter. (Visual data for two sessions were lost because of equipment failure). Although earlier experiments had revealed a tendency for a monotonic increase in mean error rate across repeated sessions, in these sessions no such trend was observed.

Note the strong similarity (r = 0.73) between the mean auditory and visual performance records. These results confirm results of our earlier auditory detection experiments (Makeig et al., 1990), and of classical vigilance experiments (Mackworth, 1948). In individual sessions, however, local error rates fluctuated widely, as shown in Figure 4, in which each local error rate trace may vary from a low of 0% (all correct) to a high of 100% (no responses). Note in Figure 4 that the four subjects have different mean error rates and no consistent pattern of fluctuations within sessions. On average, two of the subjects (d1 and d2) had lower mean error rates than the other two (d3 and d4).

Error Rate Estimation

Figure 5 below shows estimated and actual error rates for one feedback session each from three of the four subjects. A qualitative fit between the estimated and actual error rates is apparent. In each case, during segments in which estimated error rate was above the alarm threshold, local error rate was high relative to the rest of the session.

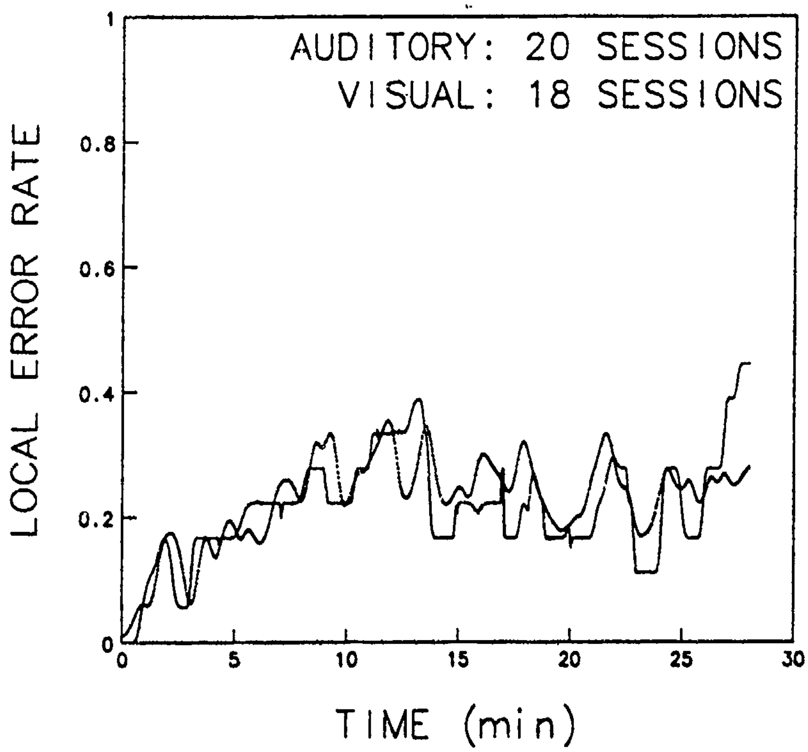

Figure 3. Group means in auditory and visual task error rates. For the auditory task, *(dotted line)* average of 20 sessions on four subjects. Visual error rate data *(solid line)* from two sessions was lost due to mechanical difficulties, so the visual error rate curve averages data from 18 sessions. These curves reproduce those found in earlier experiments (Makeig & Inlow, 1993), and follow trends reported in previous vigilance research.

Discussion

These experiments have demonstrated the feasibility of objective real-time alertness estimation in a laboratory AMM system by delivering auditory alarm signals to operators when estimated alertness became lower than a preset threshold. Alertness level is difficult to estimate subjectively, since loss of alertness is accompanied also by loss of self-awareness and memory. This may also be one reason that individual fluctuations in alertness have not been studied much in previous vigilance research. Nevertheless, at the end of the three sessions shown in Figure 5, each of the three subjects reported that the auditory feedback helped them to maintain alertness by giving them objective confirmation of changes in their levels of alertness.

However, differences between the estimated and actual error rate curves remain larger than expected on the basis of Monte Carlo simulations and earlier work using linear regression (Makeig & Inlow, 1993). As the system design was changed during the sessions, quantitative

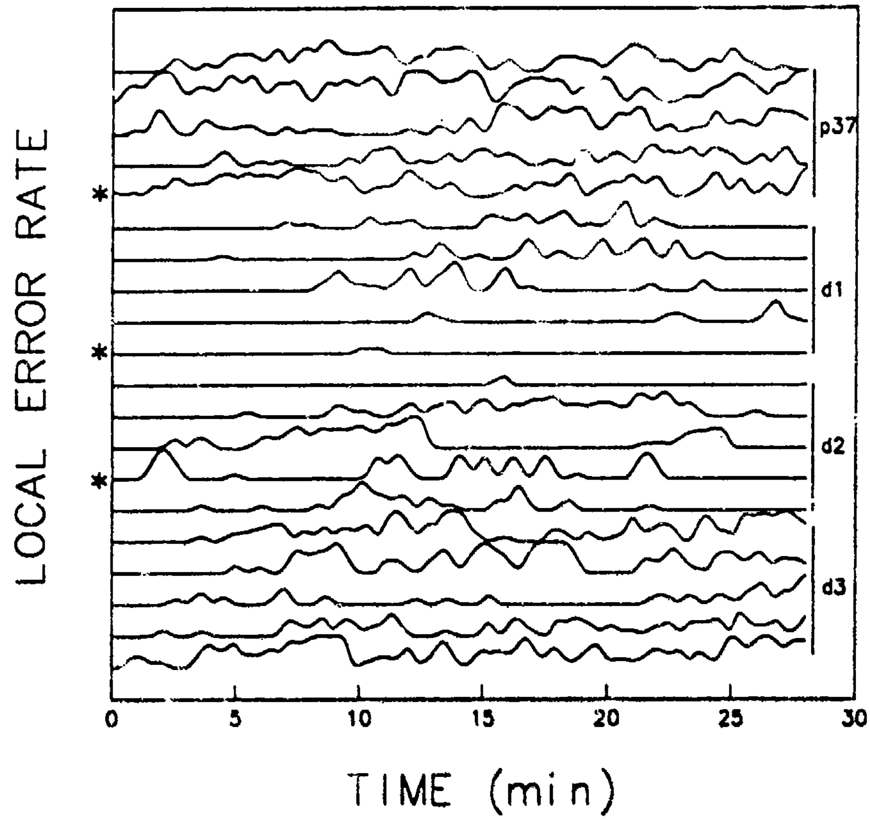

Figure 4. Local error rate time series for individual sessions. Note that despite the predictability of the group mean rates in Figure 3, individual error rate dynamics vary considerably. The three starred runs are examined in more detail in Figure 5 below.

performance statistics will be reported later, after a larger number of subjects are run using the same version of the system. In the two feedback sessions for the fourth subject, the error rate estimate did not follow the actual error rate or subjective alertness changes accurately enough for the subject to report that feedback had been useful in managing his alertness. Recently, we trained a network using Version 1.1 of the system and tested this net on data from this subject's other four sessions. Results suggest that Version 1.1 is capable of producing useful alertness estimates for this subject. Review of the results of 30 sessions has prompted several ideas for improving the alertness estimation and training procedure:

1. In Version 1.0 of the system, which was tested on the first four subjects (Fig. 5), the smoothing window used to compute local error rate and train the neural net algorithms consisted of a two-sided or noncausal bell-shaped window. In real-time spectral estimation, however, noncausal filtering cannot be used, because it requires "future" data which become available only after an unacceptable delay. In the real-time system, therefore, a trailing or causal exponential window was used to smooth the spectral data. In system Version 1.1, which so far has been tested on only two subjects (figs. 6 and 7), the error rate and EEG windows used to train and run the estimation network are causal.

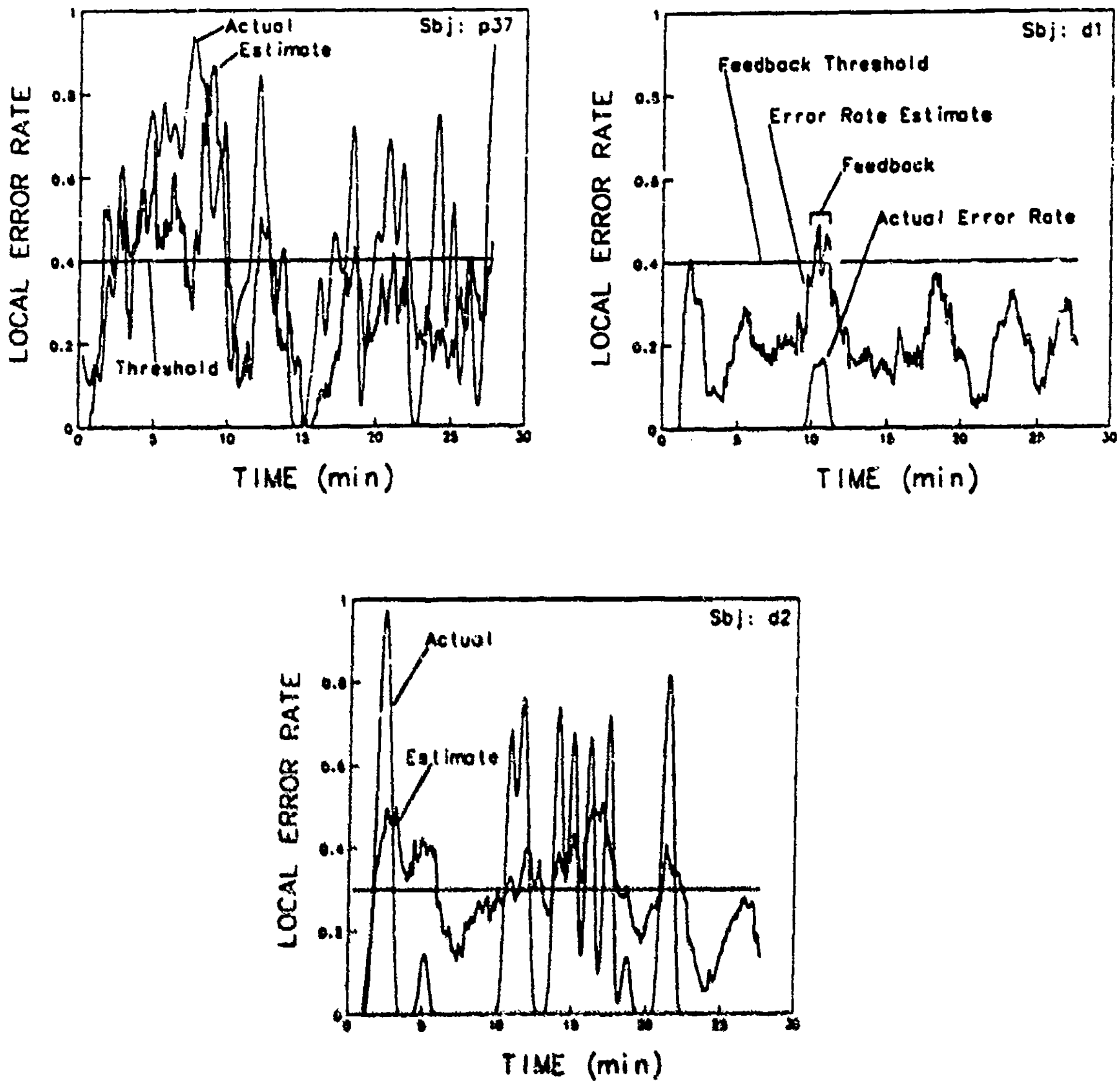

Figure 5. Actual and estimated local error rate for the auditory task for three feedback sessions on three of the four subjects. The alarm threshold level is also shown on each plot. (See text for further details).

2. In Version 1.0 of the system, there were differences between the artifact rejection and median smoothing in the computation of the EEG spectra for training and running the networks. In Version 1.1, both algorithms are identical.

3. The subjects judged the feedback alarm signals used in these experiments as sufficiently arousing on first presentation. However, in Version 1.0 the signals tended to mask auditory targets that were sometimes presented simultaneously. Therefore, in Version 1.1, the alarm signals are shortened. In future versions of the system, the alarms should increase in intensity or urgency if the error rate estimate remains high. The subject should also be able to control the volume of the alarm. Finally, use of more complex auditory feedback, including prerecorded voice messages, should be explored. In more advanced versions of the system, continuous visual alertness information could be made available to the subject.

4. It may be that improved estimation accuracy can be attained if the neural network used power at more than five EEG frequencies. and/or if it employed a more complex structure and training procedure. Other network structures and training procedures are being investigated.

5. While monitoring from a single scalp channel might be optimum form a practical point of view, it is likely that significant improvements in statistical power and noise cancellation can be achieved by collecting data from two or more scalp channels, by monitoring evoked responses to the task-irrelevant probe tones (Makeig et al., 1990) and/or by monitoring other psychophysiological channels (such as eye movements and heart rate). Data from multiple physiological information channels could then be combined in a more advanced alertness estimation algorithm. We are now evaluating these possibilities.

6. Earlier research (Makeig & Inlow, 1993) showed that to ensure accurate error rate estimation, the training data must contain a wide range of local error rates. For at least two subjects (d2 and d4), local error rate remained low throughout the training sessions and therefore could not be used to train the estimation network. Future experiments should be long enough, and/or should manipulate subject sleep, caffeine intake, etc., so as to maximize the range of local error rates seen in the training sessions. Optimum methods of taking advantage of partial between-subject similarities to combine training session data from more than one subject should also be studied.

Test of First Enhancements

To date, Version 1.1 of the system has been tested on two subjects. The fifth subject, a professional sonar operator, did not produce enough detection lapses in his training sessions to allow an estimation algorithm to be trained on his data. Results of one feedback run on the sixth subject are presented in Figures 6 and 7. This subject was also a sonar operator (but unlike the previous subject, not a coffee drinker). Figure 6 shows actual and estimated error rate for a 28-min session with alertness feedback. The alarm threshold and alarm record are also shown, along with head drop, which was noted from the videotape taken of the subject during the session. The estimation algorithm used power at five EEG frequencies chosen and smoothed individually for the subject on the basis of a previous training session. As anticipated, at the beginning of the run, the estimation algorithm needed 95 s of data to begin making accurate error rate estimates.

In Figure 6, note the following points: After the session, the subject resported that the early alarm (at minute 2) was appropriate, since he was feeling drowsy at the time, and actually missed one auditory target. The subject stated that the alarm reminded him to be more attentive. The error rate estimate began to rise in minutes 5 to 9, i.e. about 4 minutes before the subject had another response lapse. Head drop became visible almost 12 min after the alarm threshold was reached by the error rate estimate (at minute 10), and the subject began making errors. The estimation algorithm accurately estimated the subject's return to alertness at minute 19, which was aided by an unplanned verbal query to the subject from the experimenter, who thought the sound delivery system might be malfunctioning. Error rate during the final portion of the run (minutes 20 to 28) was accurately tracked by the system without use of any information except the subject's EEG spectrum.

Figure 7 shows a *post hoc* alertness estimate of error rate in this same session, computed from EEG power estimates at the same five frequencies and using the same smoothing constants as were used in the real-time algorithm (Fig. 6). The same training session was also used to train the revised estimation net. In the revised net, however, normalization constants were based on two training sessions rather than one. In Figure 7, note that the root mean square (RMS) error between the estimated and observed error rates (0.08) is less than that observed in Monte Carlo simulations using a predetermined error rate probability curve to generate stochastic series of Hits and Lapses with the same temporal statistics as in the actual experiments. Hence, statistically, the prediction error in Figure 7 is as small a possible, and further improvement in estimation accuracy cannot be expected. The extremely close fit of the estimated and actual error rate record in Figure 7 demonstrates that precise information about the exact alertness trend is contained in the EEG power spectrum, as reported by Makeig and Inlow (1993). However, power normalization remains an important area for further research.

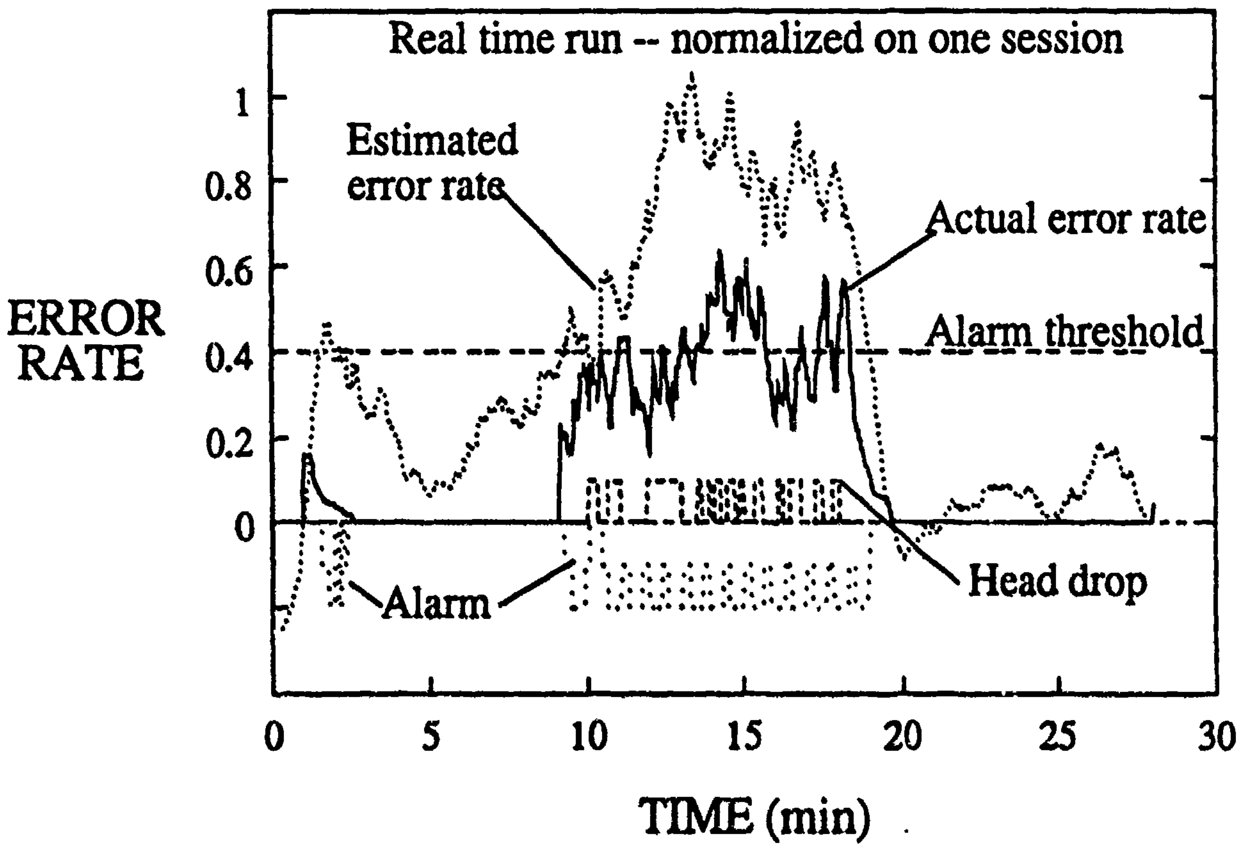

Figure 6. Results of a session on a fifth subject employing real-time alertness monitoring and feedback. In this experiment, the feedback alarm threshold was set at 0.4. The actual alarm signal record is shown in the lowest trace. The square pulses in the dashed trace above it indicate moments when the subject's head was observed to drop in the videotape of the session. The dark trace shows the actual local error rate on the auditory detection task, smoothed using a 95-s exponential window whose center of gravity trailed the leading edge of the window by 22 s. The top trace shows the error rate estimate computed from the subject's EEG spectrum during the run.

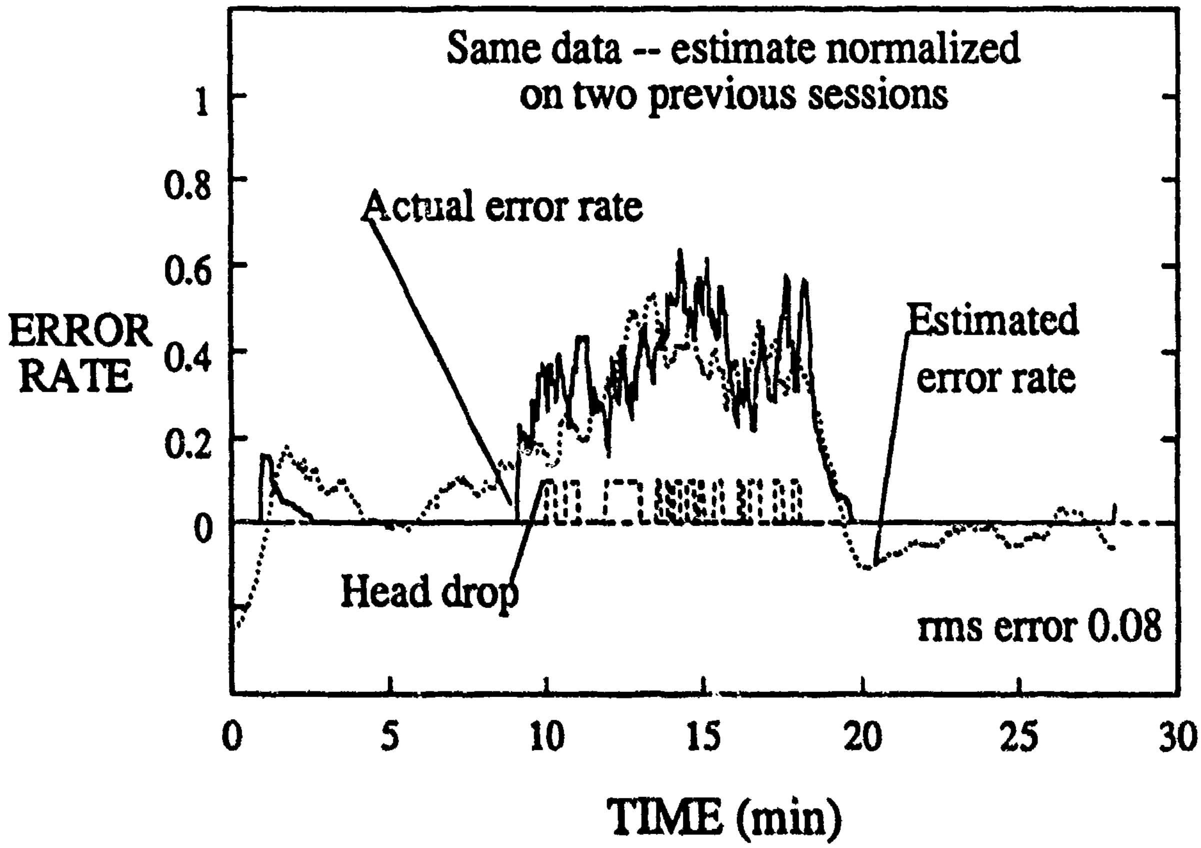

Figure 7. This figure demonstrates that significant improvement in estimation accuracy would have been obtained had the normalization constants for preprocessing the smoothed EEG power been computed using data from two training runs rather than from just one. The figure shows *post hoc* error rate estimation using the same EEG data as in Figure 6, but with revised normalization parameters computed using Version 1.1 of the training algorithm on the basis of two previous training runs on this subject. Estimation accuracy for this session cannot be distinguished from measurement error for the local error rate measure.

Conclusions

First laboratory versions of an individualized Alertness Monitoring/Management (AMM) system have been demonstrated and have been tested, to date (January, 1993), in a total of 30 half-hour test sessions on six subjects. Results of these experiments suggest that sufficient information is available in the EEG spectrum recorded from a single scalp derivation to accurately estimate probability of detection in an auditory detection task performed by subjects concurrently with a visual task simulating one aspect of sonar operations. Future research will attempt to determine the optimum number and location of electrodes, the best way to collect and process individualized net training data, the generalization of the alertness estimate to performance on more complex tasks, and ways to combine more than one channel of psychophysiological information to make the alertness estimates more accurate and noise-resistant. Practical questions regarding easily applied electrode montages and automatic artifact cancellation also remain to be investigated. If reliable and convenient automatic alertness monitoring can be developed, the range of potentially useful applications appears to be wide.

References

Beatty, J., Greenberg, A., Deibler, W.P., & O'Hanlon, J.F. (1974) Operant control of occipital theta rhythm affects performance in a radar monitoring task. *Science*, 183:871-873.

Belyavin, A. & Wright, N.A. (1987) Changes in electrical activity of the brain with vigilance. *Electroenceph. clin. Neurophysiol.*, 66:137-144.

Hillyard, S. A., & Johnston, P. (submitted) *Event-related brain potentials as predictors of target detection performance in a moving waterfall display simulating passive, braod-band sonar monitoring*, (Technical Report). San Diego, CA: Naval Medical Research and Development Command, Naval Health Research Center.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quart. J. Exper. Psychol.*, 1, 6-21.

Makeig, S., Elliott, F. S., Inlow, M., & Kobus, D. A. (1990). *Predicting lapses in vigilance using brain evoked responses to irrelevant auditory probes* (Technical Report 90-39). San Diego, CA: Naval Medical Research and Development Command, Naval Health Research Center.

Makeig, S., & Inlow, M. (1993). Lapses in Alertness: Coherence of fluctuations in vigilance and the EEG spectrum. *Electroenceph. clin. Neurophysiol.*, 86:23-35.

Makeig, S., Inlow, M., & Galambos, R. (1991). Coherence of EEG spectral changes and vigilance. *Soc. Neurosci. Abst.*, 17(1), 833.

Matousek, M. & Petersen, I. (1983) A method for assessing alertness fluctuations from EEG spectra. *Electroenceph. clin. Neurophysiol.*, 55:108-113.

Stroud, J. M. (1966). The fine structure of psychological time. *Ann. N.Y. Acad. Sci.*, 138, 623-631.

Venturini, R., Inlow, M., Makeig, S., Elliott, F.S., Lytton, W.W., & Sejnowski, T. (submitted) *Monitoring alertness: a neural network perspective.* (Technical Report). San Diego, CA: Naval Medical Research and Development Command, Naval Health Research Center.