Old Dominion University Research Foundation

DEPARTMENT OF PSYCHOLOGY
COLLEGE OF SCIENCES
OLD DOMINION UNIVERSITY
NORFOLK, VA 23529

BIOCYBERNETIC CONTROL OF VIGILANCE TASK PARAMETERS

By
Dr. Frederick G. Freeman, Principal Investigator
Department of Psychology

FINAL REPORT
For the period ending December 31, 1999

Prepared for
NASA Langley Research Center
Attn.: Dr. Alan T. Pope
Technical Officer
Mail Stop 152
Hampton, VA 23681-2199

Under
NASA Grant No. NAG-1-2105
ODURF File No. 183881

February 2000

Final Report

NASA Grant No. NAG-1-2105
ODURF No. 183881

Biocybernetic control of Vigilance Task Parameters

Dr. Frederick G. Freeman, Principal Investigator
Department of Psychology
Old Dominion University

Submitted by
Old Dominion University
Research Foundation
800 West 46th Street
Norfolk, Virginia 23508

Submitted to
NASA Langley Research Center
Hampton, VA 23681-2199

February 2000

FINAL REPORT

NASA Project: NAG-1-2105

Title: "Biocybernetic Control of Vigilance Task Parameters

ODU Project# 183881

Introduction

The major focus of the present proposal was to examine psychophysiological variables that are related to hazardous states of awareness induced by monitoring automated systems. With the increased use of automation in today's work environment, people's roles in the work place are being redefined from that of active participant to one of passive monitor. Although the introduction of automated systems has a number of benefits, there are also a number of disadvantages regarding the worker performance (Wiener, 1988). Byrne and Parasuraman (1996) have argued for the use of psychophysiological measures in both the development and the implementation of adaptive automation. While both performance based and model based adaptive autmation have been studied (e.g. Parasuraman, Mouloua, and Molloy, 1996), the use of psychophysiological measures, especially EEG, offers the advantage of real time evaluation of the state of the subject.

Previous investigations of the closed-loop adaptive automation system in our laboratory, supported by NASA, have employed a compensatory tracking task which involved the use of a joystick to maintain the position of a cursor in the middle of a video screen. This research demonstrated that, in an adaptive automation, closed-loop environment, subjects perform a tracking task better under a negative, compared to a positive, feedback condition. While tracking is comparable to some aspects of flying an airplane, it does not simulate the environment found in the cockpit of modern commercial airplanes. Since a large part of the flying responsibilities in commercial airplanes is automated, the primary responsibility of pilots is to monitor the automation and to respond when the automation fails. Because failures are relatively rare, pilots often suffer from hazardous states of awareness induced by long term vigilance of the automated system. Consequently, the aim of the current study was to investigate the ability of the closed-loop, adaptive automation system in a vigilance paradigm. It is also important to note that tracking involves a continuous, though low level, motor response. Since it is not clear how such activity might affect performance of the adaptive automation system, it was thought to be important to evaluate how the system functioned when there was minimal motor output by the subjects.

The current study used the closed-loop system, developed at NASA-Langley Research Center, to control the state of awareness of subjects while they performed a vigilance task. Several experiments were conducted to examine the use of EEG feedback to control a target dimension used in the task. Changes in a subject's arousal, as defined by specific EEG indexes, produced stimulus changes known to affect task performance. In addition, different electrode sites, compared to previous research, were sampled to determine the optimum configuration with regard to the following criteria: 1. task performance and 2. EEG index.

In the first experiment, subjects performed a vigilance task in which the event rate was altered as a function of the subjects' state of alertness based on an EEG index. Subjects were run in the same vigilance paradigm that we previously used, in which the event rate was changed as a function of the subjects' engagement index. However, the sites from which EEG was recorded differed, with two electrodes placed over the frontal lobes and two over the occipital lobes.

In our initial research involving a tracking task, defining the conditions which would produce increased arousal were more straight forward than in the vigilance monitoring task. Thus, maintaining the subject in the manual tracking mode was assummed to produce higher levels of arousal compared to the automatic mode. In a vigilance task, however, a variable which might intuitively produce higher arousal levels might also result in poorer performance (e.g. an increased event rate). Thus, even though positive and negative feedback do not necessarily apply to the use of the closed-loop system in a vigilance paradigm, it was necessary to run different groups of subjects for whom the opposite contingencies applied regarding the manipulation of task parameters. Thus, if one group has a task made easier because their index indicates a state of increased arousal, a second group should have the task made harder when such a state occurs. Low arousal should result in the opposite conditions. Since there is no data on how constant low and high event rates might affect the EEG index, these two control conditions were also evaluated.

Finally, it has been theorized that the frontal lobes are important in controlling sustained attention. Alpha activity, however, is more commonly seen over the occipital lobes. Most of the research we have conducted with the closed-loop has involved recording from parietal/central sites. Thus, the contributions of different sites and bandwidths recorded from frontal and occipital sites were evaluated.

Method

Participants

Seventy subjects ranged from 18 - 40 years in age. All subjects were either undergraduate or graduate students and were given either extra credit for their class or $30.00 for participating in the experiment.

Engagement Index

The engagement index employed was $\beta/(\alpha+\Theta)$. This was the index found to be most effective by Pope, et al. (1995). EEG was recorded from four sites: F3, F4, O1, and O2. The combined theta, alpha, and beta power from these four sites was used to derive the index. The mean index value was first determined over a 10-minute baseline practice session, along with the standard deviation. At the beginning of the task, a 20-second window (10 2-second epochs) was used to determine the subject's initial index value. It was then updated every two seconds using a sliding 20-second window. Deviations of this index, above and below the baseline index, were derived and used to determine the stimulus event rate in the vigilance task (see below).

Apparatus

EEG was recorded using an Electro-cap International lycra sensor cap. The cap consists of 22 recessed tin electrodes arranged according to the International 10-20 system (Jasper, 1958). EEG was recorded using a BIOPAC EEG100A differential amplifier module consisting of four, high gain, differential input, bio-potential amplifiers. The low and high pass filters were set at 100 and 1 Hz, respectively.

The EEG100A was connected to a Macintosh Quadra. Each amplified EEG channel was digitized at a rate of 400 samples per second. The digital signals were arranged into epochs of 1024 data points (roughly two and one half seconds) prior to conversion to a spectral power form using a Fast Fourier Transform (FFT). A LabVIEW Virtual Instrument (VI) calculated total EEG power in three bands: theta (4-8 Hz), alpha (8-13 Hz), and beta (13-22 Hz). The VI also performed the engagement index calculations and commanded the task mode changes through serial port connections to the task computer.

An artifact rejection subroutine examined the amplitudes of each epoch from the four channels of digitized EEG and compared them with a preset threshold. If the voltage in any channel exceeded the threshold by more than 25%, the epoch was marked as bad and the calculated index was replaced with a value of zero. These epochs were then ignored when computing the slope of the index. The data record resulting from an epoch containing an artifact was marked when it was written to the data file so that is could be ignored when later data analysis was done.

Subjective workload was assessed using the NASA-TLX (Hart & Staveland, 1988). This is a multi-dimensional scale that requires participants to rate tasks on six 20-point scales (Mental Demand, Physical Demand, Effort, Temporal Demand, Performance, Frustration) that are thought to be major components of subjective workload. The raw NASA-TLX requires the sum of the six subscales.

Task

Subjects performed a vigilance monitoring task which consisted of the presentation of two white vertical lines presented for 200 msec on a blue background at the rate of 20 times per minute. The size of the neutral stimuli was 2 X 72 mm separated laterally by 26 mm and subtending a visual angle of eight degrees. Critical signals were represented by an occasional increase to the top of the pair of lines. EEG was recorded continuously during the vigilance task and fed into the LabView virtual instrument which calculated the $\beta/(\alpha+\Theta)$ index.

Under the negative feedback condition, if this index was more than .4 standard deviations below the baseline (i.e. decreased arousal), determined during the five minute practice session, the frequency at which the stimuli were presented was increased from the initial 20 times per minute to 60 times per minute. If the index returned to within +/- .4 standard deviations of the baseline the stimulus event rate returned to 20 times per minute. If the index increased to more than .4 standard deviations above the baseline the event rate was decreased to six times per minute.

Under the positive feedback condition, if this index was more than .4 standard deviations above the baseline (i.e. decreased arousal), determined during the five minute practice session, the frequency at which the stimuli were presented was increased from the initial 20 times per minute to

60 times per minute. If the index returned to within +/- .4 standard deviations of the baseline the stimulus event rate returned to 20 times per minute. If the index increased to more than .4 standard deviations below the baseline the event rate was decreased to six times per minute.

A yoked control group was run for both the negative and the positive feedback conditions to evaluate the experimental subjects' performance. A subject in these groups was yoked to a specific subject in either the negative or positive group such that the order of the different event rates, as well as the duration of time spent at each event rate, was exactly the same as that of the experimental subject to whom he/she was yoked. However, the EEG of the yoked subject had no effect on the event rate presented to the subject. Thus, experimental and yoked subjects experienced the same stimulus conditions, but the nature of the conditions was controlled by the EEG index of the experimental subjects.

A random control group was also run. The stimulus presentations for these subjects was determined by taking the stimulus presentations to the positive feedback subjects and randomizing them. Thus subjects in this group received the same number of event rates as the positive group but the order was completely randomized. As with the yoked subjects, these subjects' EEG had no effect on the event rate.

Two other control groups were run. The low rate group was run the entire experimental session using the six per minute event rate. The high rate group was run the entire session using the 60 per minute event rate. EEG was recorded for these subjects but did not have any effect on the event rate.

Procedure

Upon entering the experimental suite, subjects were explained the nature of the experiment and asked to sign an informed consent form. The electrode cap and reference electrode were then attached. The left mastoid area was used as the reference. The impedance levels at the four recording sites and mastoid area were reduced below 5 kOhms.

Subjects were then seated approximately .5 m in front of the monitor on which the vigilance task was displayed. The nature of the task was explained and the subjects were allowed to practice for 10 minutes. All subjects used their right hand for the task. Following the 10-minute practice, the 20-second baseline recording was taken and the task was started. The task lasted 40 minutes.

Results

The major performance dependent measures were the reaction time to target stimuli, proportion of hits to target stimuli, D', β, A' scores, and B" scores. A' and B" are nonparametric measures comparable to D' and β used in signal detection research and is a function of the number of hits and false alarms. In addition, EEG index values and power in the beta, alpha, and theta bandwidths for the experimental subjects were compared for the different task levels. For clarity, the results for the low rate and high rate groups will be presented separately from the other groups.

Performance

As seen in Figure 1, the reaction time to targets was fairly stable for the first two 10-minute periods. For the third and fourth 10-minute periods, reaction time increased slightly for negative feedback group and its yoked control but decreased for the positive feedback and the random control. For the positive yoked group there was a marked increase for the fourth period, though this appears to be due to two outliers. There was a significant interaction effect between groups and periods $\underline{F}(12,135)= 1.97$, $\underline{p} < 0.05$. There was neither a group nor a period effect.

In Figure 2 is presented the A' scores as a function of period and group. It can be seen that for all groups, A' decreased from the first to the second period, though more dramatically for the positive feedback group and its yoked control. However, for the negative feedback group and its yoked control, A' actually increased slightly from the second to the fourth period, while for the other groups the A' scores leveled off so that by the fourth period these groups were performing much worse than the negative feedback and its yoked control. The ANOVA yielded a significant effect of periods, $\underline{F}(3,135) = 2.88$, $\underline{p} < 0.05$. There was also a marginally significant interaction between groups and period, $\underline{F}(12,135) = 1.81$, $\underline{p} < 0.06$. There was no main effect for groups.

For comparison, in Figure 3 is presented D' scores. The same pattern can be seen as was found for the A' scores. However, for D' there was a significant interaction between groups and periods, $\underline{F}(12, 135)= 2.49$, $\underline{p} < .01$.

In Figure 4 is presented the proportion of hits as a function of period and group. For the first 10 minutes, the groups performed at approximately the same levels. However, the performance of the negative feedback group actually increased from the first to the fourth period, while that of its yoked control dropped initially, then also increased. The performance of the random control group dropped moderately from the first to the second period while that of the positive feedback and its yoked control dropped below 50%. Performance then leveled off for all three groups. An ANOVA yielded a significant effect for period, $\underline{F}(3,135) = 3.84$, $\underline{p}< 0.02$, and a significant group by period interaction, $\underline{F}(12, 135) = 2.46$, $p < 0.01$. There was no main effect for group. Analyses of the false alarms, Beta, and B" data did not yield any significant effects.

For the high and low rate control groups, there was a significant group effect for reaction time, $\underline{F}(1,18)=10.68$, $\underline{p}< .01$, with the high rate group have faster reaction times across all periods. There was also a significant group and period effect for probability of hits, $\underline{F}(1,18)= 12.82$, $\underline{p}<.01$ for groups and $\underline{F}(3,54)= 3.06$, $\underline{p}<.05$ for periods. The low rate group performed significantly better across all periods. There was also a slight decrease in performance across periods. The B" scores were significantly lower for the low rate group, $\underline{F}(1,18)=28.38$, $\underline{p}<.0001$, and increased across periods, $\underline{F}(3,54)=3.04$, $\underline{p}<.05$. There was neither a significant effect of A' scores nor a group by period interaction.

EEG Index and bandwidths

For the derived EEG Index a 5 (Group) x 4 (Period) x 3 (Stimulus frequency) ANOVA yielded significant effects for each main effect and interaction excepting for a main effect for groups. In Figure 5 is presented the three-way interaction, $F(24, 270) = 4.57$, $p < .001$. It can be seen that, as predicted, the high stimulus rate yielded the highest index values, which decreased as a function of stimulus rate. Although there was a significant period effect, $F(3, 270) = 4.15$, $p < .01$, it appears from the figure that indexes were relatively stable, except for the positve feedback group, which appeared to increase over periods. The groups also appeared to be differentially affected by the stimulus rate, $F(8, 270) = 2.93$, $p < .01$. The positive feedback group and its yoked control had the highest index values. This difference increased with increases in stimulus rate, especially compared to the random control group.

The effect of stimulus rate on the derived index is further demonstrated when looking at the high and low rate control groups. The two-way interaction between groups and periods, $F(3,54) = 40.14$, $p < .001$ is presented in Figure 6. Although the groups started out virtually identically, the index for the low rate group dropped dramatically for periods two and three, but then, for some reason increased in the last period. The index for the high rate group also decreased for periods two and three but not as much as for the low rate group, then also increased for period four. In comparing the mean indexes for these two groups with those seen in Figure 5, it is obvious that switching back and forth between stimulus rates causes a much greater increase in arousal, as reflected in the index, for the high rates. When the rate remained constantly high the index fluctuated between 5.7 and 5.9. When the rates were constantly varying, the indexes fluctuated between 7 and 9.

Trying to understand the indexes by examining the individual bandwidths is very difficult. For theta, every main effect and interaction was significant for the three-way ANOVA. In Figure 7 is presented the three-way interaction, $F(24, 270) = 5.38$, $p < .001$. While the positive feedback group tended to maintain low theta levels, as might be expected, it's yoked control produced some of the highest theta levels. Conversely, the negative feedback group and it's yoked control were extremely similar, mirroring their performance. Theta levels in these groups tended to be intermediate between the positive feedback group and its yoked control and the random control. As seen in Figure 8, the exact same pattern was seen for the three-way interaction for alpha activity, $F(24, 270) = 4.39$, $p < .001$, though there was less fluctuation across periods. Oddly, an examination of beta activity, with the three-way interaction presented in Figure 9, $F(24, 270) = 5.54$, $p < .001$, shows that the positive feedback group also had the lowest beta. The negative feedback group and its yoked control were again in the middle.

For the high and low rate groups, for theta, there was also a significant two-way interaction between groups and periods, $F(3,54) = 4.23$, $p < .001$. Interestingly, the high rate group produced more theta, which increased from period one to period two, then decreased slightly (Figure 10). Conversely, the group by period interaction for alpha showed the low rate group producing higher alpha levels, which increased up to period three then decreased, $F(3,54) = 5.52$, $p < .001$ (Figure 11). For beta, the two-way interaction shows the high rate group with more beta, which increased across periods, but decreased for period four, $F(3,54) = 9.92$, $p < .001$ (Figure 12).

It would thus appear, after looking at the indexes and the different bandwidths, that negative feedback functioned as it was designed to, keeping individuals at a moderate level of arousal. Since the negative yoked group's performance was similar to that of the experimental group, it might be concluded that the changes in stimulus conditions, generated by the negative feedback group could be responsible for producing a specific state of arousal which, in turn, results in enhanced performance. This conclusion, however, will need to be verified by replicating and extending the design of the current experiment. The bandwidth data for the high and low control groups suggest that high stimulus rates induce higher states of arousal (seen with higher beta) but decreased ability to attend (higher theta). Lower stimulus rates appear to produce relatively lower arousal (lower beta and higher alpha) but relatively better attention (lower theta). This resulted in better performance by the low rate group relative to the high rate group.

Experiment Two:

Experiment two attempted to use the biocybernetic system to moderate vigilance performance by altering task demand, by varying task difficulty, based on the operator's mental engagement. Participants' EEG was recorded while they performed a vigilance task. The task again consisted of the presentation of vertical bars on a computer screen. Task demand was manipulated so that signal amplitude (i.e. height of the vertical bars) increased in the low task demand condition, and decreased under high task demand. In addition, both positive and negative feedback conditions were utilized. According to Pope et al. (1995), the purpose of negative feedback is to moderate engagement. In other words, decreasing task demand when engagement is high should bring engagement levels down to a more moderate level. Conversely, increasing task demand during periods of low engagement should increase engagement levels to a more moderate level. Such a variation in task demand resembles positive feedback in the present study. Under this contingency, the operator must be cognizant of the increase in task demand. In the present study, however, the operator may not have been aware of such a change due to the decreased discriminability between critical and neutral stimuli. Therefore, the operator needed to be informed that a critical signal had been presented. Providing Miss Knowledge of Results (KR) to indicate signal occurrence was employed in an attempt to eliminate this problem.

List of Hypotheses

1) It was expected that alerting participants with KR to missed targets would motivate them to pay more attention to the task, thus increasing the probability of correct detections. More task switches were also expected when KR was provided.

2) An interaction was also expected between feedback and KR. The best performance was expected under negative feedback/KR, with the worst performance occurring under positive feedback/No-KR. In addition, task switches were expected to occur most often during negative feedback/KR and least often during positive feedback/No-KR.

3) Under positive feedback, amplitude decreased when engagement was low. Such an increase in difficulty was expected to cause the engagement index to decrease further, especially when KR was not provided. Therefore, the engagement index was expected to be lowest during the positive feedback/ No-KR condition.

4) Participants in the positive feedback/ No-KR condition were expected to show the largest increase in perceived workload between the first and second administration of the NASA-TLX.

METHOD

Participants

Forty male and female Old Dominion University undergraduate psychology students participated in the experiment. Participants received course credit or a $30 stipend for participation. Normal (20/20) or corrected-to-normal vision was required.

Materials and Apparatus

The apparatus was the same as that used in the first experiment.

Task

The vigilance task was presented on an IBM compatible 386 personal computer with a VGA monitor. The neutral signals consisted of two vertical bars measuring 38 mm x 3 mm. Three critical signals were used: (a) low amplitude signals measured 40 mm x 3mm; (b) medium amplitude signals measured 42 mm x 3 mm; and (c) high amplitude signals measured 44 mm x 3 mm. The bars were separated by approximately 27 mm. The event rate was set at 30 stimuli per mm. Two critical stimuli appeared at random within each minute. Participants responded to the occurrence of a critical signal by pressing the space bar on a standard keyboard.

Two KR conditions were utilized: KR and No-KR. In the KR condition, a 1000 Hz tone at 68 dB sounded 1500 ms after the occurrence of a target if the participant failed to make a response. Participants in the No-KR were yoked to those in the KR condition; i.e., they heard the exact same sequence of Miss KR tones that were generated by their KR counterparts, but in the context of a different pattern of CS occurrences. In this condition, therefore, the tone was not tied to participants' performance, so it provided no performance contingent information. Therefore, this condition was used as a control manipulation.

Experimental Design

A 2 Feedback (negative, positive) x 2 KR (KR, No-KR) x 3 CS Amplitude (small, medium, large) x 4 Monitoring Period (10-min periods) mixed factorial design was implemented. Feedback and KR were manipulated between subjects, while Amplitude and Monitoring Periods were within-subjects factors.

The mean engagement index and number of task switches were the dependent variables for the EEG recordings. Vigilance performance was measured by the proportion of hits (PHIT), proportion of false alarms, (PFA), response time to hits (RTH), and the nonparametric signal detection indices, A', and B". A' ranges from 0.5 to 1.0 (Grier, 1971). However, the range for A' was set at 0.7 to 1.0 for the present study. B" ranges from -1.0 to 1.0. The NASA-TLX was used to assess participants' perceived level of mental workload.

Procedure

Upon arrival, participants were asked to read a detailed description of how they were to be fitted for the EEG cap, and how their performance was to be recorded. Alcohol pads were used to clean participants' mastoid, temples, and areas above and below their right eye. The areas were then rubbed with Omni Prep skin prepping paste. Single electrodes were secured on these sites with adhesive pads.

Participants were then fitted for the electro cap. Sites Cz, Pz, P3, and P4, which have been used in past studies that have used the biocybernetic system on tracking performance (see Freeman et. al, 1998; Pope et. al, 1995; and Prinzel et. al, 1995) were used in the present study. Impedances for all electrodes were reduced to below 5 kOhms.

Once participants were prepared for recording EEG, they were given general instructions about the task. A pair of neutral stimuli followed by a pair of medium amplitude critical stimuli were presented to show their differences. A short session, which consisted of 10 sets of 5 pairs of stimuli, was then presented to familiarize the participants with the task. Participants were asked to press the space bar if they believed a critical signal had appeared. Feedback was given at the end of each series. Two baseline measures were recorded immediately following this session. In the first baseline, participants were asked to sit calmly for 30 s with their eyes open. In the second baseline recording, they were asked to sit quietly for 30 s with their eyes closed.

Participants then performed an 8-min practice trial. During this trial, critical stimulus size remained at the medium amplitude level. Participants' mean index of engagement was calculated during this session. An A' score of 0.7 was set as the minimum criterion to proceed to the experimental session. Participants who failed to meet the minimum score performed the 8-min practice session a second time. If they again failed to meet the criterion score, the session was terminated and the participant was dismissed. The NASA-TLX was administered at the successful completion of the practice session.

Participants who met the criterion score were given instructions for the 40-mm experimental session. The experimental session began with a 40-s premeasurement period so that sufficient data to calculate the initial engagement index could be collected. The index was recalculated continuously throughout the session using data collected 20 s prior to the calculation. Data were updated every 2 s. The critical stimulus remained at or reverted to medium amplitude when the index was within the boundaries of 0.2 standard deviations (SD) above or below the mean index. Amplitude increased or decreased by 2 mm when the index was more than 0.2 SD above or below the mean. In the negative feedback condition, amplitude increased when the index fell more than 0.2 SD below the mean, and decreased when the index exceeded the mean index by more than 0.2 SD. In the positive feedback condition, amplitude increased when the index exceeded the mean by 0.2 SD and decreased when it fell below the mean by 0.2 SD. The NASA-TLX was again administered at the end of the experimental session to assess participants' perceived level of workload.

RESULTS

Few significant effects were revealed when performance data were collapsed across amplitude. Thus, all performance analyses were broken down among CS types unless stated otherwise. Confidence levels for all analyses was set a priori at 0.05. Bonferroni t-tests were used for pairwise comparisons among the means for all repeated measures. Student Newman-Keuls procedures were used for all other post hoc tests.

Engagement Index

A 2 Feedback (positive and negative) x 2 KR (KR and No-KR) x 3 CS Amplitude (small, medium, large) x 4 Monitoring Period (10-min periods) ANOVA was performed on the engagement index. A main effect for KR was observed, $\underline{F}(1, 36) = 5.65$. The engagement index was significantly higher for the KR condition (M = 7.73, SD = 4.69) than the No-KR condition (M = 4.88, SD = 2.41).

A main effect for CS Amplitude was also observed, $\underline{F}(2, 72) = 171.81$. The mean engagement index for large, medium, and small signals (with standard deviations in parentheses) were 7.60 (4.41), 6.22 (3.87), and 5.02 (3.45), respectively. Bonferroni t-tests with an adjusted $\alpha = 0.017$ revealed that the engagement index was significantly higher for large signals than for medium signals. The index was also significantly higher for medium signals than for small signals.

A KR x CS interaction was observed. Student Newman Keuls post hoc tests revealed that the index was significantly higher for the KR group for each CS type. Bonferroni t-tests with an adjusted $\alpha = 0.0\ 17$ revealed that the index was highest when large signals were presented for both the KR and No-KR conditions. The index was also higher for medium than small signals for both feedback conditions.

A Period x CS interaction also reached significance. Bonferroni tests with an adjusted $\alpha = 0.017$ revealed that the engagement index was significantly higher for large signals than for medium and small signals across all periods. The index was also higher for medium signals than small signals across periods.

CS Amplitude Switches

A 2 Feedback x 2 KR x 4 Monitoring Period ANOVA was performed on the number of changes in CS amplitude. A main effect for Period was observed, $\underline{F}\ (3, 108) = 6.11$. The mean number of amplitude switches for periods one through four (with standard deviations in parentheses) were 77.60 (18.53), 81.40 (20.09), 76.70 (20.74), and 66.75 (26.25), respectively. Pairwise comparisons with a Bonferroni correction of $\alpha = 0.008$ revealed that significantly fewer switches occurred in period four than in period two. No other effects reached significant.

CS Occurrences

CS Occurrences were analyzed using a 2 Feedback x 2 KR x 2 CS Amplitude x 4 Monitoring Period ANOVA. A main effect for CS was found, $\underline{F}(2, 58)\ 23.40$. Mean occurrences for large, medium, and small signals (with standard deviations in parentheses) were 9.92 (5.06), 3.73 (2.22), and 7.11(4.31), respectively. Bonferroni t-tests with an adjusted $\alpha = 0.017$ revealed that large signals occurred more often than medium and small signals, with medium signals occurring least often.

A Feedback x CS Amplitude interaction was also observed, $\underline{F}(2, 58) = 8.91$. Figure 13 displays the nature of this interaction. Bonferroni t-tests with an adjusted $\alpha = 0.017$ revealed that small and large signals occurred at a significantly higher rate than medium signals in both feedback conditions. Small signals also occurred at a significantly higher rate under positive feedback, whereas large signals occurred more often under negative feedback. A significant difference was also found between small and large signals under negative feedback, with large signals occurring most often.

Proportions of Hits (PHIT)

A 2 Feedback x 2 KR x 3 CS Amplitude x 4 Monitoring Period ANOVA was used to analyze PHITs. A main effect for CS type was found, $\underline{F}(2, 58) = 36.43$. Mean PHITs for large, medium, and small signals (with standard deviations in parentheses) were 0.66 (0.28), 0.41 (0.34), and 0.40 (0.30), respectively. Bonferroni t-tests with an adjusted $\alpha = 0.017$ revealed that large signals were detected most often. A main effect for Period was also found, $\underline{F}(3, 87) = 5.83$. Mean PHITs for periods one through four (with standard deviations in parentheses) were 0.61 (0.28), 0.46 (0.31), 0.47 (0.29), and 0.43 (0.33), respectively. Pairwise comparisons with a Bonferroni correction of $\alpha = 0.008$ revealed that PHIITs declined significantly from period one to period two.

A KR x CS Amplitude interaction also reached significance, $\underline{F}(2, 58) = 6.14$ (see Figure 14). A Student Newman Keuls post hoc test revealed that significantly more large and medium signals were detected in the KR as compared to the No-KR condition.

Proportion of False Alarms (PFA)

A 2 Feedback x 2 KR x 4 Monitoring Period ANOVA was performed to analyze PFAs across CS types. A KR x Monitoring Period interaction was observed, $\underline{F}(3, 108) = 3.30$. Figure 15 illustrates the interaction. However, Bonferroni t-tests with an adjusted $\alpha = 0.008$ revealed no differences in PFA between periods for both the KR and No-KR conditions.

Perceptual Sensitivity (A')

Participants' A' scores were collapsed across CS types. Data were then analyzed using a 2 Feedback x 2 KR x 4 Monitoring Period mixed factorial ANOVA. A main effect was found for Period, $\underline{F}(3, 108) = 6.50$. Mean A' scores for periods one through four (with standard deviations in parentheses) were 0.90 (0.07), 0.85 (0.09), 0.86 (0.07), and 0.84 (0.13), respectively. Pairwise comparisons with a Bonferroni correction of $\alpha = 0.008$ revealed that a significant decrease in perceptual sensitivity occurred between periods one and two.

Response Criterion (B")

Participants' B" scores were analyzed using a 2 Feedback x 2 KR x 4 Monitoring Period mixed factorial ANOVA. No significant differences were found.

Response time to Hits (RTH)

A 2 Feedback x 2 KR x 3 CS Amplitude x 4 Monitoring Period ANOVA was used to analyze response times to hits (RTH). A main effect for KR was observed, $\underline{F}(1, 29) = 4.76$. Overall RTHs were faster for KR (M = 614.20, SD =102.04) as compared to No-KR (M =

688.00, SD = 237.14). A main effect for CS Amplitude also reached significance, $\underline{F}(2, 58) = 3.67$. However, Bonferroni t-tests with an adjusted $\alpha = 0.017$ revealed no differences in RTH for CS Amplitude. Mean RTH for small, medium, and large CSs (with standard deviations in parentheses) were 699.79 (417.02), 636.66 (147.25), and 617.44 (143.32), respectively.

NASA-TLX

Workload estimates for Session 1 differed between groups. Therefore, Session 2 scores were analyzed with a 2 Feedback x 2 KR Analysis of Covariance (ANCOVA) using Session 1 scores as the covariate. The covariate, Session 1 scores, were significant, $\underline{F}(1, 35) = 26.76$. Consequently, no significant differences were found for Feedback or KR when differences in Session 1 scores were taken into account.

SUMMARY

Research has shown that vigilance performance can be moderated by manipulating critical stimulus (CS) amplitude. For instance, Metzger, Warm, and Senter (1974) found that operator performance improved when larger signals were presented. Wiener (1973) found that he could maintain operator performance by manipulating signal amplitude. In his experiment, the manipulation was driven by a closed-loop system that used the operator's immediate prior performance as the criterion for change. Wiener provided evidence that the vigilance decrement could be eliminated by adjusting task difficulty. These findings prompted the present investigation into the effects of CS amplitude manipulation based on physiological measures.

In the present study, a physiological variation of Wiener's cybernetic system was used. The goal was to investigate the utility of a biocybernetic system to moderate vigilance performance by changing critical signal (CS) amplitude based on the operator's own EEG patterns. Knowledge of results (KR) to missed signals was also provided to aid in performance and engagement by motivating participants to pay closer attention to the task.

References

Byrne, E.A., Parasuraman, R., 1996. Psychophysiology and adaptive automation. Biol. Psychol. 42, 249-268.

Freeman, F.G., Mikulka, P.J., Scerbo, M.W., & Hadley, G. 1998. Comparison of two methods for use in adaptive automation. Perceptual and Motor Skills, 86, 1185-1186.

Hart, S.G., & Staveland, L.E. 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), Human Mental Workload (pp. 139-183). Amsterdam: North-Holland.

Metzger, K.R., Warm, J.S., & Senter, R.J. 1974. Effects of background event rate and critical signal amplitude on vigilance performance. Perceptual and Motor Skills, 38, 1175-1181.

Parasuraman, R., Mouloua, M., & Molloy, R. 1996. Effects of adaptive task allocation on monitoring of automated systems. Human Factors, 38, 665-679.

Pope, A.T., Bogart, E.H., & Bartolome, D.S. 1995. Biocybernetic system validates index of operator engagement in automated task. Biol. Psychol. 40, 187-195

Prinzel, L.J., Freeman, F.G., Scerbo, M.W., & Mikulka, P.J. 1995. A biocybernetic system for regulating operator engagement in an adaptively automated, multi-task flight environment. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, 1365-1369.

Wiener, E.L. 1973. Adaptive measurement of vigilance decrement. Ergonomics, 16, 353-363.

Wiener, E.L. 1988. Cockpit automation. In E.L. Wiener & D.C. Nagel (Eds.), Human Factors in Aviation (pp.433-461). San Diego: Academic Press.

Property on loan from NASA (which we need to renew for our current grant):

1. 1091969 MICRO COMPUTER, APPLE COMPUTER INC.
   Serial # F315MUG230
   Model # M4200

2. 1083530 DISPLAY UNIT NEC INFORMATION SYSTEMS INC.
   Serial # OZM24480
   Model # JC1404HMA

3. 1091018 MICROPROCESSOR INTERFACE BIOPAC SYSTEMS
   Serial # 91080129
   Model # MP100

4. 1091019 INTERFACE MODULE BIOPAC SYSTEMS
   Model UM100

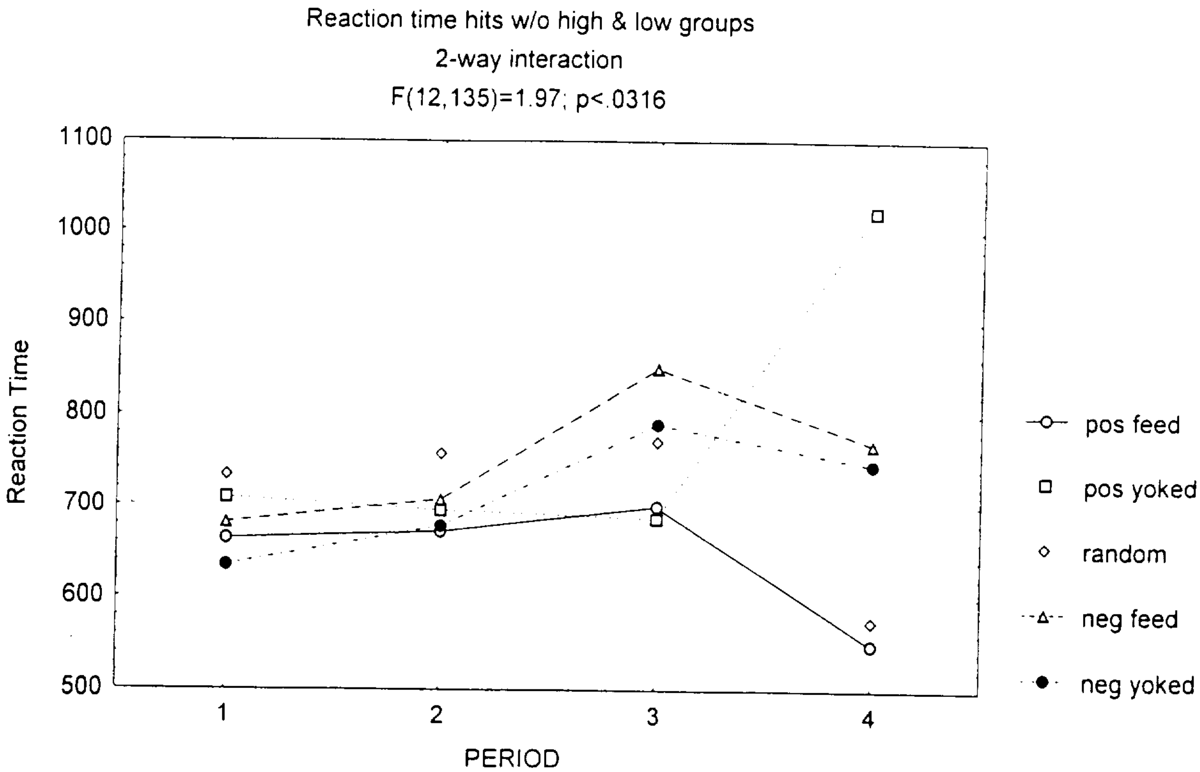

Reaction time hits w/o high & low groups
2-way interaction
F(12,135)=1.97; p<.0316
1100
1000
900
800
700
600
500
Reaction Time
1
2
3
4
PERIOD
pos feed
pos yoked
random
neg feed
neg yoked

Figure 1

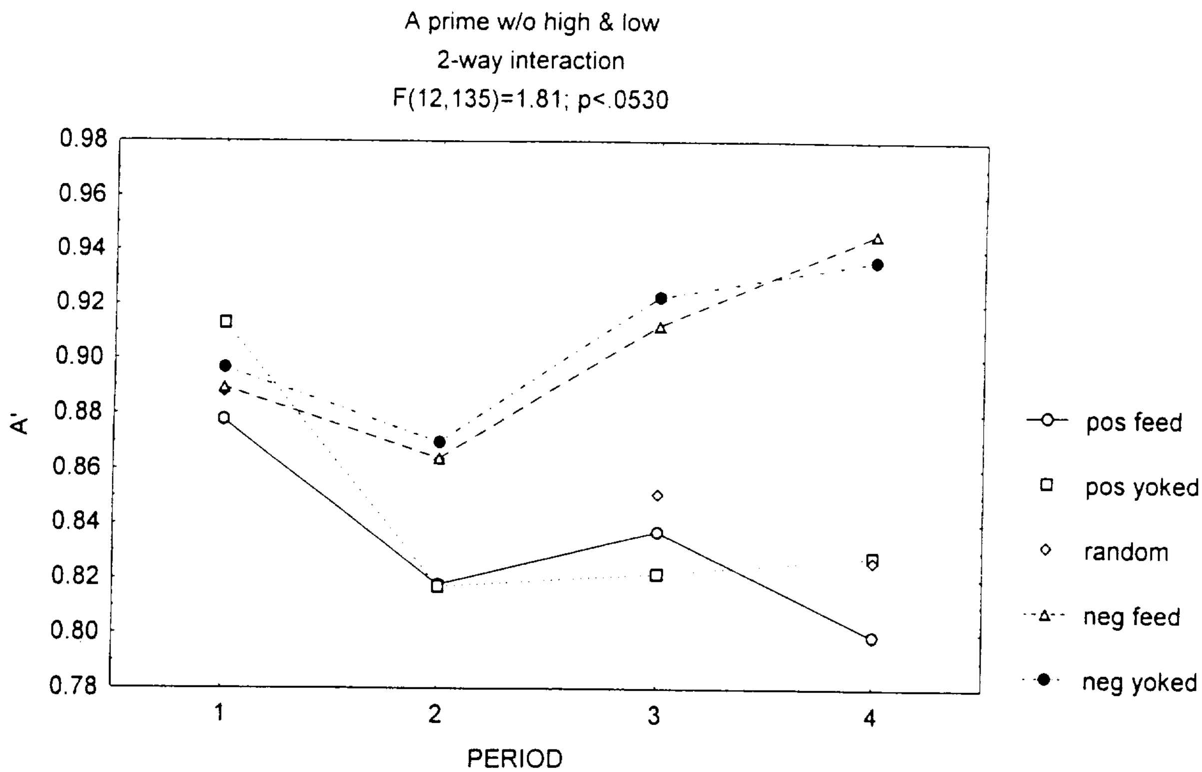

A prime w/o high & low
2-way interaction
F(12,135)=1.81; p<.0530
A'
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0.84
0.82
0.80
0.78
1
2
3
4
PERIOD
pos feed
pos yoked
random
neg feed
neg yoked

Figure 2

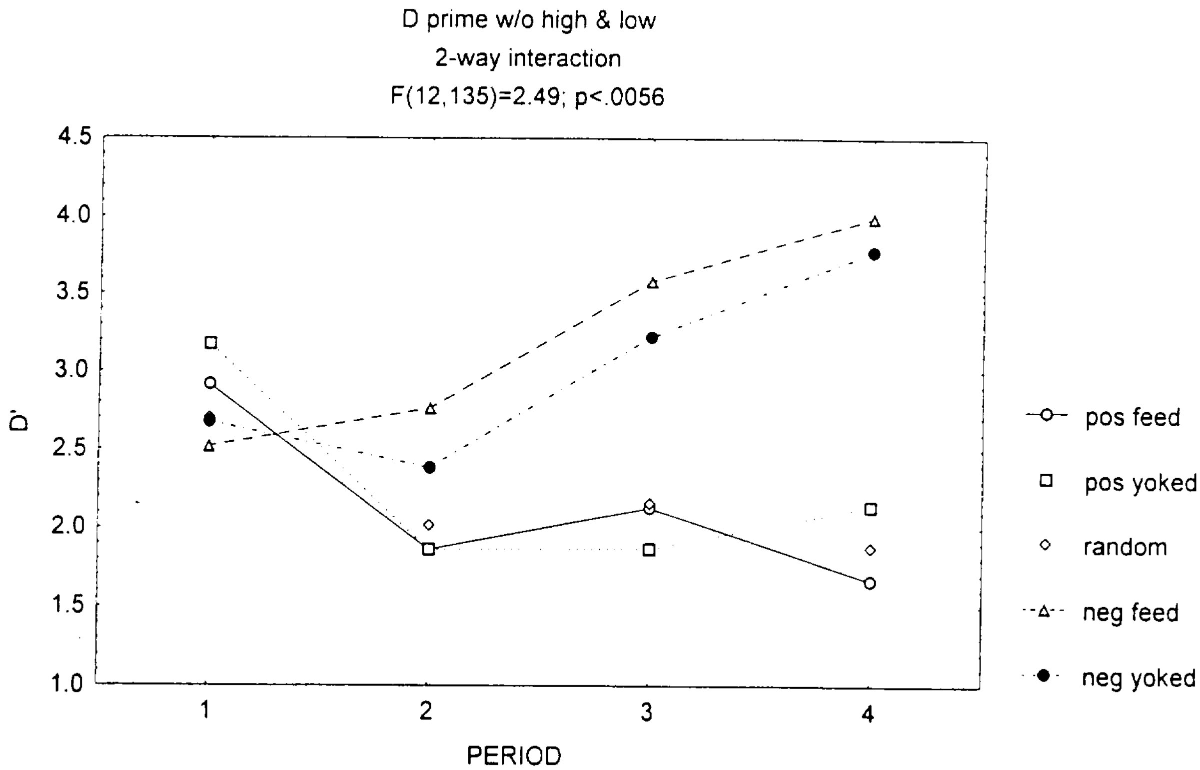

D prime w/o high & low
2-way interaction
F(12,135)=2.49; p<.0056
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
D'
1
2
3
4
PERIOD
pos feed
pos yoked
random
neg feed
neg yoked

Figure 3

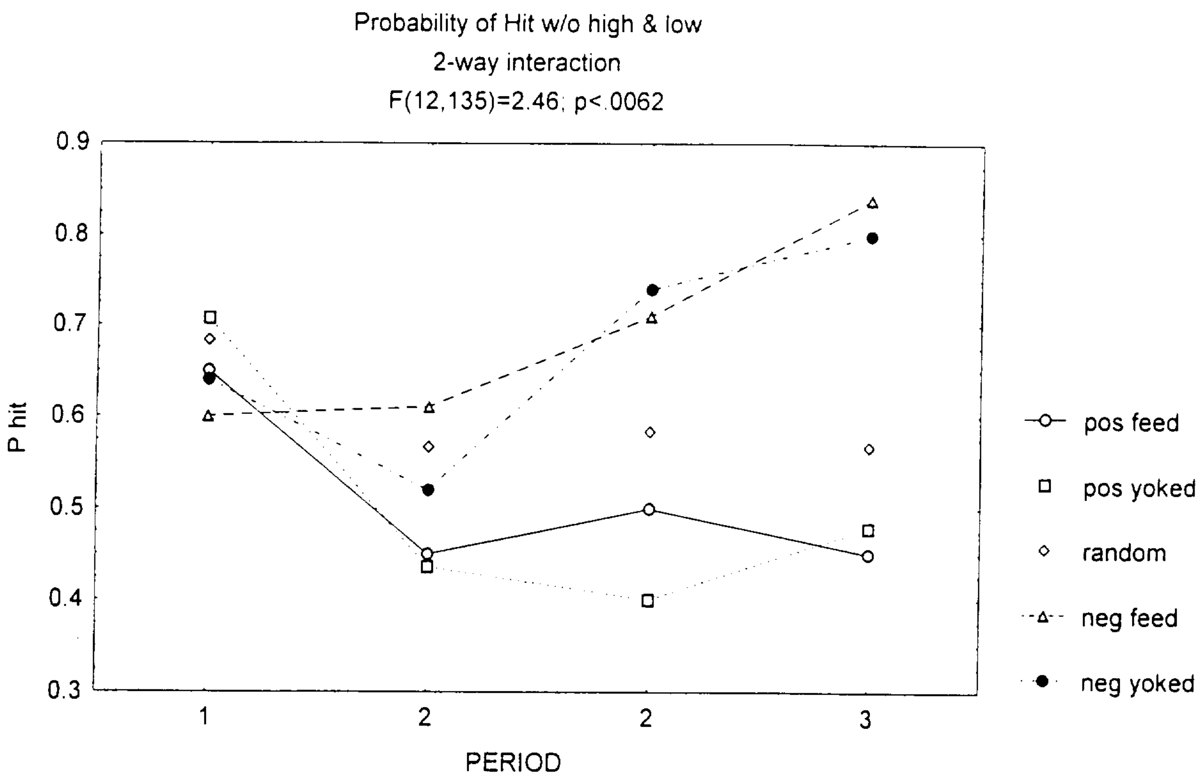

Probability of Hit w/o high & low
2-way interaction
F(12,135)=2.46; p<.0062
0.9
0.8
0.7
0.6
0.5
0.4
0.3
P hit
1
2
2
3
PERIOD
pos feed
pos yoked
random
neg feed
neg yoked

Figure 4

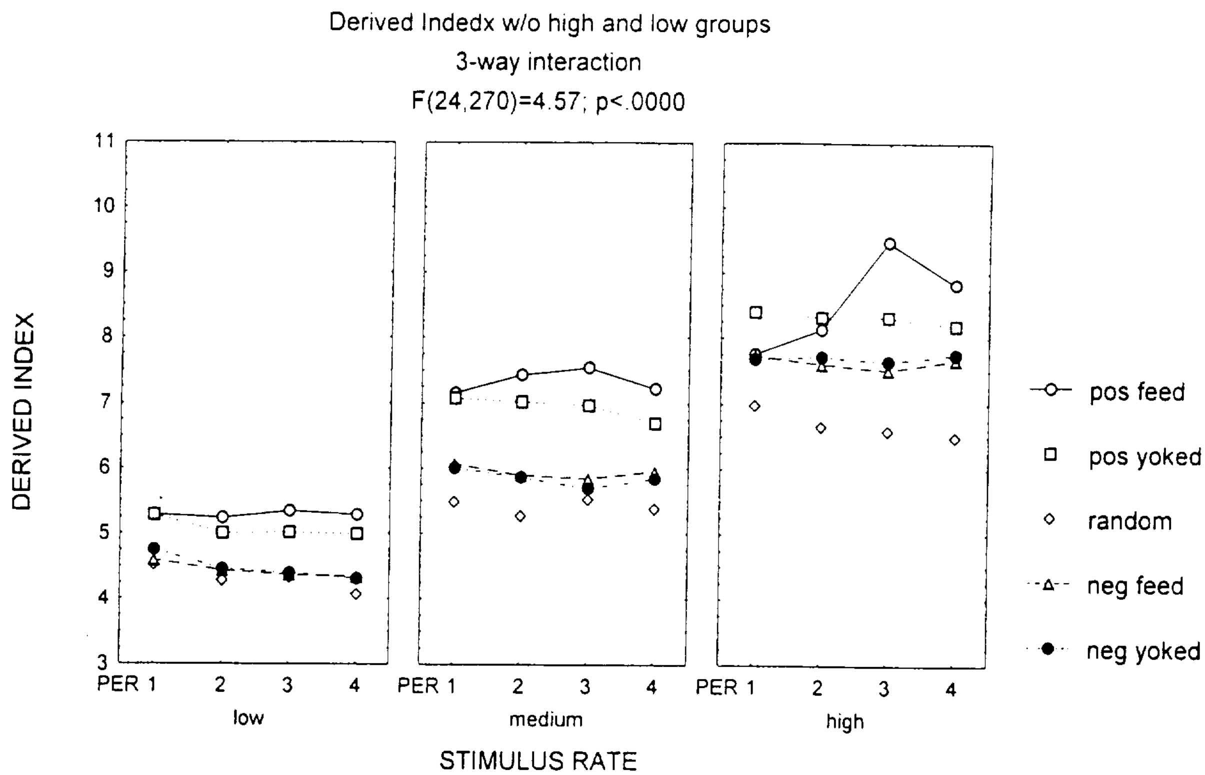

Derived Indedx w/o high and low groups
3-way interaction
F(24,270)=4.57; p<.0000
DERIVED INDEX
11
10
9
8
7
6
5
4
3
PER 1
2
3
4
low
medium
high
STIMULUS RATE
pos feed
pos yoked
random
neg feed
neg yoked

Figure 5

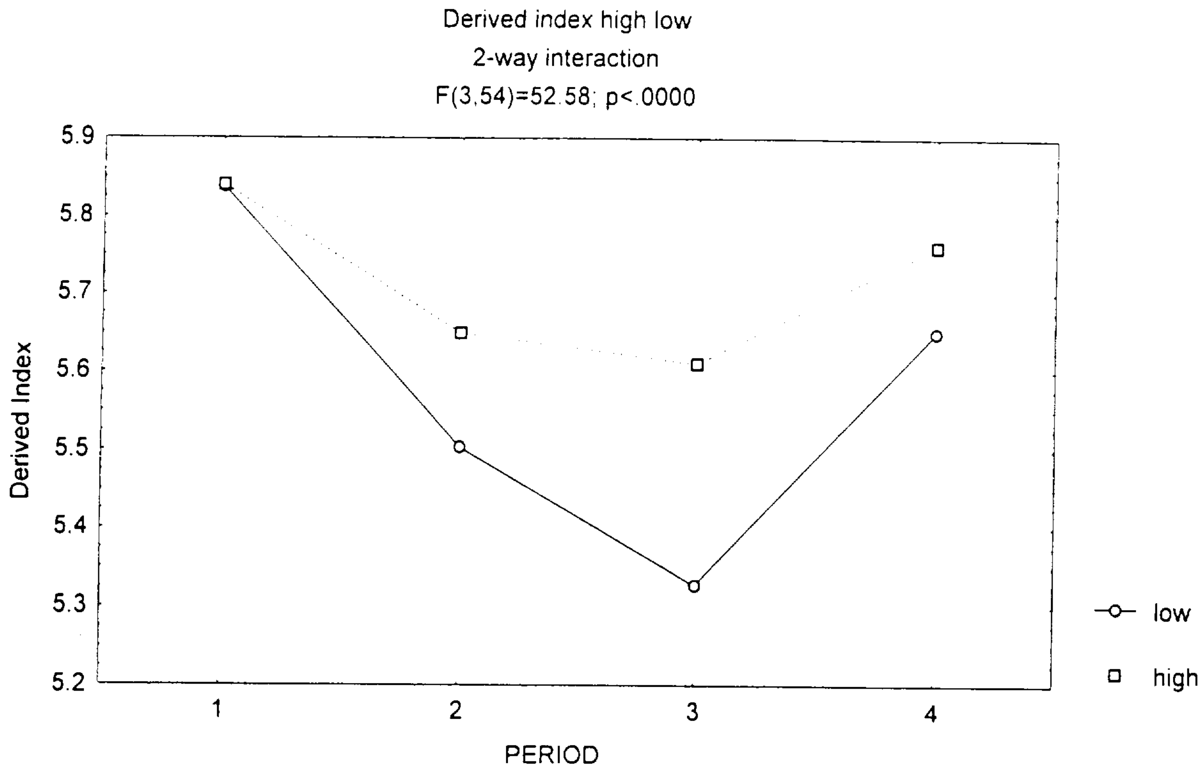
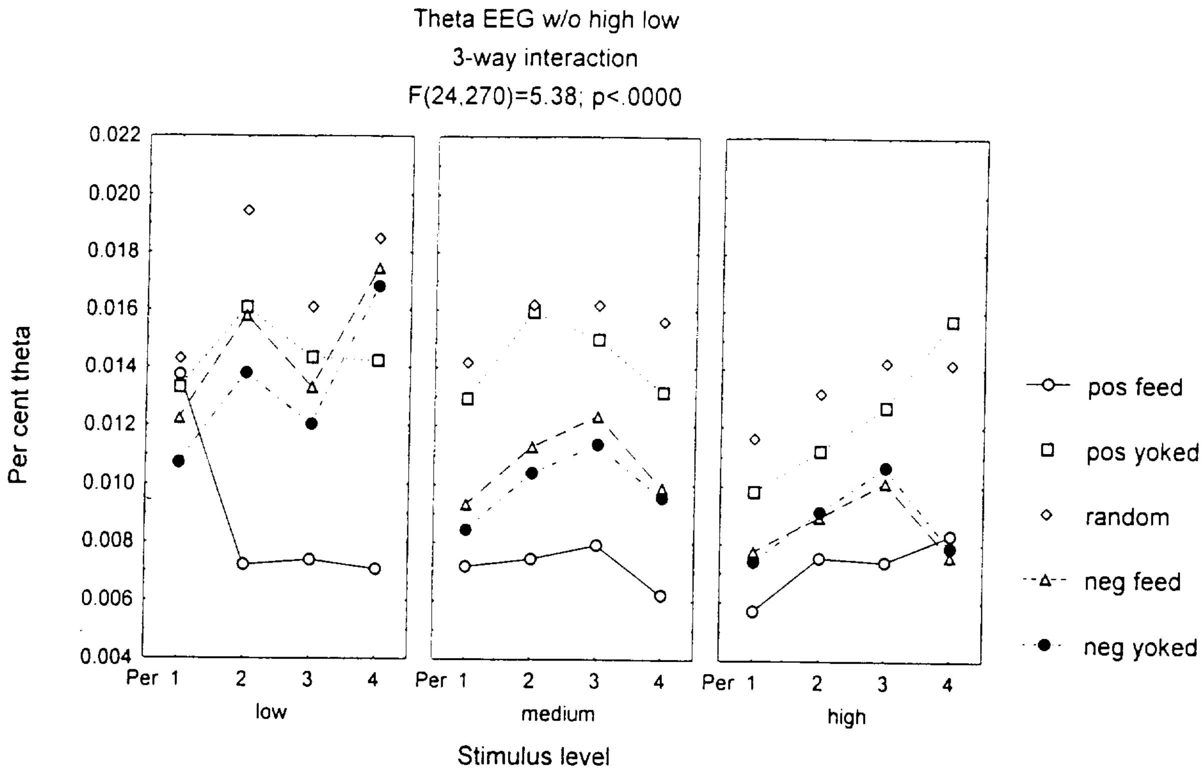

Derived index high low
2-way interaction
F(3,54)=52.58; p<.0000
5.9
5.8
5.7
5.6
5.5
5.4
5.3
5.2
Derived Index
1
2
3
4
PERIOD
low
high

Figure 6



Theta EEG w/o high low
3-way interaction
F(24,270)=5.38; p<.0000
0.022
0.020
0.018
0.016
0.014
0.012
0.010
0.008
0.006
0.004
Per cent theta
Per 1
2
3
4
low
medium
high
Stimulus level
pos feed
pos yoked
random
neg feed
neg yoked

Figure 7

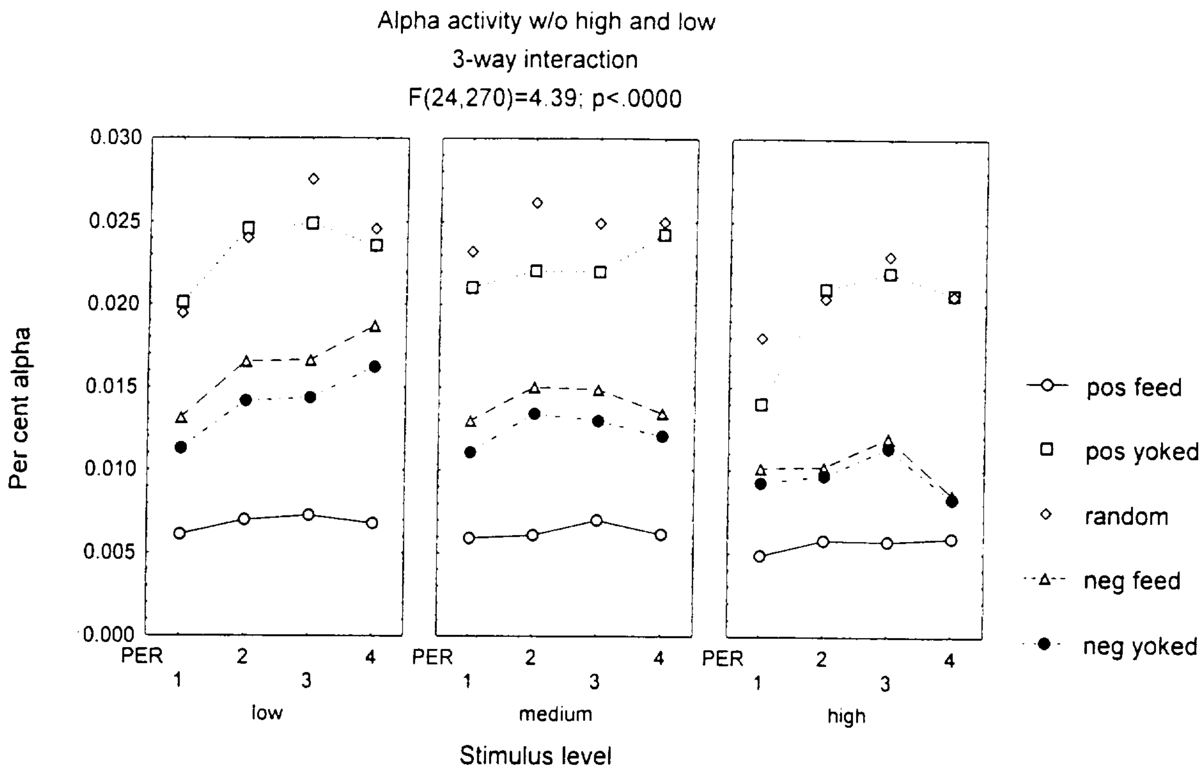
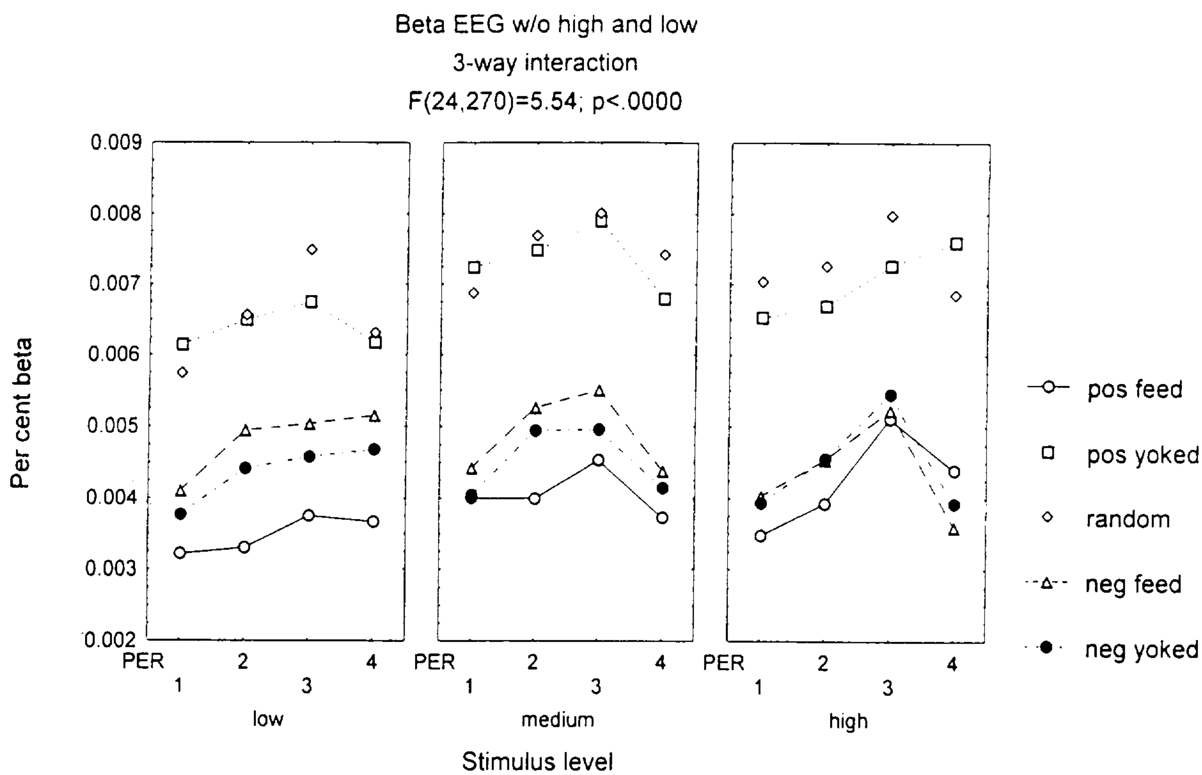

Alpha activity w/o high and low
3-way interaction
F(24,270)=4.39; p<.0000
Per cent alpha
0.030
0.025
0.020
0.015
0.010
0.005
0.000
PER
1
2
3
4
low
medium
high
Stimulus level
pos feed
pos yoked
random
neg feed
neg yoked

Figure 8



Beta EEG w/o high and low
3-way interaction
F(24,270)=5.54; p<.0000
Per cent beta
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
PER
1
2
3
4
low
medium
high
Stimulus level
pos feed
pos yoked
random
neg feed
neg yoked

Figure 9

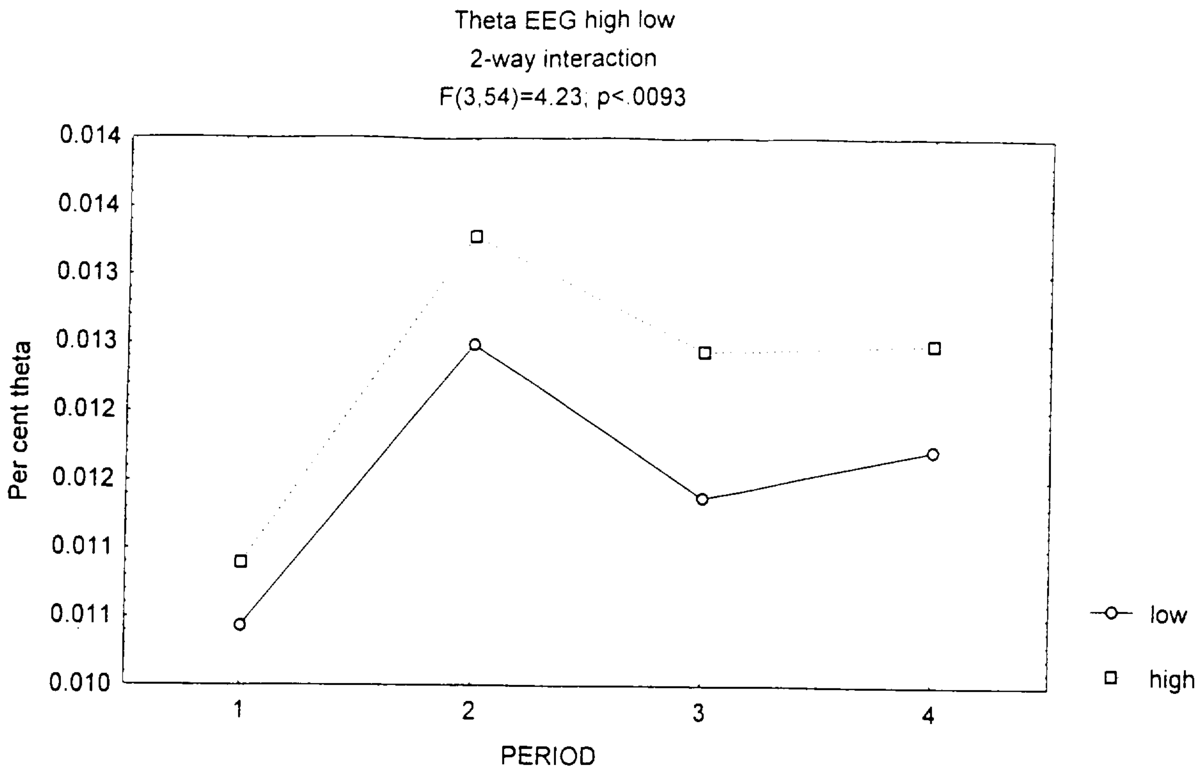

Theta EEG high low
2-way interaction
F(3,54)=4.23; p<.0093
0.014
0.014
0.013
0.013
0.012
0.012
0.011
0.011
0.010
Per cent theta
1
2
3
4
PERIOD
low
high

Figure 10

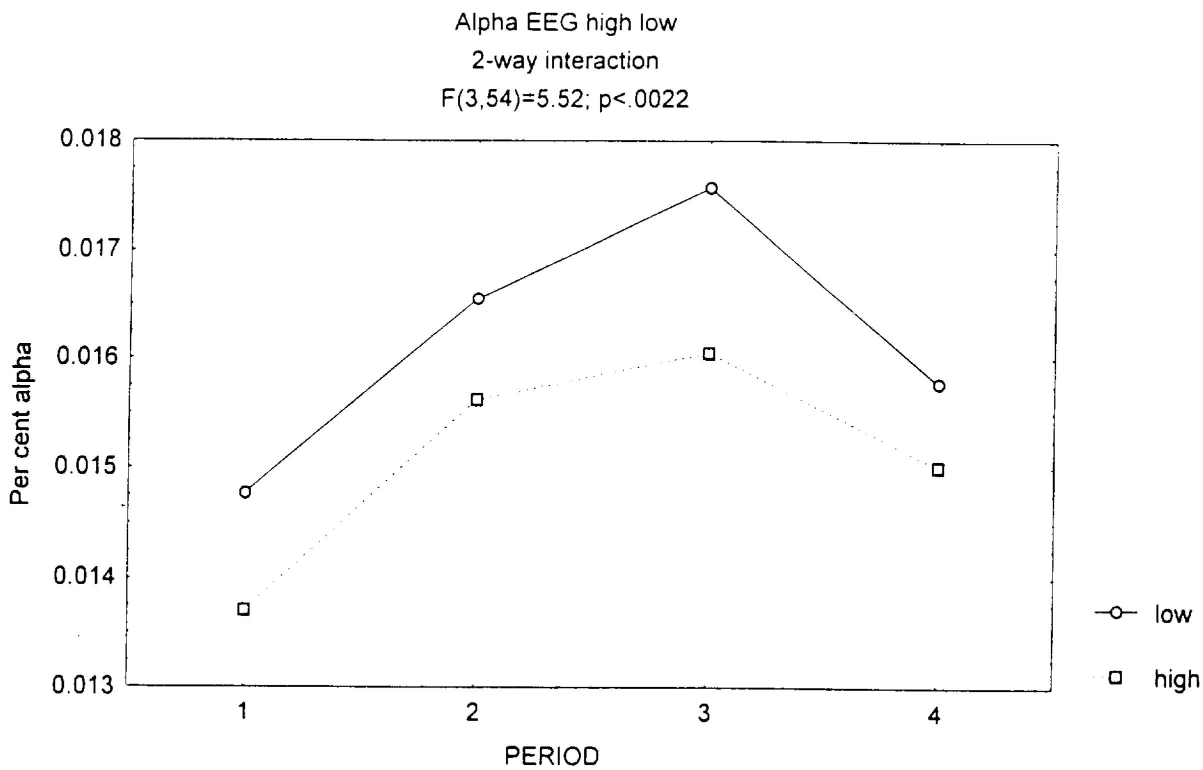

Alpha EEG high low
2-way interaction
F(3,54)=5.52; p<.0022
0.018
0.017
0.016
0.015
0.014
0.013
Per cent alpha
1
2
3
4
PERIOD
low
high

Figure 11

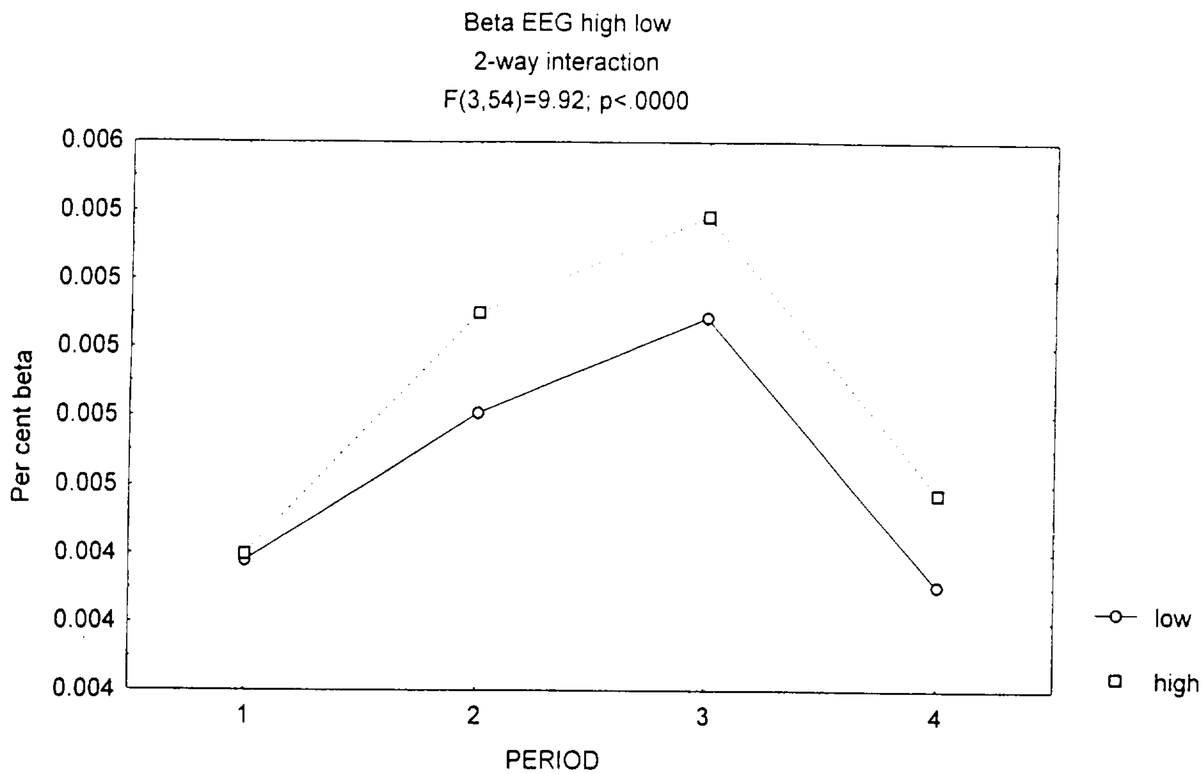

Beta EEG high low
2-way interaction
F(3,54)=9.92; p<.0000
0.006
0.005
0.005
0.005
0.005
0.005
0.004
0.004
0.004
Per cent beta
1
2
3
4
PERIOD
low
high

Figure 12

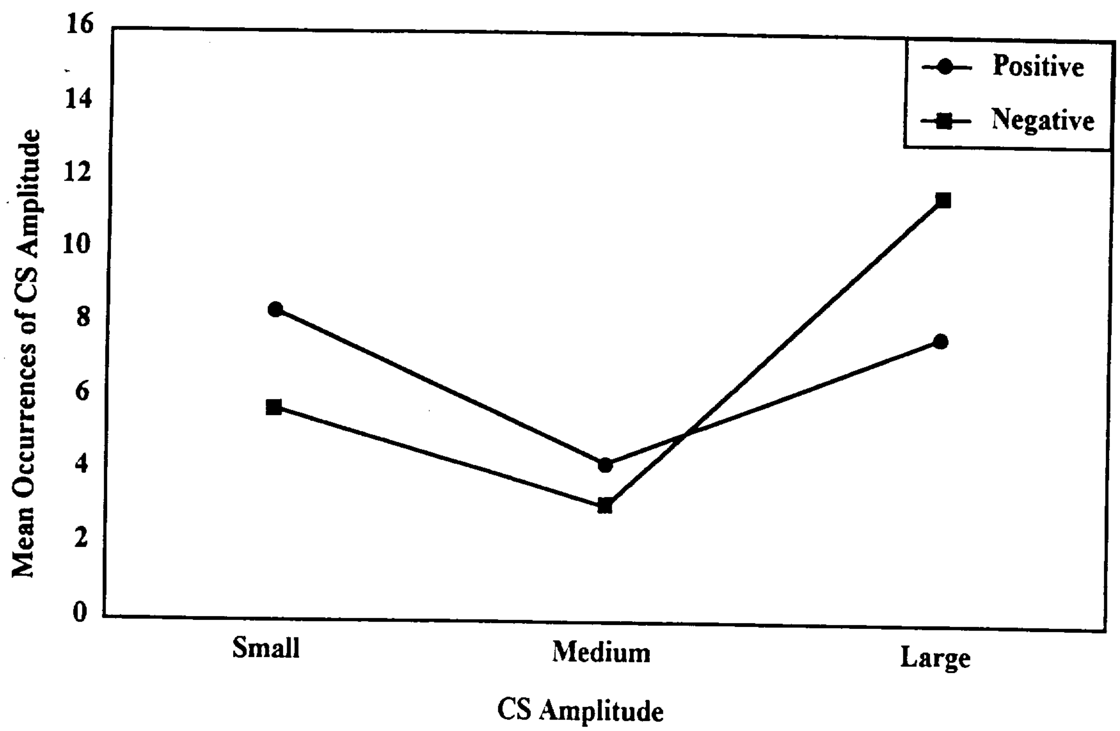


Figure13. Mean occurrences of critical signal (CS) amplitude during the experimental task under positive and negative feedback.

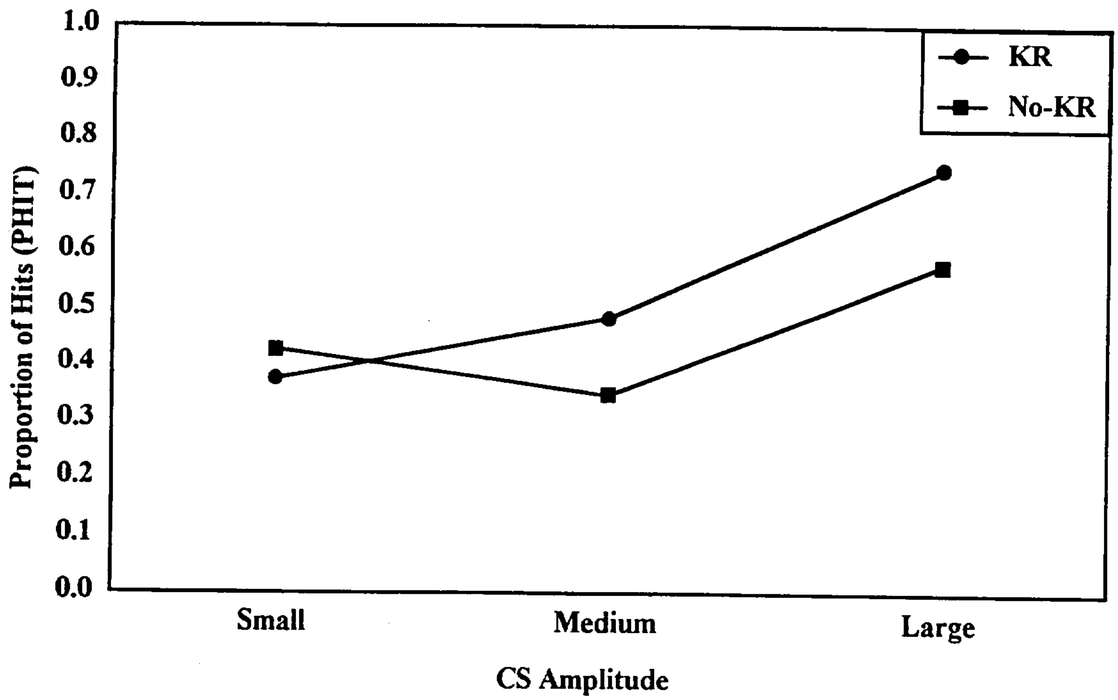


Figure 4. Overall proportion of hits (PHIT) per critical signal (CS) amplitude for the Knowledge of Results (KR) and No-KR conditions.

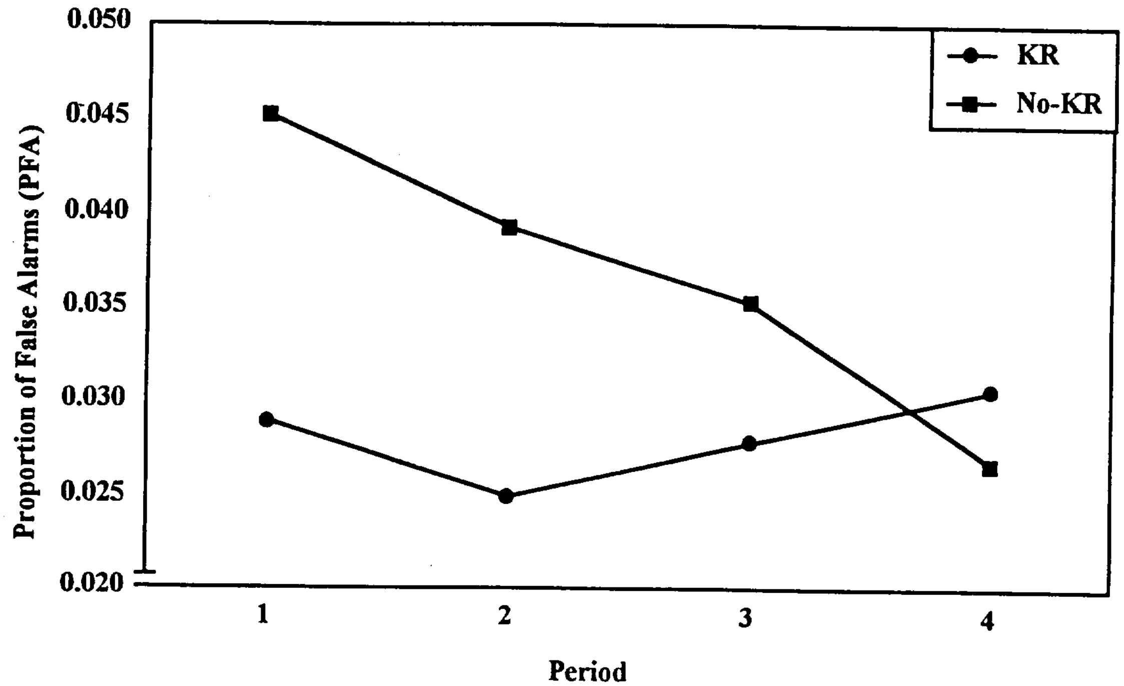


Figure 5. Proportion of false alarms (PFA) per monitoring period for the Knowledge of Results (KR) and No-KR conditions.