Response requirements and performance in a visual vigilance task

MICHAEL J. GURALNICK† and KAREN G. HARVEY
The American University, Washington, D.C. 20016

Three groups of Ss performed a visual vigilance task with either one (standard vigilance procedure), two (binary procedure), or four (rating procedure) keys available as response indicators. Data analyzed within the framework of the theory of signal detection revealed that the criterion, β, increased for all groups but was considerably lower for the rating method group. The sensitivity parameter, d′, remained constant over time and was also found to be independent of the response requirement. Results were discussed in terms of the relationship between psychophysical procedures and vigilance tasks.

Many recent studies in vigilance have analyzed their data within the framework of the signal detection theory (SDT) (e.g., Loeb & Binford, 1968). In doing so, according to the theory, one is able to obtain independent estimates of the S's willingness to report a signal or his criterion (β) and his sensory acuity or sensitivity (d′) (Green & Swets, 1966). SDT further suggests that the d′ performance measure should be independent of the psychophysical procedure employed. Psychophysical studies comparing the binary procedure (S reports only "yes," he sees a signal, or "no," no signal present, on each trial) with the rating procedure (S is provided with a number of response alternatives and is asked to report his confidence as to the existence of a signal on each trial) have tested this latter assumption, with generally favorable results. For auditory tasks, Egan, Schulman & Greenberg (1959), Markowitz & Swets (1967), and Emmerich (1968) found excellent agreement between the two procedures, whereas Watson, Rilling, & Bourbon (1964) found d′ values to be slightly higher for the binary procedure. For visual detection tasks, the original Swets, Tanner, & Birdsall (1961) studies produced equivocal results; however, Nachmias (1968) did find sensitivity to be independent of the number of response alternatives employed (two vs four).

Although a number of vigilance studies carried out within the SDT framework have utilized the rating and binary methods (or variations of these procedures), an adequate assessment of the relationship between these procedures and vigilance performance is not available. The standard vigilance task generally provides S with a single response button and instruction to respond only on those trials when he believes he detects a signal. Comparisons between this one button case and the rating procedure by Loeb & Binford (1964) indicated that these methods may not be equivalent. They found that the rating method yielded a much larger proportion of false alarms than the simple detection procedure, with no differences with respect to correct detections. A further study (Binford & Loeb, 1966) did not support this result and firm conclusions relating SDT indices to response requirements were not achieved in either study.

The purpose of the present experiment was to determine the relationship between response factors and performance in a visual vigilance task. One group of Ss was provided with a single button as in the standard vigilance task, a second group was provided with two buttons, one labeled "yes" and the other "no," and required to respond on every trial (binary procedure), and a third group was given four buttons and required to respond on every trial (rating procedure). Both conventional and SDT measures were used to compare these procedures.

SUBJECTS

Thirty-six students, obtained from the introductory psychology subject pool at The American University served as Ss. Ss were divided equally among the three experimental groups on a random basis.

APPARATUS

The visual display consisted of an internally illuminated Model 200 Simpson microamp meter with the scale markings removed. The S, seated approximately 4 ft from the display in a dimly illuminated subject room, received both instructions and white noise (approximately 75 dB provided by a Grason-Stadler Model 901B white-noise generator) through a Koss (Model SP-3XC) headset. When a stimulus event occurred, the pointer of the meter was deflected briefly from its null position. If this event was a nonsignal or "noise" trial, the pointer would deflect to a marked center position. For a signal trial, the pointer would deflect to a position approximately 1 mm beyond this center point. The right arm of the subject chair held the response buttons. (The left hand was used for recording the GSR and heart rate. Respiration rate was also obtained by means of a chest bellows. However, these data are not reported here.) All events were automatically controlled by the appropriate electromechanical programming and recording apparatus located in an adjacent room.

PROCEDURE

The task for each S was to detect any pointer deflection that went beyond the marked middle position of the meter. All Ss received the same training. Taped instructions, correlated with the programming apparatus, explained the nature of the task to Ss and presented 15 signal and noise trials. Ss were informed as to whether the next event was a signal or noise before each trial at this stage. Finally, 15 trials without feedback, five of which were signals (randomly interspersed), were presented. If the S reached a criterion of four correct detections with no more than one false alarm, he participated in the main portion of the task that immediately followed. (Almost all Ss reached criterion immediately here since the signal was readily detectable for an alert individual.) At this point, Ss were cautioned that signals would occur much less frequently and that feedback would not be provided.

The Ss in the one-button group (1) merely had to detect the presence of a signal and were to respond only on those trials. Group 2 was required to respond on each trial by depressing either the button marked "yes" or the one marked "no" to indicate the presence or absence of a signal (binary procedure). Group 3, as in the binary procedure, had two main category headings labeled "yes" and "no" and were required to respond after each trial. However, there were four buttons, two corresponding to the responses "Sure yes"

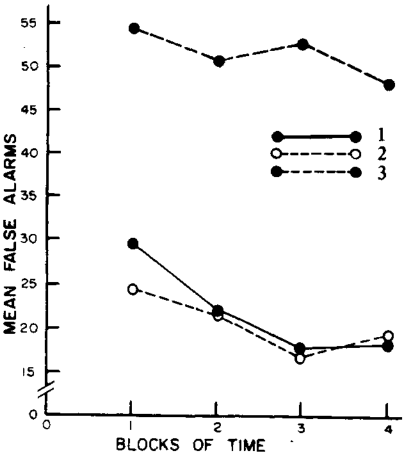

Fig. 1. Mean number of false alarms for each condition as a function of time on task.

*Financial support from the College of Arts and Sciences, The American University, to the senior author is gratefully acknowledged.

Table 1
Mean TSD Measures as a Function of Response Requirement and Time on Task

Response Require-ment	Blocks of Time 1	2	3	4
Group 1				
d′	2.60	2.32	2.41	2.52
β	3.84	8.40	8.86	8.81
Group 2				
d′	2.71	2.84	2.46	2.72
β	2.31	3.61	7.21	7.31
Group 3				
d′	2.21	2.17	1.96	2.05
β	0.59	0.83	1.13	1.02

and "Unsure yes" and the remaining two corresponding to the responses "Sure no" and "Unsure no" (rating procedure).

The main vigilance task lasted 80 min and was divided into four 20-min segments for programming and analysis purposes. Events were presented at a rate of one every 6 sec. Ten of the total 200 events per 20-min time block were signals, yielding a mean intersignal interval of 2 min. The signal schedule was obtained by randomly selecting from a rectangular distribution of intersignal intervals, with the restriction that no intersignal interval be more than 240 sec (twice the average interval) or less than 20 sec. This schedule was repeated four times throughout the task.

RESULTS

An analysis of variance carried out on the false alarm data obtained from each S (see Fig. 1) yielded a significant decrease in false alarms across the four time blocks ($F = 4.09$, $df = 3/99$, $p < .025$) and a significant difference for the three experimental groups ($F = 4.49$, $df = 2/33$, $p < .025$), with no interaction between these two variables ($F = .70$, $df = 6/99$). Individual comparisons of means, determined by the Newman-Keuls test (Weiner, 1962), revealed that Groups 1 and 2 were both different from Group 3 but not from each other.

Hit data were analyzed using nonparametric tests (Siegal, 1956) since Group 3 data were negatively skewed, presumably due to the low criterion adopted by these Ss (see below). A Friedman two-way analysis of variance was carried out separately for each experimental group to determine the effects of time on the number of hits (vigilance decrement). All groups showed a decline in hits as a function of time on tasks (see Fig. 2). However, the effect was significant only for Group 1 ($X_r^2 = 8.93$, $df = 3$, $p < .05$), with Group 2 approaching significance ($X_r^2 = 7.67$, $df = 3$, $p > .05$). A Kruskal-Wallis analysis of variance by ranks demonstrated that the experimental groups differed significantly as a function of response requirement ($H = 6.36$, $df = 2$, $p < .05$). Individual comparisons, however, revealed that the only significant difference was between Groups 1 and 3 (Mann-Whitney $U = 31$, $p < .02$, two-tailed).

The SDT measures, d′ and β, were derived from the hit and false alarm data for each S for each of the four time blocks (see Table 1). An interpolation procedure (to the next possible measurable value) was employed to obtain these measures whenever the percentage of hits or false alarms was 0 or 100 (see Loeb & Binford, 1968). An analysis of variance revealed that d′ remained constant over time ($F = 2.25$, $df = 2/33$). Also, the interaction mean square and the main effect for the experimental groups did not differ significantly from chance ($F = 1.52$, $df = 6/99$, and $F = 1.57$, $df = 3/99$, respectively).

The Kruskal-Wallis test indicated that β varied as a function of response requirement ($H = 10.15$, $df = 2$, $p < .01$). Two-tailed Mann-Whitney U tests showed that Groups 1 and 2 both differed from Group 3 at the .01 level but not from each other at the .05 level. β increased significantly as a function of time on task for all experimental conditions, as revealed by separate Friedman tests (Group 1, $X_r^2 = 11.75$, $p < .01$; Group 2, $X_r^2 = 8.93$, $p < .05$; Group 3, $X_r^2 = 9.08$, $p < .05$; $df = 3$ for each test).

DISCUSSION

The Ss who performed the task using the rating procedure (3) produced a considerably larger number of false alarms than either of the other groups. A similar result was obtained by Loeb & Binford (1964). The high rate of false alarms obtained here was probably due to the difficulty of the signal and the unusually low criterion adopted by Group 3 Ss. An analysis utilizing SDT indices suggested that Group 3 Ss did not differ in sensitivity (d′) but, rather, maintained a much lower criterion, β, than the other two groups.

The performance of the binary procedure group (2) and the standard vigilance group (1) was generally very similar, although only Group 1 had a significantly lower hit rate than Group 3. This latter finding can be understood by noting that the β value for Group 1 increased markedly from Block 1 to Block 2 (and stayed at that level for the remainder of the task), whereas Group 2 did not reach their high and stable level until Block 3.

The fact that manipulating the response requirement produced changes in Ss' willingness to report a signal, but did not affect the sensitivity parameter, d′, is consistent with a number of psychophysical studies reviewed above. It is important to note, however, that the event rate of one every 6 sec used here does not produce a very substantial demand on observing behavior, with the consequence that Ss are able to thoroughly attend to each stimulus event (Jerison, 1967). Since this condition approximated a psychophysical procedure, it is not surprising that similar results were obtained. It is quite possible that experiments employing faster event rates will alter the relationship between response requirements and vigilance performance obtained in this study.

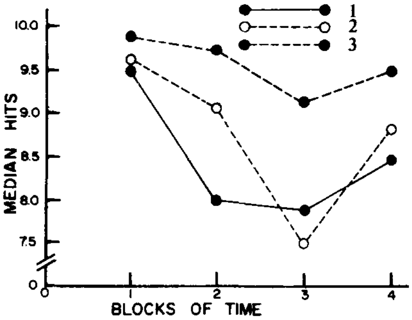

Fig. 2. Median number of hits for each condition as a function of time on task.

REFERENCES

BINFORD, J. R., & LOEB, M. Changes within and over repeated sessions in criterion and effective sensitivity in an auditory vigilance task. Journal of Experimental Psychology, 1966, 72, 339-345.

EGAN, J. P., SCHULMAN, A. I., & GREENBERG, G. Z. Operating characteristics determined by binary decisions and ratings. Journal of the Acoustical Society of America, 1959, 31, 768-773.

EMMERICH, D. S. ROCs obtained with two signal intensities presented in random order, and a comparison between yes-no and rating ROCs. Perception & Psychophysics, 1968, 3, 35-40.

GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics.* New York: Wiley, 1966.

JERISON, H. J. Activation and long-term performance. Acta Psychologica, 1967, 27, 373-389.

LOEB, M., & BINFORD, J. R. Vigilance for auditory intensity changes as a function of preliminary feedback and confidence level. Human Factors, 1964, 6, 445-458.

LOEB, M., & BINFORD, J. R. Variation in performance on auditory and visual monitoring tasks as a function of signal and stimulus frequencies. Sensory Research Laboratory Report, University of Louisville, Louisville, Kentucky, 1968.

MARKOWITZ, J., & SWETS, J. A. Factors affecting the slope of empirical ROC curves: Comparison of binary and rating responses. Perception & Psychophysics, 1967, 2, 91-100.

NACHMIAS, J. Effects of presentation probability and number of response alternatives on simple visual detection. Perception & Psychophysics, 1968, 3, 151-155.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

SWETS, J. A., TANNER, W. P., JR., & BIRDSALL, T. G. Decision processes in perception. Psychological Review, 1961, 68, 301-340.
WATSON, C. S., RILLING, M. E., & BOURBON, W. T. Receiver operating characteristics determined by a mechanical analogue to the rating scale. Journal of the Acoustical Society of America, 1964, 36, 283-288.
WINER, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

The isolation effect and verbal discrimination learning*

SHARRON S. WIKE and EDWARD L. WIKE
The University of Kansas, Lawrence, Kans. 66044

The present study was a 2 by 2 factorial design of verbal-discrimination learning with isolation vs no isolation of a pair of items as one factor and isolation vs no isolation of the correct item feedback as the other factor. The Ss were 36 undergraduates, 9 to a group. It was found that there were significantly fewer errors on the isolated item when the pair was isolated, but not when the feedback was isolated.

If an item is made different in some way from other items, as for example, by being printed in red, that item is learned faster or with fewer errors. This isolation effect, as it is called, has been demonstrated with serial learning, paired-associate learning, and free-recall learning. Stimulus isolation in paired-associate learning may have a greater effect than response isolation (Erickson, 1965).

In the verbal-discrimination task, the stimulus consists of a pair of items and the response consists of one of the items of the pair. Response learning is unimportant for verbal-discrimination learning because the correct response is given when the pair is presented.

The present experiment investigated the effects of isolation upon verbal-discrimination learning. The design of the experiment was a 2 by 2 factorial, with isolation vs no isolation of a pair of items as one factor and isolation vs no isolation of the correct item feedback as the other factor.

METHOD

The Ss were 38 volunteers from the introductory psychology course at the University of Kansas. Two Ss were lost—one due to misunderstood instructions and one due to an outside interruption while the experiment was in progress. There were 36 Ss, 14 males and 20 females, in the experimant proper. The Ss were assigned randomly and equally to four groups.

The apparatus consisted of a Sawyer Rotomatic slide projector programmed by three timers activated by a power supply. The material consisted of 50 pairs of low meaningful CVCs. The items had a rated meaningfulness of 1.46 to 2.95 (Noble, 1961). Two items were used in the instructions; the remaining items were combined to minimize intrapair similarity. The same list had been used in a previous study (Wike & Wike, 1970 Experiment 2). There were two interpair orders of the 24 pairs. The isolated pair was the 6th pair in the first order and it was the 15th pair in the second order. The order of the pairs was determined randomly. The item chosen to be isolated was selected because in the previous study it was one of the more difficult items. Since verbal-discrimination learning is relatively rapid and the probability of making an error by chance is only .50, one of the more difficult items was chosen. The pair that was isolated was zum and qal, with zum correct.

Ten trials were administered; the rate was 1:1. There were four groups: Group 1 had neither the stimulus pair nor the feedback item isolated; Group 2 had both the stimulus pair isolated and the feedback item isolated; Group 3 had only the stimulus pair isolated; and Group 4 had only the feedback item isolated. Isolation was accomplished by making the background of the slide red. The instructions were standard verbal-discrimination instructions in which S was directed to guess on the first trial. No mention in the instructions was made of a slide's being red.

Table 1
Mean Number of Errors for the Four Groups

Group	Isolation	Mean
I	None	2.64
II	P + F*	.64
III	P	.82
IV	F	2.73

P = pair; F = feedback

RESULTS AND DISCUSSION

The measure used was the number of errors made on the isolated pair on Trials 2-10 (the first trial was a guessing trial). The data were submitted to a 2 by 2 analysis of variance. The factors were (1) stimulus isolation and (2) feedback isolation. Only stimulus isolation was significant ($F = 10.88$, $df = 1/32$, $p < .005$), with isolation leading to fewer errors. The F for feedback isolation was below 1. The F for the interaction was also below 1. The means for the four groups are presented in Table 1.

Contrary to what has been normally found in other learning situations where isolation was manipulated, there seems to be no effect of isolation of the feedback term. Isolation usually is a fairly influential variable. It is somewhat unusual to fail to get an effect at all of isolation. Perhaps part of the isolation effect is to facilitate response learning. In the verbal-discrimination task, response-learning requirements are at a minimum.

A substantial isolation effect was observed when the isolation of the stimulus pair was manipulated. Thus, the isolation effect would appear to be a stimulus phenomenon, at least in verbal-discrimination learning. The finding that stimulus isolation is more important than response-term isolation agrees with the results of Erickson (1965) on paired-associate learning.

*These studies were supported by a grant, HD 00870, from the National Institute of Child Health and Human Development and a Biomedical Sciences Support grant, FR-07037.

REFERENCES

ERICKSON, R. L. Differential effects of stimulus and response isolation in paired associate learning. Journal of Experimental Psychology, 1965, 69, 317-323.
NOBLE, C. E. Measurements of association value (a), rated associations (a'), and scaled meaningfulness (m') for the 2100 CVC combinations of the English alphabet. Psychological Reports, 1961, 8, 487-521.
WIKE, S. S., & WIKE, E. L. Presentation rates and verbal discrimination learning. Psychonomic Science, 1970, 19, 253-254.