

Sense mode and coupling in a vigilance task¹

JIMMY L. HATFIELD AND MICHEL LOEB²

US ARMY MEDICAL RESEARCH LABORATORY, Fort Knox, Kentucky

This investigation examined the performance of 36 Ss on three 90-min vigilance tasks. As time on task increased, there was a significant decrease in number of detections and false responses and a corresponding significant increase in response latency. The consistent rank order relationship between stimulus conditions suggests that there were uniform trends among these conventional response measures. There was a significant decline in sensitivity (d') for the closely coupled tasks, regardless of the sense mode involved. However, d' remained fairly stable for the loosely coupled visual task. The significant increase in criterion values (β) suggest that Ss adopt a more conservative mode of responding with increasing time on task. Further, these data indicate that a lack of control of coupling effects, rather than sense mode specificity, may have confounded the interpretation of prior research on visual and auditory tasks.

Egan, Greenberg, and Schulman (1961), viewing the monitoring situation within the framework of signal detectability theory have suggested that there is no real decrement in operator efficiency on vigilance tasks; instead, there is a progressive change in the O's criterion in the direction of greater conservativeness in responding such that later in a monitoring period there is a reduced tendency to report faint signals and also a reduced tendency to report environmental or physiological noise as a signal.

In recent years, numerous experiments have demonstrated the relevance and applicability of the theory of signal detectability to the study of vigilance phenomena. However, research which has related the measures of sensitivity (d') and criterion (β) values to performance in visual and auditory tasks has been largely contradictory. As noted by the diverse findings, a decrement in performance during a vigilance task may be due to changes in the O's criteria (Broadbent & Gregory, 1963; Egan et al, 1961; Loeb & Binford, 1964) or to decreased sensitivity (Mackworth & Taylor, 1963).

Due to the unique operating characteristics of each sense modality, it is certainly plausible that there might be two or more types of vigilance "decrements" produced by different underlying mechanisms. In some vigilance tasks there may be a true decline in effective sensitivity. This might be reflected primarily in a change in detections, especially in a "loosely coupled" task like the Mackworth Clock Test and related visual tasks, in which the S may easily make a response incompatible with observation of the display. On other tasks involving more "closely coupled" responses, e.g., auditory tasks, in which Os could not readily fail to receive stimulation from the display, the apparent decrements might be due to changes in the O's criteria.

It is also apparent that the control of the difficulty level of signals in prior research involving two different sense modes has been largely unsuccessful. This is especially true if both visual and auditory vigilance tasks are involved. The experiments involving more than one modality have clearly demonstrated the differential effects upon task performance as a function of differences in task difficulty.

Few vigilance studies have examined the performance of Os monitoring single and multiple signal levels. This manipulation is of special importance because in the real world situation most signal levels are variable. A specific question of interest is whether performance efficiency is approximately the same for single and multiple signal-level conditions, particularly when identical signal intensities have occurred.

The combined results of prior research have shown that the reliabilities of individual performances on a particular task are high, that the correlations between alerted (pretest and posttest) and watch (vigilance) performances are high, but that the correlation between individual performances from one task or

sensory mode to the other is very low (Buckner, Harabedian, & McGrath, 1965). This latter finding must be resolved before the term "vigilance" can be used to represent a general trait or a fundamental characteristic of the individual.

The major objective of this experiment was to examine each of the critical research problems mentioned above. Accordingly, the experimental design incorporated both visual and auditory vigilance tasks, manipulated coupling as an independent variable, compared single and multiple signal level conditions, and permitted the computation of TSD parameters, d' and β .

METHOD

Preliminary Study

Before the main experiment a pilot study was conducted to determine an index of discriminability for each experimental condition. The primary purpose was to obtain performance data which could be used in determining comparable levels of detectability for the visual and auditory signals to be used in the main experiment. Subjects were given extensive practice and a 1-h vigilance task on each sense modality. Test conditions were identical with the auditory and visual vigilance tasks to be used in the main experiment.

For the auditory task, difficulty levels for the Difficult, Moderate, and Easy signals were defined as being increments of 0.8, 1.3, and 1.8 dB (SPL), respectively. Prior research on comparable detection and vigilance tasks (Binford & Loeb, 1963; Green & Swets, 1966; Loeb & Binford, 1963, 1964) indicates that .4-.8 dB increment shifts can produce sizeable differences in detection performance. Using this performance data, specifically the d' values, the physical stimulus values for the visual signals were adjusted accordingly in order to provide a comparable task for each stimulus condition. It should be noted that the rank order of stimulus conditions reported in the main experiment may be largely dependent upon these initial d' values. In addition, differences attributable to stimulus conditions and time blocks may be confounded by differences in mean d' values, particularly those differences across stimulus conditions.

Subjects

Thirty-six volunteer male personnel assigned to the U.S. Army Medical Research Laboratory were employed as Ss in this experiment. Since these individuals had recently completed the standard induction physical examination, no rigorous screening procedure was necessary. The Ss were randomly assigned to a one or three signal level group, 18 Ss in each. There were no specific age controls, but the subject population was limited to Caucasians in order to eliminate the effects of skin pigmentation of the eyelids in the closely coupled visual task.

Vigilance Tasks

Auditory session (closely coupled): Each S listened to a series of pulses of white noise, having an intensity of 70 dB (SPL), a duration of .5 sec, initiated 2.5 sec apart. Aperiodically, increments in intensity were added to the pulses. These more intense pulses were defined as signals and the S's task was to acknowledge detection by pressing a telegraph key. Responses within 2.5 sec after initiation of the signal were scored as *detections*; no response during this interval as a *miss*; other responses as *false alarms*.

Visual session (closely coupled): S's eyelids were taped closed with transparent plastic tape to eliminate eyeblink, increase "coupling" and minimize the effects of "observing responses" (Baker, 1960; Buckner et al, 1965; Elliott, 1960; Holland, 1958; Jerison & Pickett, 1964; Jerison, Pickett, & Stenson, 1965). It is entirely plausible that the low correlations obtained on vigilance performance in different sense modes has been primarily due to

these confounding variables. Each S sat in a straight-back chair with arm rest, approximately 4 ft from a light source. The room was dark, except when the light source was switched on for 0.5 sec intervals during the experimental session. Illumination was provided by various combinations of four incandescent bulbs in an L-shaped plywood box. The open sides of the box permitted light pulses to be reflected on a flat white wall. Since the light source was located on the floor of the Industrial Acoustics Company (IAC) booth (bottom of the diffusing area), there was a vertical gradient of illumination in the immediate area of the light source. However, S's eyes were taped closed and this gradient was not perceptible. The non-signal light pulses were produced by a single, shaded 100-W bulb. Occasionally, increments in intensity were added to the non-signal pulses. These more intense pulses served as signals, and the S's task was to respond by pressing a telegraph key. Recording of responses was identical to that previously mentioned for the auditory session.

Visual session (loosely coupled): Each S sat in a straight-back chair with arm rest, approximately 4 ft from a light source located at eye level in a black plywood box. The S was required to detect a change in the brightness of a light appearing in a 1-in. ground-glass (neutral density filter) covered aperture. Ambient illumination in the test room was furnished by the overhead ceiling light in the IAC room.

Procedure

Each S was available for four consecutive days of testing; an orientation and practice session on the first day and a 90-min watch on the next three days (auditory or visual conditions in counterbalanced order between Ss). On the first day, Ss were given instructions about the tasks and a complete briefing on all experimental requirements. Immediately following this orientation all Ss were required to complete two brief watch periods on each vigilance task. These practice sessions were identical with the auditory and visual conditions to be used in the experimental sessions, except that each watch period was brief (5 min) and the signal rate high (20 signals per 5 min). There was a 1-min rest period between each practice session for each stimulus condition and a 5-min rest period between the sessions for the different sensory modes. Immediate knowledge of results was given for each signal (by the E using the booth intercom system). The Ss in Group 1 (one signal level) were required to detect a signal of

Moderate difficulty, whereas Ss in Group 2 (three signal levels) were presented with an equal number of Difficult, Moderate, and Easy signals during each session.

Each 90-min watch period was preceded by a 5-min warm-up period during which immediate knowledge of results was given for each signal. The Ss were given a 2-min rest period followed immediately by the 90-min main watch. Signal rate averaged one per min, with the intersignal interval varying from 10-120 sec. The distribution between intervals was rectangular, with a mean intersignal interval of 60 sec. Within the limits of the interval range, signal occurrence was random with the restriction that five signals of each difficulty level (for Ss in Group 2) must occur within each 15-min period of the watch. No knowledge of results was presented in these experimental sessions.

Apparatus

The experiment was conducted in an IAC sound-shielded room in which overall ambient sound level is approximately 30-50 dB (SPL). The random noise for the auditory stimuli was generated by a random-noise generator and presented through a pair of earphones embedded in Grason Stadler muffs, the latter producing an additional 15-40 dB attenuation (frequency dependent) of this ambient sound. Auditory signal intensities were adjusted to the three difficulty levels by Daven variable attenuators. The signal and non-signal pulses were measured (calibrated) with a Bruel and Kjaer 6 cc coupler, using a B & K condenser microphone. The light source (brightness-illumination) was furnished by an appropriate combination of incandescent light bulbs contained in a plywood box. Illumination (brightness) levels were measured by a Spectra Brightness Spot Meter for the loosely coupled visual task and a Weston Foot-Candle Meter for the closely coupled visual task. Since the S's eyes were taped closed in the closely coupled visual task, the exact values of illumination transmitted to the cornea are unknown, but the ratios between intensities presumably remained fairly constant for each S. Duration of signals was controlled by Hunter timers and the intersignal intervals by a Gerbrands program timer.

Design

The experimental design was a 3 by 6 by 2 mixed factorial design with repeated measures on the first two factors (stimulus condition—"closely coupled" Auditory, "closely coupled" Visual,

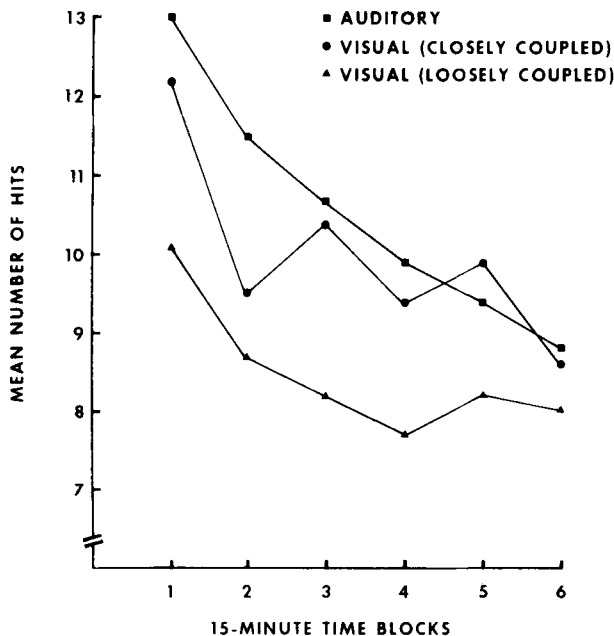


Fig. 1. Mean number of signals detected by the one signal level group.

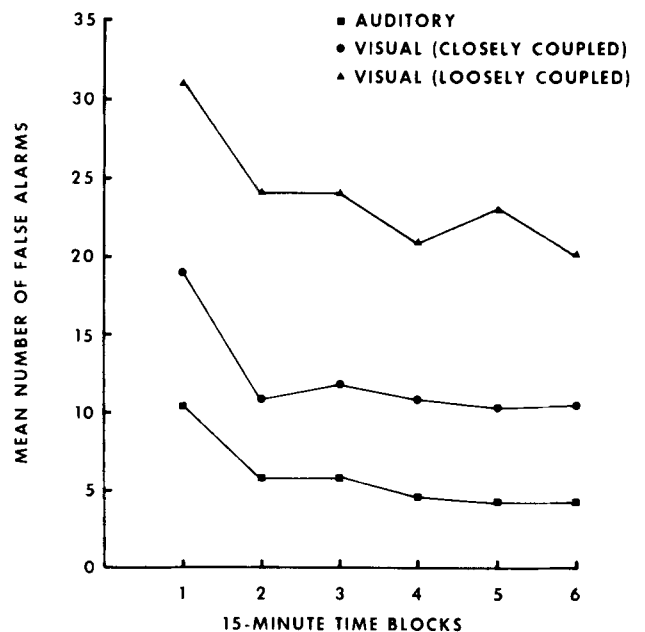


Fig. 2. Mean number of false alarms by the one signal level group.

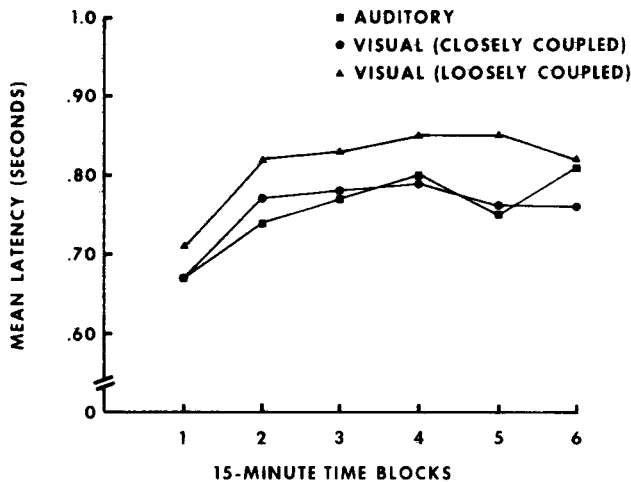


Fig. 3. Mean response latency by the one signal level group.

"loosely coupled" Visual and time blocks). Two groups of 18 Ss each define the third factor (number of difficulty levels). The relationships of primary interest are: number of difficulty levels (between Ss), auditory vs "closely coupled" visual vs "loosely coupled" visual (within Ss), and time on task (within Ss). Time on task was divided into six 15-min time blocks for analyses.

Response measures for signals occurring in each of the two modalities were available in three basic forms: (1) correct detections (Hits), (2) errors of commission (False Alarms), and (3) response time (Latency). However, the derived dependent measures of primary interest are d' and β , which were readily computed by methods described in prior research (Broadbent & Gregory, 1963; Jerison et al, 1965). Pearson product-moment correlation coefficients across sensory modalities and coupling conditions were computed on all of these measures.

RESULTS

In order to present an overall view of the major findings and to facilitate comparisons with prior research, the data are summarized into five categories: (1) conventional measures, (2) signal detection measures, (3) orthogonal comparisons, (4) correlations of performance, and (5) number of signal intensities. One- and three-signal level data are presented separately, since most prior research has employed only one signal-to-noise intensity ratio within a vigilance session.

Although analyses were performed in terms of the 3 by 6 by 2 complete design (Lindquist, 1953, p. 292), supplementary A by B by S analyses (Lindquist, 1953, p. 237) were performed to permit more detailed examination of variables associated with stimulus conditions and time on task. Orthogonal comparisons (Winer, 1962, p. 65) were performed to establish the relative importance of, and functional relationships between, sense mode specificity and coupling (nature of the task). Pearson product correlations were computed on all dependent measures to determine the correlations of performance between visual and auditory tasks.

Conventional Measures

Figures 1-3 present curves for mean number of Hits, False Alarms and Latency measures at one signal level. As time on task increased, the number of detections and false responses both exhibited a general downward trend; whereas latency showed a marked increase. The rank order of stimulus conditions was fairly uniform, with the auditory task producing the highest level of performance, and the loosely coupled visual task producing the lowest. The closely coupled visual task maintained an intermediate position and exhibited the greatest variability, particularly when Ss were observing for three signal intensities within a vigilance session.

HITS

Figure 1 shows the mean number of detections in successive 15-min periods. At one signal level, differences between stimulus conditions were highly significant ($p < .005$); whereas differences between curves were not significant for the three signal level group. However, both groups detected most signals in the auditory condition and the fewest in the loosely coupled visual task. The closely coupled visual task occupied an intermediate position, but clearly exhibits the most variability. For both groups, the decline in hits with increasing time on task was significant ($p < .001$). Stimulus Conditions by Blocks interaction was significant ($p < .05$) for the one signal level.

FALSE ALARMS

Figure 2 presents the mean number of false detections over successive time blocks. The differences between stimulus conditions and the downward trend with increasing time on task were significant ($p < .001$). The primary characteristic of these curves is the sharp decline in false alarms during the first 30 min of the session followed by a more or less gradual decline during the remaining periods. As compared with mean number of hits, there is an inverse relationship in the rank order of stimulus conditions, i.e., Ss exhibiting the largest number of hits on the auditory task also made fewer false alarms.

LATENCY

Latencies for the one signal level group are summarized in Fig. 3 as a function of successive 15-min periods. The curves show an overall increase in latency with time on task. This increase is significant for both groups, one signal level ($p < .001$) and three signal levels ($p < .005$). Differences between stimulus conditions were not significant for either group.

The preceding data indicate a relative consistency between the several conventional measures. As time on task increases, there is a significant decrease in hits and false alarms ($p < .001$) and a corresponding increase in response latency ($p < .005$). Further, there appears to be a consistent, rank order of stimulus conditions based upon difficulty of discrimination. The reason for considerable variability in the closely coupled visual task cannot be established.

Signal Detectability Measures

The relevance and applicability of the theory of signal detectability (TSD) to the study of vigilance phenomena has been well

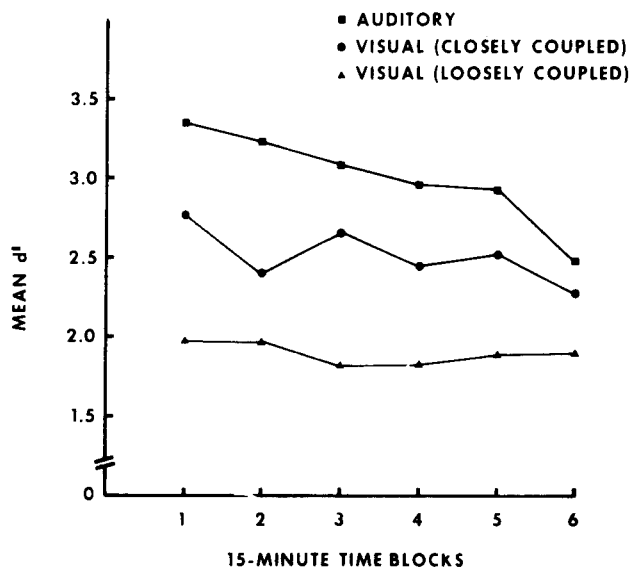


Fig. 4. Mean d' values for signals of moderate difficulty (one signal level group).

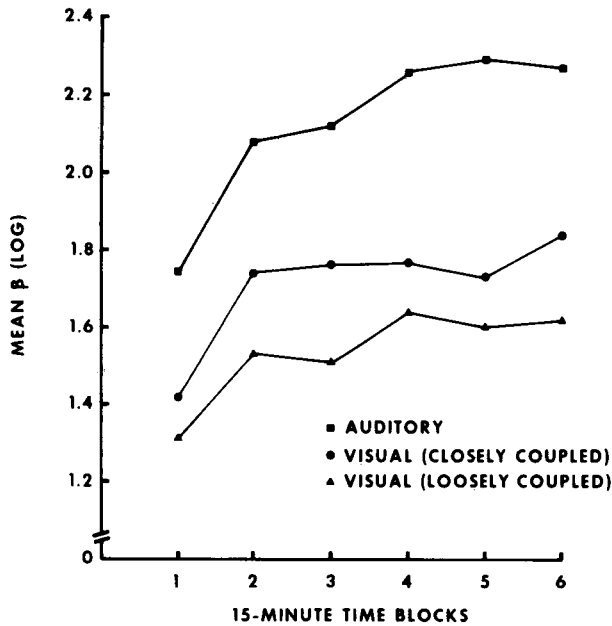


Fig. 5. Mean β values for signals of moderate difficulty (one signal level group).

established and its potential theoretical significance documented (Broadbent, 1964; Green & Swets, 1966). However, past research which has sought to relate the measure of sensitivity (d') and criterion (β) shift to performance in visual and auditory vigilance tasks has been largely contradictory. It was felt that by deriving both d' and β values from the data of the present experiment, some progress toward resolving these contradictions might be achieved.

There are methodological difficulties involved in the transformation of the data. Specifically, when proportions of detections are 0% or 100% or the percentages of false alarms are 0%, it is necessary to assign arbitrary percentages halfway between the observed 0% and 100% value and the next possible value. Such a procedure has been described by Jerison et al (1965) and permits an estimate of d' and β to be computed for vigilance data. The techniques used in this experiment did reduce the magnitude of this problem. For example, none of the Ss exhibited 0% false alarm performance for the 90-min vigil, nor 0% performance on hits for any of the 15-min or the total 90-min analyses. However, analysis by 15-min time blocks did indicate an expected increase in 0% false alarm rate. There were approximately two to nine Ss (3.7% to 16.6% of the scores) exhibiting a 0% false alarm rate; likewise, the number of Ss exhibiting 100% performance on hits follows a similar pattern (one to five Ss and approximately 1.8% to 9.2% of the scores). These figures are rather small as compared with the acute problem reported by Jerison. Of the 72 Ss in their experiment, 27 Ss did not produce a single false alarm during the 80-min vigilance task. Since three of his Ss also had 100% performance on hits, exact values of d' and β could be determined on just 42 Ss.

There are additional problems associated with the three signal level group; namely, the assignment of false alarms to a specific intensity level. Therefore, it is necessary to compute an average sensitivity (d') and criterion level (β) for the three signal levels. While there is no precedent for this procedure the technique seems reasonable, since the Moderate intensity values were equated prior to the main experiment ($\pm .11 d'$) and the Easy and Difficult intensity levels spaced approximately the same d' distance apart. This procedure will not permit a d' or β comparison of Moderate signal intensities (between groups) but offers an opportunity to

examine trends in total or overall sensitivity and criterion values for each group.

Shown in Fig. 4 are mean d' values for the one signal level group. There were significant differences between stimulus conditions for both groups ($p < .001$), and a significant decline in sensitivity for the one signal level group ($p < .005$). However, the slopes are not uniform, as evidenced by the Stimulus Conditions by Time Blocks interaction ($p < .05$). The nature of these effects are clearly shown in Fig. 4, e.g., the d' value in the loosely coupled visual task is fairly uniform throughout the vigilance session, but the two closely coupled tasks show a gradual decline in sensitivity.

Mean criterion indices ($\log \beta$) are presented in Fig. 5. The underlying data were quite variable and contained a few extreme scores in each condition. Therefore, β 's were transformed to their common logarithms prior to multivariate analysis. Differences between stimulus conditions and effects over time blocks were highly significant ($p < .001$). Stimulus Conditions by Time Blocks interaction was not significant for either group. The differences between curves and the uniform increase in β with time on task are readily apparent.

If Ss are observing for signals at a constant signal-to-noise intensity ratio, it appears that sensitivity declines with increasing time on task for the closely coupled tasks, regardless of the sense mode involved. Sensitivity associated with the loosely coupled visual task remains fairly stable throughout the vigilance session. Sensitivity does not show a significant decline over time periods for those Ss that are required to detect three signal-to-noise intensity ratios as critical signals.

Criterion values, on the other hand, increase significantly during the vigilance session for both groups, irrespective of number of signal intensities, sense mode or coupling conditions. There was a tendency, therefore, to adopt a more conservative mode of responding with increasing time on task.

Orthogonal Comparisons

In order to further explore the effects of sense mode specificity and coupling on performance in vigilance tasks, orthogonal comparisons were computed on that portion of the variance attributable to stimulus conditions. First, sense mode effects were determined by comparing the treatment means for the closely coupled auditory and closely coupled visual task. In this case, coupling was held constant and sense mode varied. If sense mode was *not* significant, the effects of coupling were examined by comparing the loosely coupled visual task with the average of the auditory and closely coupled visual tasks. If sense mode was significant, the comparison between closely and loosely coupled

Table 1
Orthogonal Comparisons between Stimulus Conditions to Determine Sense Mode and Coupling Effects

ONE INTENSITY LEVEL					
Dependent Measure					
Stimulus Condition	Hits	False Alarms	d'	β	Latency
Sense Mode	<1.00	2.86	8.09**	39.37**	<1.00
Coupling	13.76**	20.99**	11.60**	3.85	6.05*
THREE INTENSITY LEVELS					
Dependent Measure					
Stimulus Condition	Hits	False Alarms	d'	β	Latency
Sense Mode	<1.00	1.50	3.7	1.9	1.02
Coupling	<1.00	16.89**	18.94**	17.6**	2.80

F_{0b} values of each group of Ss on all dependent measures

Critical value is $F_{.99}(1,321) = 6.63$

* $p < .05$

** $p < .01$

Table 2
Correlations between Stimulus Conditions for Conventional and Detection Measures (N = 18)

ONE INTENSITY LEVEL					
Stimulus Condition	Dependent Measure				
	Hits	False Alarms	d'	β	Latency
Auditory vs Closely Coupled Visual	.65**	.15	.34	.33	.76**
Auditory vs Loosely Coupled Visual	.48*	.21	.27	.32	.75**
Closely vs Loosely Coupled Visual	.62**	.43	.47**	.92**	.70**

THREE INTENSITY LEVELS					
Stimulus Condition	Dependent Measure				
	Hits	False Alarms	d'	β	Latency
Auditory vs Closely Coupled Visual	.63**	.32	.58*	.54*	.10
Auditory vs Loosely Coupled Visual	.64**	.22	.50	.34	.11
Closely vs Loosely Coupled Visual	.38	.61**	.32	.10	.07

Pearson product-moment correlations. Critical values for two-tailed tests are as follows: $r_{.10} = .40$; $r_{.05} = .468$; $r_{.01} = .590$

* $p < .05$

** $p < .01$

visual tasks was performed to estimate coupling effects. Other comparisons are, of course, feasible. Table 1 includes F_{Oby} values of each group on all dependent measures. This F ratio has one degree of freedom for the numerator and 321 degrees of freedom for the denominator. For these data the critical value is $F_{.99} = 6.63$.

At one intensity level, coupling effects account for a significant portion of the variance in four out of five measures; whereas, sense mode effects are largely restricted to the detection measures, d' and β . If Ss, however, are required to observe for three intensity levels, the significant effects of coupling are equally dramatic, but in this case, sense mode differences did not account for a significant portion of the variance on any of the five measures.

Correlations of Performance

Investigations concerned with correlations of performance between visual and auditory vigilance tasks have generally reported low correlations (Buckner & McGrath, 1963; Buckner et al, 1965). These studies have largely emphasized the mode specificity of individual performances and have suggested that performance on one sense mode could not be predicted accurately from performance on another. Further, the findings have argued against the existence of a general mechanism as well as the status of vigilance as a general personality characteristic of the individual which determines the level of his performance regardless of the task. It should also be noted that percentage of correct detections has been the primary response measure.

In order to explore these possible relationships within the context of the present experiment, scores for each individual were computed for each 90-min session as follows: (1) total number of hits and false alarms, (2) mean response latency, and (3) overall d' and β values. Pearson product-moment correlations were then computed between individual performance scores for each stimulus condition. Correlations for β were computed after a normalizing log transformation. These data have been summarized in Table 2. The critical importance of coupling effects and the selection of dependent measures are evident.

There are numerous significant correlation coefficients and

some striking relationships between stimulus conditions, type of dependent measure and number of intensity levels. All of the correlations are positive and approximately 50% of the correlations are significant. In the one signal level condition, hits and response latency discriminate reliably between sense mode and coupling conditions. These correlations are significant and highly consistent. Between the visual tasks in the one signal level condition there is a remarkably high correlation of β values which reflects a very consistent mode of responding for the visual modality.

Number of Signal Intensities

Surprisingly few studies of vigilance performance have manipulated signal variables, such as intensity and duration, over a significant range (Jerison & Pickett, 1963). C. H. Baker (1963) did vary signal duration within a vigilance session and Loeb and Binford (1963) varied signal intensities between sessions. However, to our knowledge, signal intensity has not been manipulated within a vigilance session.

A Lindquist Type VI mixed analysis was performed on the data to determine the differences between one- and three-signal level groups. Two separate analyses were computed for Hits: (1) total number of hits and (2) number of hits on the Moderate intensity level converted to percentages for each group.

There were no significant differences between groups on any of the five dependent measures. This finding was anticipated primarily because of the following factors: first, preliminary data equated difficulty of discrimination; secondly, the methodological difficulties of isolating those effects specifically associated with the Moderate intensity level in the three signal level group; and third, the inter-subject comparisons are usually much less precise in a mixed factorial design. Most of the variance attributable to stimulus conditions, time blocks and respective interactions were highly significant. An examination of these interactions may provide valuable insight into the complex relationships involved in presenting several intensity levels as critical signals.

Similar results were obtained from the two separate analyses of Hits. Main effects for stimulus conditions and time blocks were significant ($p < .001$). The Time Blocks by Groups interaction ($p < .025$) is plotted in Fig. 6. There is a significant decline in mean number of hits, with a sharp decrease in performance during the

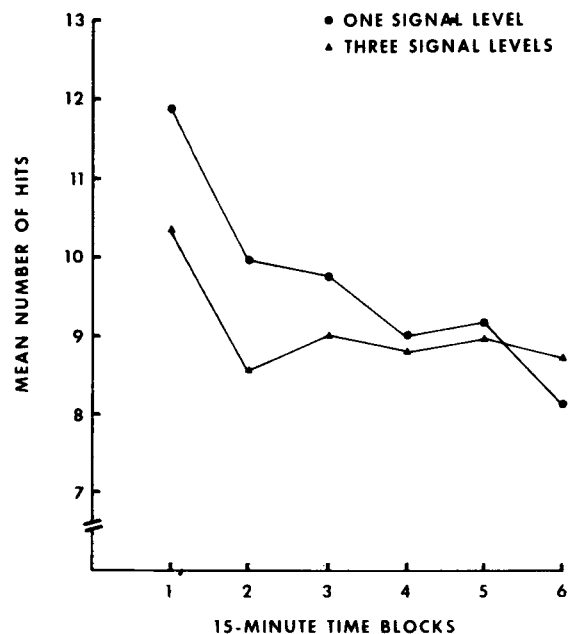


Fig. 6. Time blocks by groups interaction for mean number of hits.

first 30 min for both groups. However, while the one signal level group exhibits a continual decline in detections, Ss attending to three signal intensities maintain a consistent level of performance for the remainder of the vigilance session. A plot of the Stimulus Conditions by Time Blocks interaction ($p < .01$) is very similar to Fig. 1. The Ss detected more signals in the auditory condition, and the fewest number of hits in the loosely coupled visual condition. Performance in the closely coupled visual task occupies an intermediate position, but clearly exhibited the greatest variability across time periods.

False alarms was the only measure with a Stimulus by Groups interaction ($p < .025$). Main effects for stimulus conditions and time blocks were highly significant ($p < .001$). There was a sharp decline in false alarms during the first 30 min of the session followed by a more or less gradual decline during the remaining time periods.

Main effects for stimulus conditions ($p < .05$) and time blocks ($p < .001$) are both significant for response latency. There were no interactions, therefore a plot of these main effects would be quite similar to Fig. 3. There is a significant increase in response latency with time on task, regardless of stimulus conditions or number of signal intensities.

On the sensitivity measure (d'), main effects for stimulus conditions was significant ($p < .001$); however, main effects for time blocks was not significant. This is because the two closely coupled tasks show a gradual decline in sensitivity for one signal intensity but a corresponding average increase for three signal intensities (Fig. 4). Curves plotted for the Stimulus Conditions by Time Blocks interaction ($p < .05$) would be quite similar to the above mentioned illustration of mean d' values. The Time Blocks by Groups interaction ($p < .001$) is presented in Fig. 7. By summing within stimulus conditions, it is readily apparent that the techniques utilized in this experiment for equating overall intensity levels between groups was largely successful. Equating sensitivity levels between stimulus conditions for each intensity level is infinitely more difficult.

Results for mean criterion indices ($\log \beta$) are very similar to mean response latency. There are main effects for stimulus conditions and time blocks ($p < .001$) and no interactions. There was a uniform tendency to adopt a more conservative mode of responding with increasing time on task.

DISCUSSION

The interpretation of these data and the generality of the findings are somewhat restricted due to current methodological problems associated with this area of research.

First, equating stimulus conditions between sense modes continues to be a difficult task, whether one uses psychophysical methods (Broadbent & Gregory, 1963) or the more recent detection measures (Green & Swets, 1966). In the majority of detection experiments thousands of observations are obtained on a few, highly trained Ss. This technique has permitted very precise estimates of sensitivity for detection experiments, but has its limitations for vigilance research. Since we are often interested in the performance of large groups of relatively naive Os, several different groups of Ss were used during the preliminary study and the main experiment. Even so, it was possible to equate mean sensitivity levels (d') between stimulus conditions within a small range, e.g., Difficult intensity level $\pm .35$; Moderate intensity level $\pm .11$; and the Easy intensity level $\pm .23$. Average sensitivity levels were also maintained within a small range ($\pm .11$) for each stimulus condition (Difficult, Moderate and Easy sensitivity values combined). Equating stimulus conditions between sense modes on the basis of mean sensitivity values appears to be a most useful and promising technique.

Secondly, what is the best method for computing detection measures from vigilance data? An effort was made to reduce the magnitude of the problem by (1) specifying an observation interval to avoid the complications inherent in those designs which utilize undefined observation intervals (Egan et al, 1961; Green, 1966) and (2) using a relatively high signal rate. Since several investigations (Mackworth & Taylor, 1963) have reported similar

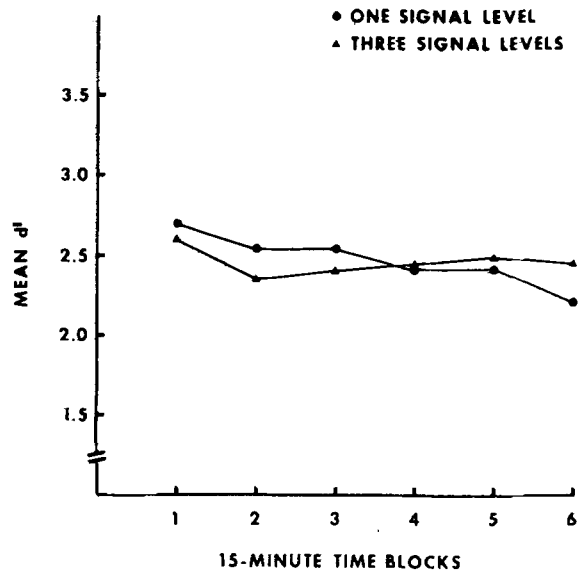


Fig. 7. Time blocks by groups interaction for overall mean sensitivity levels.

trends for high and low signal rates, the use of a high signal rate to study the vigilance decrement seems justified. Nevertheless, the absence of false alarms and 0% or 100% performance on number of detections, does require an approximation of detectability indices computed from detection and false alarm probabilities. Jerison et al (1965) discuss this persistent problem and present a method for treatment of these data. A somewhat similar procedure was adopted in this experiment to provide estimates of d' and β .

Third, if several intensity levels are presented during a vigilance session, how can one determine which false alarms are associated with a specific intensity level? This problem has not been resolved, and therefore average detection measures were computed for the three intensity level condition. Since the average sensitivity level within and between each stimulus condition was carefully equated, this technique seems tenable. At any rate, if the theory of signal detectability (TSD) is to be a substantive theory of human performance, then it must be applicable to situations studied in the laboratory and these situations typically involve the detection of several intensities in the same session (Gettys, 1966).

There is strong need for a description of the conditions under which decrements will and will not occur, both as a function of stimulus conditions and of the response measures employed (Bergum & Klein, 1961). A fairly large literature testifies to the fact that there is a decrement in performance, as measured by the proportion of correct detections, or hits, as the watch progresses. In the experiments that included an attempt to estimate the false alarm probability, this probability was found to decrease along with the hit probability (Green & Swets, 1966; Loeb & Binford, 1964). In this experiment, there was a significant decline in the number of hits and false alarms with increasing time on task, regardless of sense mode, coupling conditions, or the number of signal intensities presented to the O. Further, in this study, the rank order of stimulus conditions was fairly uniform. Ss detected most signals in the auditory task, an intermediate number in the closely coupled visual task, and achieved the fewest number of hits in the loosely coupled visual task. Of particular relevance is the noticeable inverse relationship in this rank order of stimulus conditions for mean number of false alarms. Although difficulty of discrimination appears to determine the initial and overall levels of performance for each stimulus condition, the significance of the progressive decline in hits and false alarms is not known (Buckner et al, 1965; Buckner & McGrath, 1963; Loeb & Binford, 1963). It may be due to progressive decline in responding, but this fact has not been conclusively demonstrated.

A summary of the use of reaction time or response latency as a measure of perceptual vigilance has been recently reported in the literature (Buck, 1966). In general, the experimental evidence confirms the prediction that when detection rate decreases, reaction time increases; and that as fewer signals are detected, the O takes longer to respond to those signals which he does detect (Loeb & Binford, 1964). Differential effects have been reported for difficulty of discrimination (Loeb & Binford, 1963) and as a function of the frequency of auditory signals (Binford & Loeb, 1963).

The consistent rank order relationships between the stimulus conditions of this experiment suggest that there are uniform trends among conventional response measures. The auditory task had the greatest number of hits, fewest false alarms, and the shortest response time. Conversely, the loosely coupled visual task was characterized by the fewest number of hits, most false alarms and the longest latency.

Recently, there have been several attempts to apply signal detection theory to the vigilance situation. A direct application of TSD is feasible by considering vigilance effects to be due to changes in criterion by the O or to changes in signal or noise variance with time. Results with different sense modes have been largely contradictory. Visual experiments reported by Mackworth and Taylor (1963) indicate that the detectability of the signal decreases with time on watch. Further, the decrease in $\log d'$ from its initial value was linearly related to the square root of the time on watch and the rate of decline was independent of most experimental variables employed. Broadbent and Gregory (1963) found no significant change in d' for either a visual or auditory task but an increase in the O's criterion during a vigil. However, the increase in caution was highly dependent upon the initial mode of responding, e.g., cautious Os became even more conservative; whereas, liberal Os maintained the same criterion throughout the vigilance session. Results from additional research with auditory signals (Binford & Loeb, 1966; Loeb & Binford, 1963, 1964) have not been conclusive. Initially, these investigators obtained data explicable in terms of changes in criteria for responding. There was some reason to believe, however, that sensitivity was also changing. Their most recent publication (1966) reported a small decrease in d' and a slight increase in β within sessions. Although a direct comparison of visual and auditory experiments is most difficult due to differences in experimental and computational procedures, it has been suggested that there may be at least two kinds of vigilance "decrements" produced by different underlying mechanisms. Further, it has been argued that type of signal or mode was the more important variable and that as a result of the nature of the "coupling," the auditory display was more demanding of the S's attention. The present experiment was designed to examine these various hypotheses.

If the critical signal remains at a constant signal-to-noise intensity ratio, there is a significant decline in sensitivity with increasing time on task for the closely coupled tasks, regardless of the sense mode involved. However, sensitivity associated with a loosely coupled visual task remained fairly stable throughout the vigilance session. This evidence is not compatible with the explicit assumptions of conditioning, reinforcement or observing response theory. In a loosely coupled task, the observer may easily make a response incompatible with observation of the display. It has been postulated that non-reinforcement of such responses produces a decline in observing responses, thereby reducing the number of detections (Holland, 1958; Jerison & Pickett, 1963; Jerison et al, 1965). ROC curves derived from this kind of data should reflect decrements in sensitivity with time (Loeb & Binford, 1964). In the present experiment, however, there was no significant decline in sensitivity on the loosely coupled visual task. An alternate explanation, in terms of neural mechanisms, e.g., habituation, seems plausible. Specifically, it is suggested that the irrelevant observing responses may actually inhibit habituation effects and thereby permit a relatively stable sensitivity level to be maintained throughout the vigilance session. The habituation hypothesis is certainly consistent with the data of those Ss who were required to detect three signal-to-noise intensity ratios as critical signals. In this experimental condition, there was no significant decline in

sensitivity for any of the vigilance tasks. Results of this study may be viewed as somewhat analogous to research on multiple displays and response complexity. Although these data have rarely been analyzed in TSD terms, it is generally agreed that the vigilance decrement is negligible (Adams, 1963). These investigators found damaging decrements for simple tasks, but demonstrated that the vigilance decrement can be virtually eliminated for some complex monitoring tasks. Likewise, in this experiment, observing for three intensity levels prevented a decline in sensitivity with time on task. At least, that portion of the decrement due to changes in sensitivity was reduced. The use of multiple intensity levels as an effective means for sustaining alertness finds theoretical support from the "filter theory" hypothesis and the arousal or activation hypothesis (Frankmann & Adams, 1962). Finally, from a pragmatic viewpoint, it may not be necessary to alternate presentations to different sense modes (Gruber, 1964) or to present redundant signal information (Buckner & McGrath, 1963) in order to achieve improved detection performance.

The Ss' criteria of what constitutes a signal seems to exert a major influence in monitoring tasks. In this experiment, criterion values increased significantly during the vigilance session, irrespective of number of signal intensities, sense mode or coupling conditions. There is a tendency, therefore, for the O to adopt a more conservative mode of responding with increasing time on task. In signal detection theory, sensitivity and criterion are fixed for a given set of initial conditions. The theory does not specify how an O might modify estimates of the initial conditions. However, it is obvious that such changes would have relatively simple consequences. If β increases, then there would be a corresponding decrease in hits and false alarms, a fact which is summarized in the operating characteristic, or isosensitivity curve. The Egan et al (1961) hypothesis of criterion change considers both the progressive change in hits and false detections, and the results of this experiment agree with that prediction. A systematic set of experiments should be performed to determine those factors influencing criterion changes within a vigilance session. Until such definitive tests are accomplished, it would be presumptuous to postulate explanatory concepts.

One of the major objectives of this study was to demonstrate the importance of "coupling" effects in vigilance research (Broadbent & Gregory, 1965; Buckner & McGrath, 1963; Elliott, 1960; Loeb & Binford, 1964). Coupling has always been confounded with sense modality and has not been manipulated as an experimental variable. Its effects have been largely suggestive, as a post hoc explanation of the disparate results between sense modes. Therefore, a closely coupled visual task was designed to permit an examination of both sense mode and coupling effects. The experimental design was not symmetrical, due to a lack of an appropriate loosely coupled auditory task. Nevertheless, orthogonal comparisons of that portion of the variance due to stimulus conditions identify coupling as an important independent variable. At one intensity level, coupling effects account for a significant portion of the variance in four out of five measures; whereas, sense mode effects are largely restricted to the detection measures, d' and β . The coupling effects are even more dramatic if Ss are required to observe for three intensity levels. In this experimental condition, sense mode differences do not account for a significant portion of the variance on any of the five response measures; whereas, coupling effects were significant for false alarms, d' and β values. It would appear that a lack of control of coupling effects rather than sense mode specificity may have confounded the interpretation of prior research on visual and auditory tasks. Additionally, the hypothesis of two kinds of vigilance decrements produced by different underlying mechanisms may be incorrect. The implication of these results is that the factors operating to cause a decrement in vigilance performance probably are central in origin. This suggests that the mechanics underlying the vigilance decrement are not associated specifically with characteristics of the receptors themselves, but rather are to be found in neural centers common to all modes of responding.

The generality of the concept of vigilance has been frequently questioned, primarily due to low correlations between performance measures for visual and auditory monitoring tasks. Accord-

ing to Buckner et al (1965), the correlation between individual performances on the visual and auditory tasks was .30 under alerted conditions and .24 under watch conditions, indicating that individual performances were highly mode-specific. It was evident that performance on one mode could not be predicted accurately from performance on the other, in spite of high performance reliabilities within mode. The reasons for the sense mode specific differences among Ss has long been recognized as an important target for research. Jerison and Pickett (1963) account for the low correlations on the assumption that auditory and visual tasks share a common component of observing response variance with respect to the neural attention systems, but that visual monitoring has a major special factor associated with orientation responses. In other words, correlations among vigilance tasks should be related to the similarity of type of observing behavior required. Since this experiment has emphasized coupling effects, an attempt was made to examine the observing response hypothesis. It was also apparent that the choice of a dependent measure may have been a critical variable in prior research. Therefore, after signal detection and false alarm rates were tabulated for each condition for each O, the corresponding normal deviates and ordinates of the normal curve were obtained, and from these the signal detection parameters reflecting sensitivity (d') and conservatism in responding (β) were computed. These detection measures were used to examine the hypothesis of common factors underlying efficiency of detection and common response biases involved in the detection of visual and auditory signals. Even if sensitivity is not correlated across modalities, it is reasonable to hypothesize that a person adopting a particular criterion for responding in one modality will in all probability do so when he responds in another modality.

Numerous significant correlation coefficients were obtained but it is difficult to identify and isolate uniform effects for a specific factor. This is not surprising, since many of the so-called independent variables of vigilance performance are interdependent (Jerison & Pickett, 1963). In addition, a proper analysis of the correlations between TSD and conventional measures required more than a simple application of TSD notions in which d' is treated as a measure of signal detectability and β as a measure of the criterion (Loeb & Binford, 1964; Jerison et al, 1965; McGrath, 1965). Although a precise interpretation of these correlations of performance cannot be attempted at this time, there are some important, discernible relationships. First, all of the correlations are positive and approximately 50% of the correlations are significant ($p < .01$). Within a factor analytic framework, we would at least assume that all of these correlations are measuring some common variance. Secondly, five of the six correlations for total number of hits are significant, with the last correlation approaching significance. Performance can be predicted between both sense mode and coupling conditions, using one or three intensity levels as critical signals. Third, at one intensity level, mean response latency discriminates reliably between sense mode and coupling conditions. All of these correlations are significant and highly consistent. Fourth, in the one intensity level condition, performance between visual tasks could be predicted on all five dependent measures. The highly significant correlation between criterion values reflects a consistent mode of responding across tasks involving the visual modality. The moderate but not statistically significant correlations of d' and β between auditory and visual tasks should not be ignored. These cross-modality correlations are particularly interesting and potentially of theoretical importance. It may signify that there is a cognitive or other central component of signal detection transcending modalities.

Due to methodological problems, a concise interpretation of correlations of performance for the three intensity level condition is not feasible for false alarms, latency, d' and β . Until more reliable techniques are available, an analysis of the effects for each intensity increment would be presumptuous.

REFERENCES

ADAMS, J. A. Experimental studies of human vigilance (Final Report), tech. Rep. No. ESD-TDR-63-320, Aviation Psychology Laboratory, University of Illinois, Urbana, Illinois, February, 1963.

- BAKER, C. H. Observing behavior in a vigilance task. *Science*, 1960, 132, 674-675.
- BAKER, C. H. Signal duration as a factor in vigilance tasks. *Science*, 1963, 141, 1196-1197.
- BERGUM, B. O., research, U. S. Army Air Defense, Human Research Unit, Fort Bliss, Texas, HRU Report No. 8, 1961.
- BINFORD, J. R., and detection of obscure visual signals. *Percept. mot. Skills*, 1963, 17, 735-746.
- BINFORD, J. R., & LOEB, M. Changes within and over repeated sessions in criterion and effective sensitivity in an auditory vigilance task. *J. exp. Psychol.*, 1966, 72, 339-345.
- BROADBENT, D. E. Vigilance. *Brit. Med. Bull.*, 1964, 20, 17-20.
- BROADBENT, D. E., & GREGORY, M. Vigilance considered as a statistical decision. *Brit. J. Psychol.*, 1963, 54, 309-323.
- BROADBENT, D. E., & GREGORY, M. Effects of noise and of signal rate upon vigilance analyzed by means of decision theory. *Hum. Factors*, 1965, 7, 155-162.
- BUCK, L. Reaction time as a measure of perceptual vigilance. *Psychol. Bull.*, 1966, 65, 291-304.
- BUCKNER, D. N., HARABEDIAN, A., & McGRATH, J. J. Individual differences in vigilance performance. *J. enrg. Psychol.*, 1965, 4, 69-85.
- BUCKNER, D. N., & McGRATH, J. J. A comparison of performances on single and dual sensory mode vigilance tasks. In D. N. Buckner and J. J. McGrath (Eds.), *Vigilance: A symposium*. New York: Holt, 1963. Pp. 53-71.
- EGAN, J. P., GREENBERG, G. Z., & SCHULMAN, A. I. Operating characteristic, signal detectability and the method of free response. *J. Acoust. Soc. Amer.*, 1961, 33, 993-1007.
- ELLIOTT, E. Perception and alertness. *Ergonomics*, 1960, 3, 357-364.
- FRANKMANN, J., & ADAMS, J. A. Theories of vigilance. *Psychol. Bull.*, 1962, 59, 257-272.
- GETTYS, C. F. M. The theory of signal detectability extended to the case of amplitude known statistically. Doctoral dissertation. University of Louisville, August, 1966.
- GREEN, D. M. (Chairman). American Association for the Advancement of Science, symposium on signal detection with an undefined observation interval, Washington, D. C., 29 December 1966.
- GREEN, D. M., & SWETS, J. A. (Eds.), *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- GRUBER, A. Sensory alternation and performance in a vigilance task. *Hum. Factors*, 1964, 6, 3-12.
- HOLLAND, J. G. Human vigilance. *Science*, 1958, 128, 61-67.
- JERISON, H. J., & PICKETT, R. M. Vigilance: A review and evaluation. *Hum. Factors*, 1963, 5, 211-238.
- JERISON, H. J., & PICKETT, R. M. The importance of the elicited observing rate. *Science*, 1964, 143, 970-971.
- JERISON, H. J., PICKETT, R. M., & STENSON, H. H. The elicited observing rate and decision processes in vigilance. *Hum. Factors*, 1965, 7, 107-128.
- LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. New York: Houghton Mifflin Co., 1953.
- LOEB, M., & BINFORD, J. R. Some factors influencing the effective auditory intensive difference limen. *J. Acoust. Soc. Amer.*, 1963, 35, 884-891.
- LOEB, M., & BINFORD, J. R. Vigilance for auditory intensity changes as a function of preliminary feedback and confidence level. *Hum. Factors*, 1964, 6, 445-458.
- MACKWORTH, J., & TAYLOR, M. M. The d' measure of signal detectability in vigilance-like situations. *Canad. J. Psychol.*, 1963, 17, 302-325.
- McGRATH, J. J. Performance sharing in an audio-visual vigilance task. *Hum. Factors*, 1965, 7, 141-154.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

NOTES

1. This report is based on a dissertation submitted by the senior author in partial fulfillment of the requirements for the Ph.D. degree at Vanderbilt University. The authors wish to acknowledge the assistance of Dr. Leland E. Thune, and other members of the dissertation committee, for their suggestions and corrections on the manuscript.
2. The authors wish to acknowledge the technical assistance of Mr. Gerald B. Ryan during the planning and execution of this experiment.

(Accepted for publication March 8, 1968.)