November 2, 2001

Final Report for: **NASA Agreement No. NCC-1-364**

"Psychophysiological Control of Acognitive Task Using Adaptive Automation"

Submitted by: Dr. Frederick Freeman
Principal Investigator

!
●

Introduction

The major focus of the present proposal was to examine psychophysiological variables 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 worker performance (Wiener, 1988). Byrne and Parasuraman (1996) have argued for the use of psychophysiological measures in the development and the implementation of adaptive automation. While both performance based and model based adaptive automation have been studied (e.g. Parasuraman, Mouloua, & Molloy, 1996), the use of psychophysiological measures, especially EEG, offers the advantage of real time evaluation of the state of the subject.

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 cognitive vigilance task. Previous research in our laboratory, supported by NASA, has demonstrated that, in an adaptive automation, closed-loop environment, subjects perform a tracking task better under a negative than a positive, feedback condition. In addition, this

condition produces less subjective workload and larger P300 event related potentials to auditory stimuli presented in a concurrent oddball task. We have also recently shown that the closed-loop system used to control the level of automation in a tracking task can also be used to control the event rate of stimuli in a vigilance monitoring task. By changing the event rate based on the subject's index of arousal, we have been able to produce improved monitoring, relative to various control groups.

We have demonstrated in our initial closed-loop experiments with the the vigilance paradigm that using a negative feedback contingency (i.e. increasing event rates when the EEG index is low and decreasing event rates when the EEG index is high) results in a marked decrease of the vigilance decrement over a 40 minute session. This effect is in direct contrast to performance of a positive feedback group, as well as a number of other control groups which demonstrated the typical vigilance decrement. Interestingly, however, the negative feedback group performed at virtually the same level as a yoked control group. The yoked control group received the same order of changes in event rate that were generated by the negative feedback subjects using the closed-loop system. Thus it would appear to be possible to optimize vigilance performance by controlling the stimuli which subjects are asked to process.

EXPERIMENT 1:

The experiment conducted had subjects perform a vigilance monitoring task. The task was cognitive in nature and adopted the visual search paradigm employed by Treisman (1988). In this paradigm subjects were presented with displays containing multidimensional objects, consisting of colored letters. Stimulus trials were always presented at the rate of once every five seconds (i.e. 12 trials per minute) with either two, five or nine colored letters being shown on a given trial. The target stimulus was always a green "K" which was presented on one trial per minute at random times within that minute. The stimuli consisted of either an "N", "K", "R", or "X" presented at random positions on the computer screen with the only stipulation being that the stimuli did not overlap each other. The color of the stimuli was either green, yellow-green or blue-green, except for the "K" which was either yellow-green or blue-green on non-target trials. Under the negative feedback condition, when a subject's EEG index was at least 0.2 standard deviations above baseline they were presented with a simple feature search condition in which only two stimuli were presented on the computer screen. The subject was instructed to respond, by pressing a button only when the target stimulus appeared. If the subject's index was within +/- 0.2 standard deviations of the baseline five different stimuli were presented on the computer screen. If the subject's index was more than 0.2 standard deviations below their baseline,

9 stimuli were displayed. Research by Treisman and her associates (Treisman, 1988; Treisman & Gelade, 1980) has demonstrated that conditions which require a "conjunctive search" are more difficult than simple feature searches because they require processing that involves multidimensional percepts. The more stimuli presented, the more processing that is required and the longer the reaction times. Thus, under negative feedback, requiring subjects to do more cognitive processing when their index is low should produce greater task involvement and a subsequent increase in the engagement index. Under positive feedback, the reverse contingencies will be in place and should result in poorer task performance and a reverse relationship between stimulus conditions and the engagement index.

EEG was recorded continuously from four sites (F3, F4, O1 and O2, using the international 10-20 system) during the 40 minute task and fed into a Lab-View virtual instrument which calculated the $\beta/(\alpha+\Theta)$ index. These sites were chosen because 1. the frontal lobes have been theorized to be involved in attentional processing (cf. Lubar, 1991; Posner & Raichle, 1994) and 2. the occipital lobes generating high levels of theta during lapses in attention (e.g. Beatty & O'Hanlon, 1979).

Subjects were brought into the laboratory and explained the nature of the experiment. They were then asked sign an informed consent, indicating that they understood the purpose and procedure of the study. The electrode cap was then placed on the

subject's head and a reference electrode placed on the left

mastoid. All electrode impedances were lowered to below 5kohms.

The subjects were then placed in front of the task computer and

given six trials that demonstrated the nature of the stimuli that

would be used in the task. If requested, the subject could

receive six more trials to insure they understood what they would

be looking for. They were then given an 8-minute practice trial

in which the three different stimulus arrays were randomly

presented, once every five seconds, with a target trial occurring

once per minute. At the end of the eight minute practice the

subject's A' score was determined. If A' was at least .8, the

regular 40-minute session was begun. If the A' score was below

.8, the subject was given a second eight minute practice session.

If their score was still below .8, they were not used in the

experiment. All subjects that completed the experiment were

either given two hours of extra credit for their class or $20.00,

whichever they preferred. Subjects that did not get past the

practice were given either one hour extra credit or $5.00.

Results

    First, A' scores were calculated to assess performance. In

Table 1 is presented the ANOVA for the A' scores. The only

significant effect was for periods, $F(3,87) = 16.6$, $p < .001$.

The scores are plotted in Figure 1. It can be seen that A'

improved for both positive and negative feedback groups from the

first 10-minute block to the second 10-minute block. While the scores for the negative feedback group continued to improve slightly over the remainder of the session, those of the positive feedback group declined slightly. In Table 2 is presented the ANOVA for B" scores. Again, the only significant effect was for periods, F (3,87) = 17.1, p < .01. As seen in Figure 2, subjects became more conservative as the session progressed, regardless of which group they were in.

To account for the lack of group differences observed in the A' and B" scores, we examined the false alarm and probability of hit data. In Table 3 is presented the ANOVA for false alarms. The only significant effect was for periods, F (3, 87) = 75.3, p < .01. As can be seen in Figure 3, the false alarm rate was relatively high for the first 10-minute period, but then dropped to close to zero for the last three 10-minute periods. Thus, the A' and B" data can be accounted for by the fact that the subjects appeared to still be learning the task for the first 10-minute period. Since false alarms dropped to near zero after that time, it can be argued that the only true assessment of performance is the probability of a hit. A' scores are not appropriate if subjects' false alarm rate is virtually zero.

In Table 4 is presented the ANOVA for probability of a hit for the negative and positive feedback groups. A significant group effect was found F (1,29) = 4.21, p < .05. In Figure 4 are presented the data for probability of a hit. It can be seen that

the negative feedback group improved slightly from the first to the last 10-minute period while the positive feedback group's performance decreased dramatically over the session. Curiously, the interaction between groups and periods was not significant, $F(3, 87) = 1.88$, $p > .10$. Apparently, the variability across groups was large enough to washout any effect.

The analysis of the EEG index data is presented in Table 5. The critical effect that needs to be examined is the group by array size interaction. In all of our previous studies, including tracking and perceptual vigilance, there was a significant interaction between feedback and either the manual/automatic mode (for tracking) or the stimulus rate (for perceptual vigilance). This interaction is a basic validation of the system's functioning. When subjects in the negative feedback condition become underaroused, the system should switch them to the condition that increases arousal. Conversely, when they are overaroused, the system should switch them to the opposite condition. For the positive feedback group, if subjects are highly aroused the system should keep them in that condition, and vice versa. The group by array size interaction was found to be significant, $F(2,40) = 97.7$, $p < .001$. As seen in Figure 5, the typical crossover effect that we have found in all of our earlier experiments was also produced by the cognitive vigilance task. In Figure 6 is presented this interaction across periods ($F(6, 120) = 7.92$, $p < .01$.

Finally, we compared the amount of time each group spent at each array size. We had originally reasoned that the smallest array size (i.e. two stimuli) would be the easiest to process and thus would not cause any decrement. Conversely, we expected that when subjects tried to process the large array (i.e. nine stimuli) their performance would deteriorate over the 40 minute session due to greater fatigue induced by having to constantly search for the target. Since the positive feedback group did significantly poorer than the negative feedback group, we assumed the positive feedback group probably spent more time (as reflected by number of non-target presentations) with the larger array size. The results of the ANOVA for number of non-target presentations is presented in Table 6. As expected, there was a significant group effect (F (1,20) = 4.66 p < .05) and a significant array size effect (F (2, 40) = 26.4, p < .001). There was also a significant period effect (F (3, 60) = 8.2, p < .001). There was only one significant interaction, the group by array size (F (2, 40) = 22.7, p < .001). This interaction is presented in Figure 7. As can be seen in the figure, the results were the exact opposite of what we expected. The negative feedback group, which had performed significantly better, as measured by the probability of a hit, spent the majority of the time with the large array, while the positive feedback group spent the majority of the time with the small array. Neither group spent much time at the middle array size.

Discussion.

The results of the first experiment have demonstrated that negative feedback in the closed-loop psychophysiological system can improve performance on a cognitive vigilance task just as it can for a motor task (manual tracking) and perceptual vigilance task (line length). Further, changes in the EEG index, which validate system functioning, are comparable across all three tasks. Surprisingly, the stimulus condition which produced the greatest vigilance decrement was the two-stimulus array. It would seem that when requiring subjects to continuously engage in a visual search task, presenting subjects with more stimuli to search, rather than inducing fatigue, kept subjects sufficiently aroused to maintain a high level of performance. Apparently, a two-stimulus array, while easy to search for a target, is not sufficiently stimulating, resulting in a vigilance performance decrement. Since this type of cognitive vigilance paradigm has never been used before, we felt it was necessary to confirm that, in a non-closed loop design, a vigilance decrement would indeed be seen with a two, but not a nine stimulus array, cognitive vigilance task. Experiement two tested this hypothesis.

Experiment 2

In Experiment 2, two groups were run: group one was always presented the two-stimulus array while group two was always

presented the nine stimulus array. All other stimulus parameters, intervals, electrode sites, EEG index, etc., were the same as in Experiment 1.

Results

In Table 8 is presented the ANOVA for the A' scores. There was a significant effect for both periods (F (3, 87) = 20.1, $p <$ .001) and for the group by period interaction (F (3, 87) = 3.79, $p <$ .02). As was seen for Experiment 1, A' scores, presented in Figure 9, were lowest in the first 10-minute period for both groups. However, the nine-stimulus array group maintained its high performance level for the remainder of the session while the two stimulus array group's performance began to decrease.

The poorer performance for the first 10-minute period can again be attributed to the high false alarm rate for that time. In Table 9 is presented the ANOVA for the false alarms. The only significant effect is for periods (F (3, 87) = 40.2, $p <$ .001). As seen in Figure 10, false alarms were high for the first 10 minutes then decreased to approximately zero for the remainder of the task. Because of the low false alarm rate for the majority of the task, we again felt that the probability of a hit best reflected subject performance.

In Table 10 is presented the ANOVA for probability of a hit for the two groups. The only significant effect was the interaction between groups and period (F (3, 87) = 3.67, $p <$ .02). As seen in Figure 11, for the nine-stimulus array group,

performance was slightly lower than that of the two-stimulus

array group for the first period. However, the nine-stimulus

group improved its performance slightly over the experimental

session while the two-stimulus group showed a vigilance

decrement.

Finally, the EEG index was analyzed. Despite the strong

differences observed in Experiment 1, there were no significant

differences involving either group or period for Experiment 2.

All Fs were below 1.0.

Discussion

The results of Experiment 2 confirmed the hypothesis that

the vigilance decrement seen in the positive feedback group of

Experiment 1 may be attributed to the fact that this group spent

approximately two-thirds of the experimental session observing

the two-stimulus array. That is, the positive feedback subjects

became under aroused by the task, and because of their task

contingencies (i.e. continue to show the two-stimulus array if

the EEG index is low) caused them to remain in a lowered

arousal/EEG index condition. Because the negative feedback

subjects were required to engage in more cognitive functioning

(i.e. constantly search the more complex array for the target)

their arousal level/EEG index remained higher and their

performance did not deteriorate.

It should be noted that the reason the subjects typically

had such a high false alarm rate for the first 10 minutes of the

task may be due to setting the cutoff for being in the experiment at A'=> .8. If Figures 1 and 9 are examined, it will be noted that even thought the false alarm rates were high for the first 10 minutes, A' scores were still no lower than .9. Thus, subjects, on average, were performing well above the cutoff, even for the first period.

Finally, the value of the EEG index needs to be addressed. Although the closed-loop adaptive automation system drove the EEG index as expected, without any feedback, the index remained relatively stable. In other words, it did not appear to be the nature of the stimuli or the amount of cognitive processing that is important, but rather, something about the contingencies of the closed-loop system. Since the task, regardless of the size of the stimulus array, would seem to be boring in nature, one might expect low indexes. Only with the closed-loop were we able to drive the index up or down. We have never examined how the index would be affected by a truly arousing task. We have done some pilot work using the MAT battery where subjects were required to perform three tasks at one time. Although subjects reported higher workloads compared to just performing one task, the EEG indexes still did not reflect any differences in conditions. However, even in that pilot study, it is not clear whether the high workload was due to the boring nature of the task. It would be of interest to look at how the index might be affected by a truly highly stressful task.

References

Beatty, J. & O'Hanlon, J.F. (1979) Operant control of posterior theta rhythms and vigilance performance: Repeated treatments and transfer of training. In Progress in Operant Conditioning of Physiological Events. Birbaumer, N. & Kimmel, H.D. (Eds.) Erlbaum, Hillsdale, NJ, (pp. 235-246).

Byrne, E.A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. Biological Psychology, 42,249-268.

Lubar, J.(1991). Discourse on the Development of EEG Diagnostics and Biofeedback for Attention-Deficit/Hyperactivity Disorders. Biofeedback and Self-Regulation, 16, 201-225

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

Posner, M.I. & Raichle, M.E. (1994). Images of the Mind. New York: Scientific American Library

Treisman, A. (1988). Features and Objects: The Fourteenth Bartlett Memorial Lecture. The Quarterly Journal of Experimental Psychology, 40A, 201-237.

Treisman, A. & Gelade, G.(1980). A feature integration theory of attention. Cognitive Psychology, 12, 97-136.

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.

FIGURE 1.

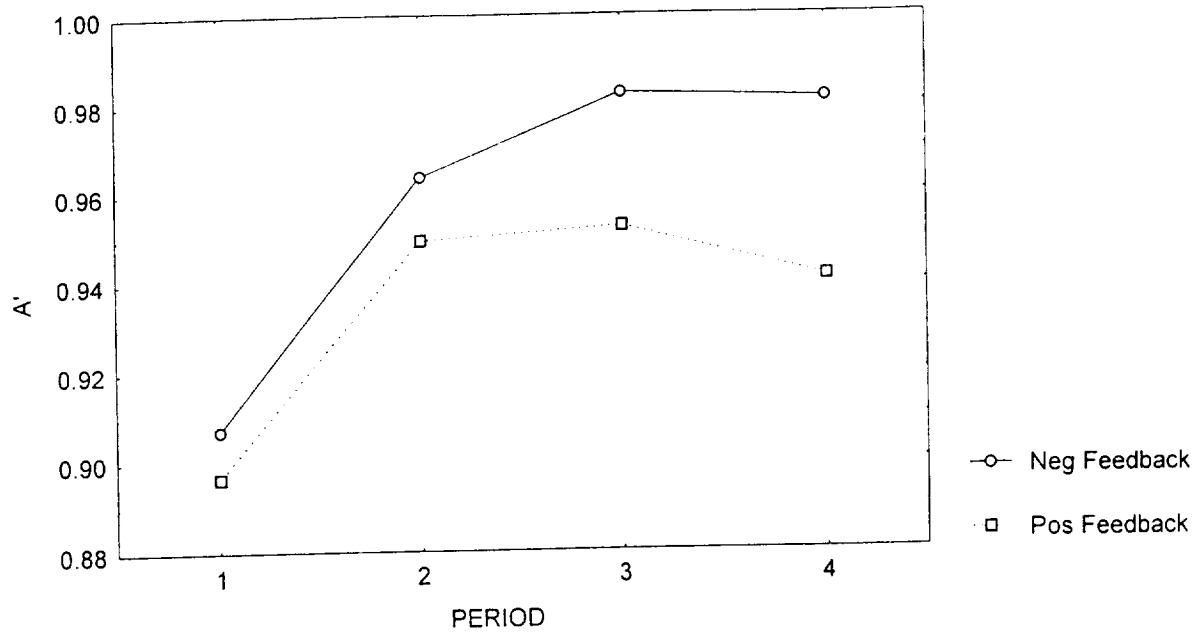A prime Treisman Closed Loop
2-way interaction
$F(3,87)=.85; p<.4697$

FIGURE 2.

B" Treisman Closed Loop
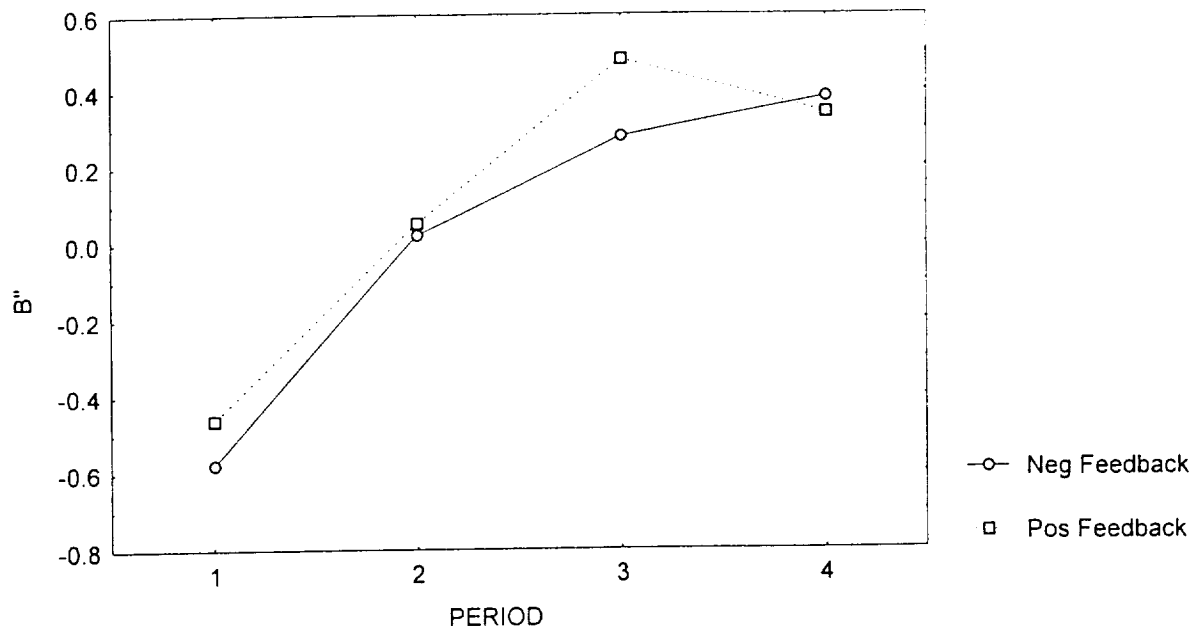2-way interaction
$F(3,87)=.27; p<.8454$

Neg Feedback

Pos Feedback

FIGURE 3.   Probability of a False Alarm
2-way interaction
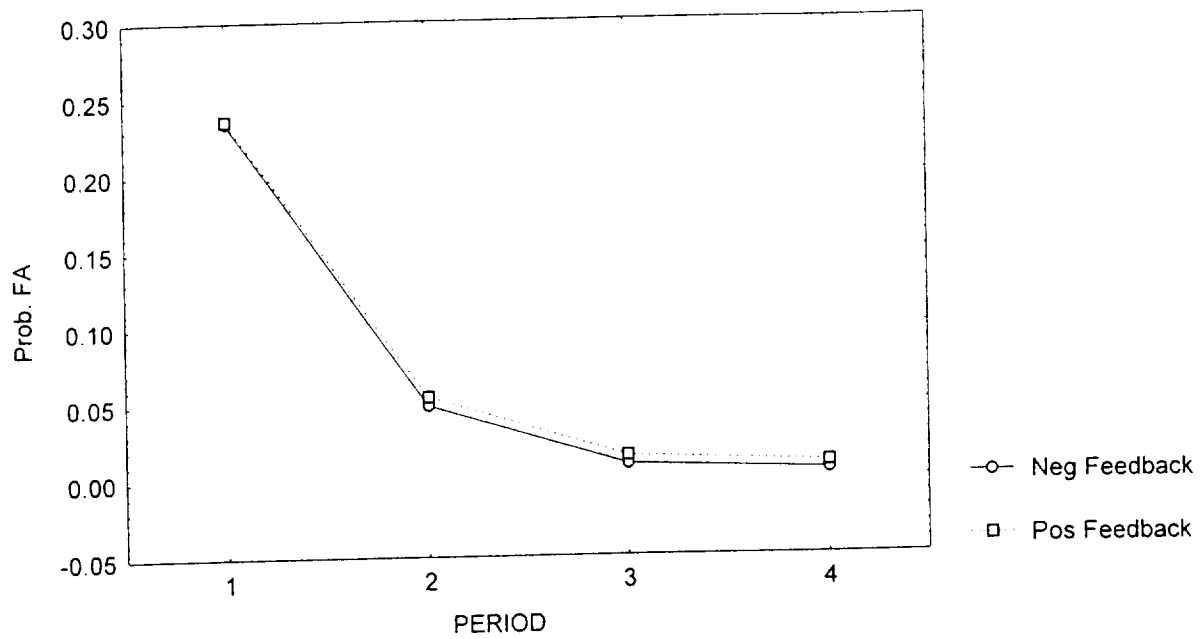$F(3,87)=.01$; $p<.9992$

FIGURE 4.   Probability of a Hit--Treisman Closed Loop
2-way interaction
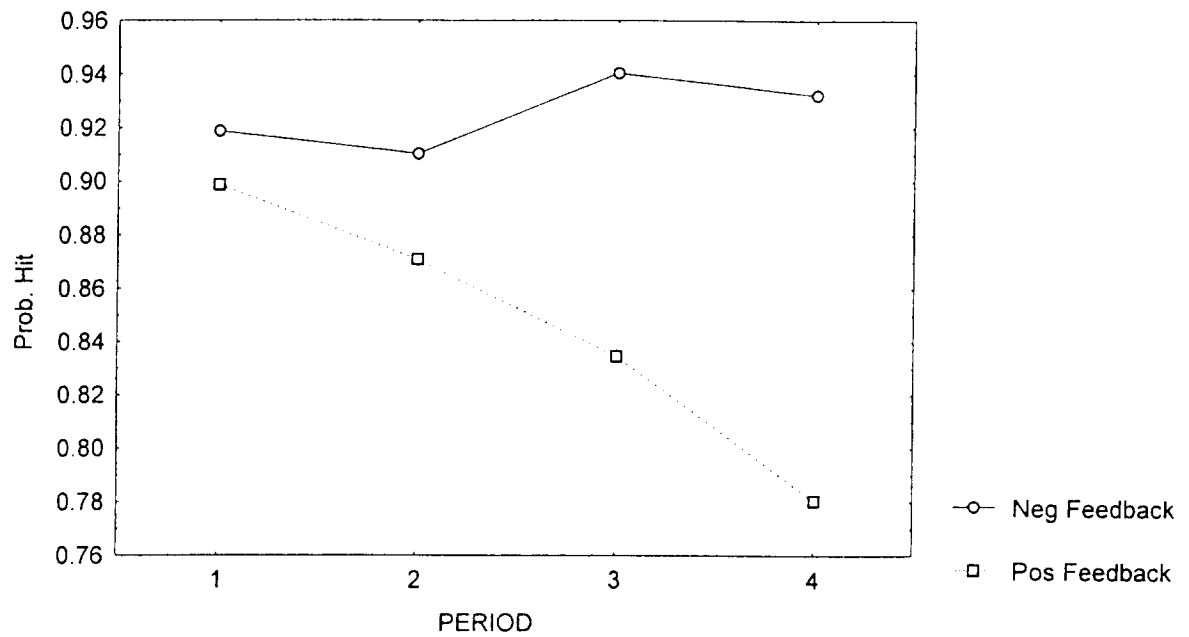$F(3,87)=1.88$; $p<.1383$

FIGURE 5.

EEG index Group x Array size
2-way interaction
F(2,40)=97.68; p<.0000
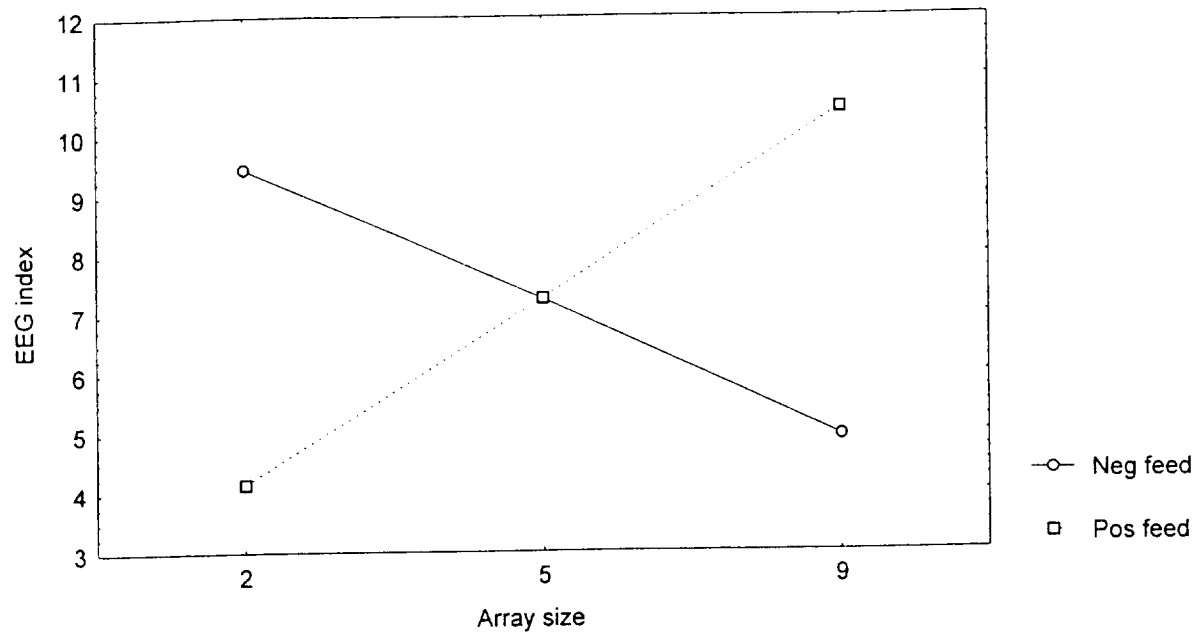
FIGURE 6.   EEG index as a function of array and period
3-way interaction
$F(6,120)=7.93; p<.0000$
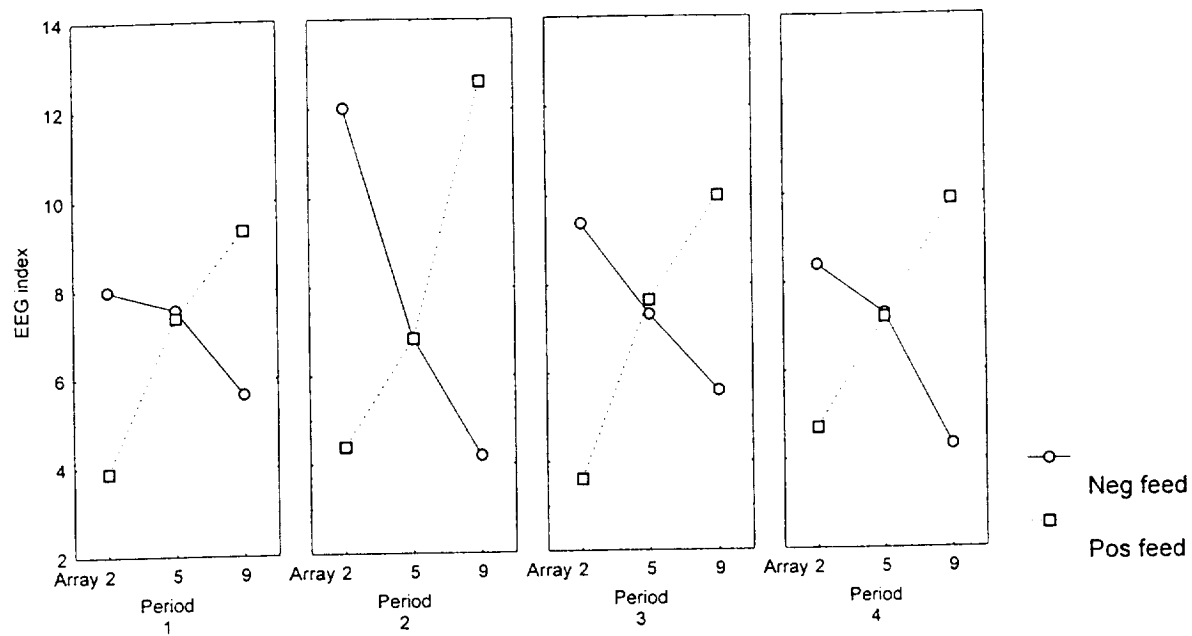
FIGURE 7.   Frequency of Non-targets
                    2-way interaction
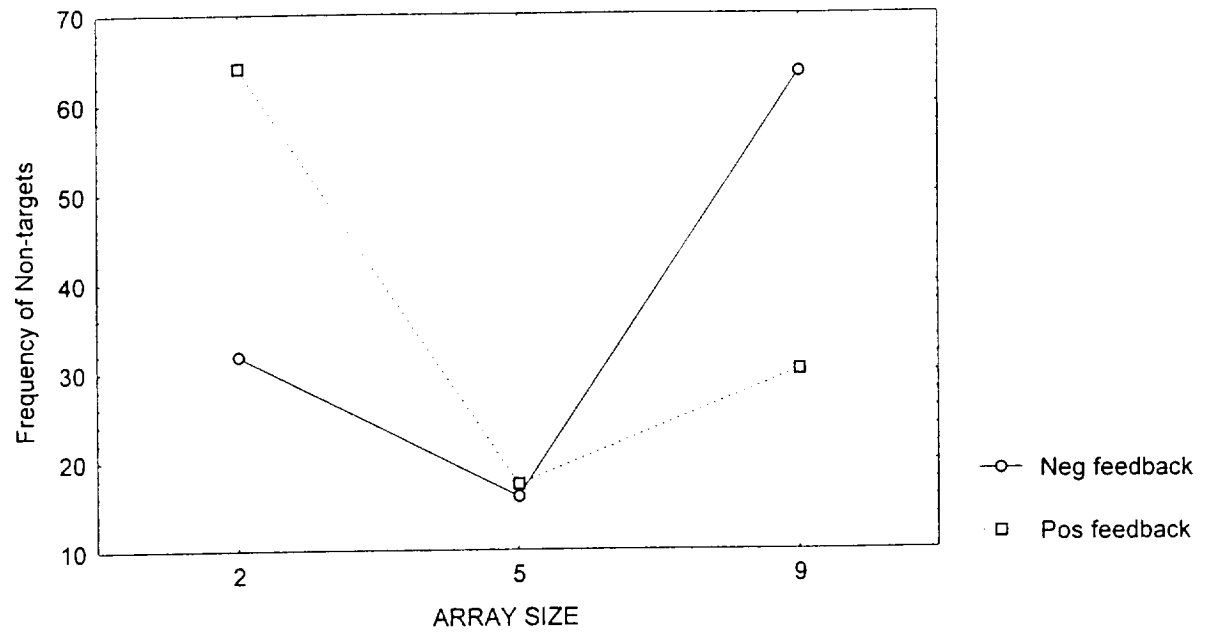   F(2,40)=22.66; p<.0000

FIGURE 8.   A' Treisman Constant 9 vs 2 Array
2-way interaction
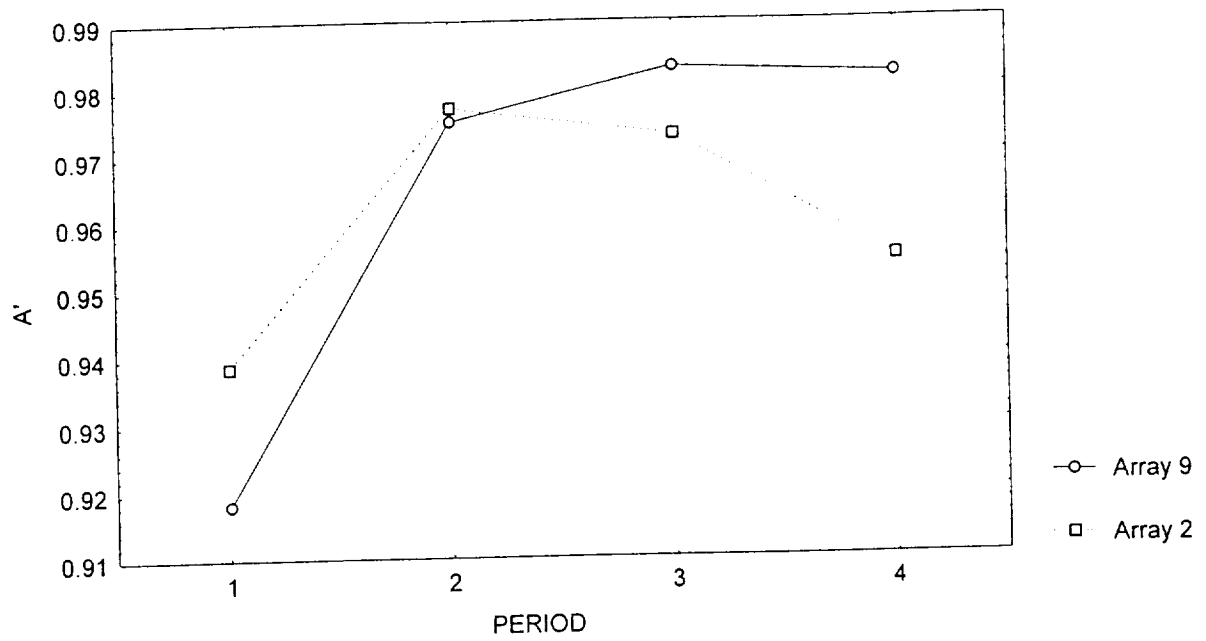$F(3,87)=3.79; p<.0132$

FIGURE 9.   Probability of FA: Treisman Constant Large Small Array
2-way interaction
$F(3,87)=.62; p<.6039$

FIGURE 10.   Probability of Hit: Treisman Constant 9 vs 2 array
2-way interaction
F(3,87)=3.67; p<.0153

Table 1. A' Scores Closed-loop Experiment 1.

SS	df	MS	F	p
Group	1	.017	2.82	.10
Period	3	.028	16.6	.0001
GroupxPer	3	.001	.85	.47

Table 2. B" Scores Closed-loop Experiment 1.

SS	df	MS	F	p
Group	1	.183	.227	.637
Period	3	5.44	17.1	.0001
GroupxPer	3	.086	.27	.84

Table 3. False Alarms Closed-loop Experiment 1.

SS	df	MS	F	p
Group	1	.0057	.055	.81
Period	3	.356	75.3	.0001
GroupxPer	3	.00003	.007	.99

Table 4. Probability of a Hit Closed-loop Experiment 1.

SS	df	MS	F	p
Group	1	.19	4.2	.049
Period	3	.014	.97	.41
GroupxPer	3	.028	1.88	.14

Table 5. EEG Index Scores Closed-loop Experiment 1.

SS	df	MS	F	p
Group	1	1.07	.01	.91
Period	3	31.87	2.67	.054
Array	2	52.4	2.67	.08
Grp x Ar	2	1914.6	97.7	.0001
GrpxPxAr	6	57.4	7.9	.0001

Table 6. Time spent at Each Array Closed-loop Experiment 1.

SS	df	MS	F	p
Group	1	.93	4.66	.04
Period	3	3.3	8.2	.0001
Array	2	27413.3	26.4	.0001
GrpxAr	2	23499.3	22.7	.0001
GrpxPxAr	6	134.8	1.3	.25

Table 7. A' Scores 9 vs 2 Array Experiment 2.

SS	df	MS	F	p
Group	1	.0004	.05	.82
Period	3	.017	20.1	.0001
GroupxPer	3	.003	3.8	.013

Table 8. False Alarms 9 vs 2 Array Experiment 2.

SS	df	MS	F	p
Group	1	.0005	.05	.82
Period	3	.176	40.2	.0001
Group x per	3	.003	.62	.60

Table 9. Probability of a Hit 9 vs 2 Array Experiment 2.

SS	df	MS	F	p
Group	1	.02	.20	.65
Period	3	.01	2.02	.11
Group x per	3	.023	3.67	.015

Table 10. EEG Index Scores 9 vs 2 Array Experiment 2.

SS	df	MS	F	p
Group	1	2.1	.01	.91
Period	3	.49	.14	.92
Array	2	3.19	2.2	.12
Grp x Ar	2	.01	.008	.99
GrpxPxAr	6	1.56	.69	.66