

Detection of auditory signals presented at random times¹

DAVID M. GREEN, UNIVERSITY OF CALIFORNIA, SAN DIEGO
R. DUNCAN LUCE, UNIVERSITY OF PENNSYLVANIA

One hundred msec tones of 1000 Hz at four intensities were presented according to two Poisson schedules in a background of wide band noise and as increments to a 1000 Hz tone. Each 15 minute test session was run under fixed conditions and one of three instructions to detect the signals; the two Ss were free to respond at any time. The data analyzed were several inter-response and signal-response distributions, and they were compared with a theory due to Luce. Discrepancies between the theory and data indicate needed modifications of both the theory and the procedure.

Most research on the detection of faint signals has involved specifying for the S well defined intervals during which a signal may be presented (for a summary, see Green & Swets, 1966; and Swets, 1964). This is patently unrealistic since in most practical detection situations signals occur without warning at irregular times. The two sources of data on these more realistic tasks come from studies of vigilance (Broadbent & Gregory, 1963; and the references given there) and studies of the method of free response (Egan, Greenberg, & Schulman, 1961; Watson & Nichols, 1966). The major difference between these two types of design is in the density of signal occurrences: They are exceedingly rare in vigilance experiments, and the focus of attention is on the deterioration of performance that results from long watches, whereas they are relatively frequent—every few seconds—in the other studies.

The present experiment falls in the latter class. It differs in two respects from what has been done before. First, the signals were presented truly at random, i.e., according to a Poisson schedule, because this leads to certain theoretical simplifications. Second, the analysis does not attempt to treat the situation as a series of fixed interval experiments, but rather accepts the signal presentations and the responses as a time series to be analyzed as such. In particular, we examine with some care certain observed inter-response time (IRT) and signal-response time (SRT) distributions. The analysis is guided by a theory that is, perhaps, the simplest one that could conceivably account for the data. As it has been described in detail elsewhere (Luce, 1966), we outline it only briefly here.

THEORY

The experimental paradigm is this: Brief, physically identical signals are presented according to some probabilistic temporal schedule that is unknown to S, and he is instructed to respond immediately whenever

he believes that a signal has been presented. He may respond whenever and as often as he wishes.

The theory postulates that at each instant S is in just one of two possible sensory states, D or \bar{D} , which are interpreted to mean "detect" and "not detect." Most of the time he is in \bar{D} , but every now and again he enters briefly into a D state as a result either of a signal presentation or of "noise" (i.e., a background of sensory activity provided by E, inadvertently generated by the apparatus, or inherent to S himself). The conditional probability that at time t he enters D when a signal is presented, $P_t(D|s)$, is assumed to be a parameter that depends on the sensory criterion of S and on physical properties of the signal and noise, but that is independent of t. We denote it by q. The noise induced D states are assumed to occur randomly in time independent of the presentation of signals and of S's responses, i.e., they form a Poisson process with some time independent intensity which we denote by ν . This parameter ν also depends on S's sensory criterion and on the physical properties of the signal and noise. As is well known, the Poisson assumption implies that $\nu e^{-\nu t}$ is the density of times between the successive D states that are due to the noise alone. The two parameters q and ν describe S's sensitivity to signals in noise in a fashion that is somewhat analogous to the hit and false alarm rates of ordinary Yes-No detection experiments.

The response process has two aspects: the conditions under which it is initiated and, given that it is, the time taken before it terminates in a response. The tendency for it to be initiated—the response bias—is assumed to be a probability that varies systematically in time; we denote it by $b(t)$. Specifically, we assume it varies as follows. If S is in the \bar{D} state at time t, then $b(t) = 0$. Also, if an incompleting response process is underway, then $b(t) = 0$. If, however, no response process is underway and if S enters the D state at time t (whether it is induced by a signal or by noise does not matter since he cannot distinguish among D states), then there is some probability b, independent of t, that S takes it seriously and initiates the response process. Once initiated, the process terminates in a response after some random time x. The random variable x has a density r that is assumed to be independent of the time t at which the process was initiated and that is also assumed to be zero for all times $x \geq \tau$, i.e.,

$$\int_0^{\tau} r(x) dx = 1, \quad (1)$$

where τ is some finite time. The latter assumption is stronger than the one made in Luce (1966) about the tail of r .

Aside from the strong assumptions that q , ν , b , and r are all independent of time, the most significant assumption is that $b(t) = 0$ whenever a response process is underway. This means that following each response S acts as if he were no longer aware of any past events. Thus, once a response occurs we, as theorists, can act as if the whole process describing S begins anew at this point; all past history prior to this response can be ignored. Therefore, we know that all events leading to the next response must have followed the given response. This makes the analysis relatively easy because Poisson processes have the happy, and characteristic, property that the distribution of times from an arbitrary time (e.g., the time of occurrence of a response) to the next Poisson event is independent of time and is, in fact, identical to the inter-event distribution. Thus, the distribution to the next noise induced D state is $\nu e^{-\nu t}$. If we elect to present the signals according to an independent Poisson process with intensity λ and if each signal has a fixed probability q of inducing a D state, then these D states form a Poisson process with intensity $q\lambda$. Thus, the distribution of the next signal induced D state following a response is $q\lambda e^{-q\lambda t}$. Moreover, since each type of D state is governed by a Poisson process and since they are independent, the overall generation of D states is Poisson with intensity $\eta = \nu + q\lambda$. Thus, if the model is correct and if we present the signals according to a Poisson process, we can treat the system at each response as totally without a past.

Luce (1966) derived some properties of this model, including the form of some IRT and SRT distributions and the equations for some of the parameters in terms

of moments of these observable distributions. Experience in analyzing the present experiment has shown that the latter are of little practical value in estimating the parameters and that far more can be learned by examining carefully (parts of) several distributions. In particular, suppose that Eq. (1) is true for some finite τ , and define the constants

$$R_\theta = \int_0^\tau e^{\theta t} r(t) dt, \quad (2)$$

where θ will take on the values $b\eta$, $b\nu$, and $-\lambda$. The results shown in Table 1 can be derived (see the Appendix). The main feature of this table is that for times $t \geq \tau$ each distribution is exponential and is independent of r , whereas for times from 0 to τ , where the unknown response time distribution plays a significant role in the form of the observable distribution, the equation for the probability of a response is simple. The tails of all four observable distributions are exponential, with one IRT and one SRT distribution each having the time constant $b\eta = b\nu + bq\lambda$ and with the other two having the time constant $b\nu + \lambda$. Since λ is known or is easily estimated from the observed pattern of signal presentations, these estimates yield immediate estimates for bq and $b\nu$. Thus, the only unknowns in the right hand probabilities are the three constants $R_{b\eta}$, $R_{b\nu}$, and $R_{-\lambda}$. So, for example, for each experimental condition we may use the $(0, \tau)$ data of Distribution 2 to estimate $R_{b\nu}/R_{-\lambda}$ and then use that estimate and our estimates of bq , $b\nu$, and λ to make a parameter-free prediction of the probability of a response in $(0, \tau)$ for Distribution 4.

The prediction of exponential tails accords with the data of Lucas (1966).

METHOD

Two female Ss were tested in a number of conditions on two auditory tasks. In one task, called the "sine-wave" experiment, the signal was a 0.10 sec increment in a continuous background tone of 1000 Hz at 70 dB SPL. To avoid transients at onset and offset, the signal was passed through a 50 Hz bandpass filter centered at 1000 Hz. In the second, called the "noise" experiment, the signal was a 0.10 sec 1000 Hz tone added to a continuous, wideband noise with a constant spectrum level of 15 dB. The bandwidth of the noise was essentially that of the earphones. The S, who was seated in a double walled, Industrial Acoustics sound attenuating chamber, listened binaurally and in-phase to the sounds presented by two matched PDR-8 earphones.

Four signal levels were selected so that S achieved, approximately, 60%, 75%, 90%, and 100% correct detections in a two alternative, forced choice test. During these preliminary tests, possible occurrences of the signal were marked by two lights about 1 sec apart. For the sine wave experiment, the values of the ratio of the increment in the background pressure to the background pressure, $\Delta p/p$, were 0.020, 0.034, 0.050,

Table 1. Deductions from the model.

The four distributions are of the following random variables:

1. Inter-response time t .
2. Inter-response time t conditional on no signal occurring between the two responses.
3. The time t from a signal until the next response conditional on no observable event (either a signal or response) during the τ sec prior to the signal. (Note: other signals may occur between the given signal and the response.)
4. The time t from a signal until the next response conditional on no observable event (either a signal or a response) during the τ sec prior to the given signal and an no observable event (a signal) from the given signal until the response.

Distribution for $t \geq \tau$	Probability of Response in $(0, \tau)$
1. $R_{b\eta} b\eta e^{-b\eta t}$	$1 - R_{b\eta} e^{-b\eta \tau}$
2. $\frac{R_{b\nu}}{R_{-\lambda}} (b\nu + \lambda) e^{-(b\nu + \lambda)t}$	$1 - \frac{R_{b\nu}}{R_{-\lambda}} e^{-(b\nu + \lambda)\tau}$
3. $\frac{R_{b\eta}}{R_{b\nu}} (1 - bq) b\eta e^{-b\eta t}$	$1 - \frac{R_{b\eta}}{R_{b\nu}} (1 - bq) e^{-b\eta \tau}$
4. $\frac{1 - bq}{1 + \lambda \frac{bq}{b\nu} \frac{R_{-\lambda}}{R_{b\nu}}} (b\nu + \lambda) e^{-(b\nu + \lambda)t}$	$1 - \frac{(1 - bq) e^{-(b\nu + \lambda)\tau}}{1 + \lambda \frac{bq}{b\nu} \frac{R_{-\lambda}}{R_{b\nu}}}$

Table 2. Ratio of number of responses to number of signals

	λ	Subject SW Signals				Subject RT Signals			
		60	75	90	100	60	75	90	100
Noise	.09	0.62	1.00	0.75	1.08	0.42	0.67	0.78	1.05
	.09	1.51	0.96	1.73	1.05	0.42	0.62	0.50	1.06
	.20	1.05	0.93	1.40	1.07	0.45	0.63	0.60	1.02
Liberal	.20	2.15				0.76			
Conservative	.20	0.70				0.26			
Sine Wave	.09	1.61	1.40	1.20	1.03	0.75	1.09	0.95	1.15
	.09	0.58	0.95	2.13	1.02	0.55	0.83	0.75	1.11
	.20	1.09	1.23	0.85	1.05	0.69	0.66	0.78	1.13
Liberal	.20	2.16				0.94			
Conservative	.20	0.46				0.26			

and 1.778, respectively, for 60%, 75%, 90%, and 100% detection. These values varied by ± 0.001 for each S. For the noise experiment, the values of $10 \log_{10} E/N_0$, where E is signal energy and N_0 is noise power density, were 9.0, 11.0, 13.0, and 41.0 dB, respectively, for the four levels. These values varied by ± 0.5 dB for each S.

The occurrence of a signal was determined by a pulse of a Geiger counter that monitored a small radioactive source. The distribution of times between presentations was, therefore, exponential with an intensity λ that depended upon the distance between the source and the counter. This dependence was determined empirically, and the distribution of inter-event times was verified to be exponential by a χ^2 test. Two signal densities were used: $\lambda = 0.20$, or about one signal every 5 sec on the average, and $\lambda = 0.09$, or about one signal every 11 sec on the average. The latter values were replicated on different days to test the stability of the behavior.

The response key was spring loaded and required 5 oz of dead weight to record a response. It exhibited hysteresis to the extent that it had to be raised 1/8 in. before a second response could be recorded. Inadvertent double pressing was virtually impossible. Delay multivibrators locked the key out for about 70 msec after depression, thereby avoiding the recording of key chatter. The maximum rate at which responses could be generated during brief intervals of from 1 to 2 sec was about 8 per sec.

The occurrence of either a signal or a response was timed from the preceding event by means of a decimal counter that was driven at a rate of 1000 pulses per sec. These times were punched onto a paper tape. Since each recording of a time required about 70 msec and since a signal might follow a response by as little as 1 msec, two clocks were used so that one was timing while the other was recording. If three or more events occurred within 70 msec, those beyond the second event were ignored. The various temporal distributions used in the analysis were calculated from the times recorded on the tape.

In addition to the two tasks, sine wave and noise, the four signal levels, and the two densities of signal

presentations, the instructions given to S were varied. For most conditions, S was simply told to push the key each time she thought she heard a signal. No restriction was imposed upon the number of responses that could be made. In addition, with $\lambda = 0.20$ and at the 75% signal level, two modifications of the basic instructions were used. In the "liberal" condition, S was told to respond whenever she heard any event resembling a signal, and in the "conservative" condition, she was asked to respond only when she was certain that a signal had occurred. During the experiment, each S on at least two occasions asked what she should do when a second signal occurred before she had responded to a prior signal. We reiterated the original instruction, "Respond each time you hear the signal." As we shall see, this was probably an error on our part.

Table 2 in the Results section summarizes the four manipulations.

The Ss were run for about 2 h per day in just one experimental condition. Each daily session was divided into four subsessions of approximately 15 min, between which there were generous rest periods. The Ss were paid at an hourly rate.

RESULTS

Response Criterion

When a signal interval is specified it is commonly acknowledged that S must establish a criterion and that he can choose it over such a wide range that, under fixed stimulus conditions, the conditional response probabilities can be varied from 0 to 1, thereby generating what is known as the ROC curve. It is also known that such a criterion, or bias, can be systematically affected by payoffs, presentation probabilities, and instructions. When the criterion is not being studied, it has become usual to stabilize it by using information feedback after each trial, or short block of trials, and by payoffs. Such stabilizing techniques were not attempted in the present experiment for fear of disrupting the ongoing process, and so we relied simply on the instructions to maintain a constant criterion. Our first question, then, is: How well did this work? Since relaxing the criterion increases the tendency to respond both to signals and to noise, an obvious model-free measure of response bias is the ratio of the total number of responses to the total number of stimuli presented. These ratios are shown in Table 2.

Table 3. Comparison of within day and across day average σ/μ ratios for the number of responses per signal in the replicated $\lambda = 0.09$ conditions at signal levels of 60%, 75%, and 90%.

	Subject			
	SW		RT	
	Noise	Sine Wave	Noise	Sine Wave
Within days	0.130	0.237	0.272	0.183
Across days	0.360	0.369	0.232	0.206

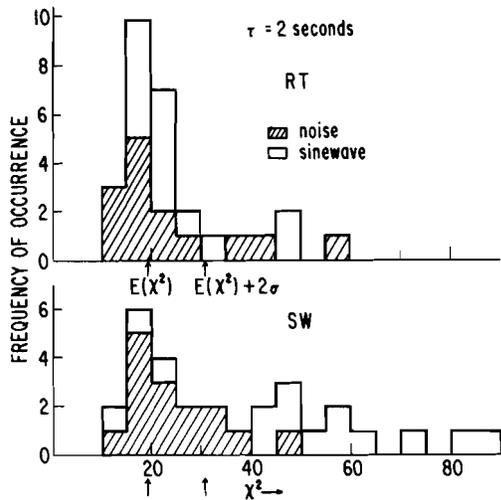


Fig. 1. Histograms of χ^2 for the maximum likelihood fit of an exponential to the empirical distributions 1, 2, and 4 (Table 1) of times ≥ 2 sec. The expected value and the upper 2σ points are indicated.

Three results stand out. First, the ratio does not seem to change systematically with signal strength except that it always ends up in the neighborhood of 1.05 for the perfectly detectable signals. Second, when a condition is replicated on different days (the two $\lambda = 0.09$ rows in each subtable), the ratios differ by factors as large as 3. This is an enormous effect, and it makes clear that our instructions failed to achieve the stability we would have liked. Third, changing the instructions does have a pronounced effect (a factor of from 3 to 5) in the expected direction.

Because the criterion was unstable over replications, we must admit the possibility that there were comparable fluctuations within the daily sessions. To estimate its magnitude, we computed the response/signal ratio for each of the four, 15 min subsessions. The standard deviation of this ratio, averaged over both Ss in all conditions, is about 0.10 and since the ratio itself is about 1, the σ/μ ratio is about 0.10. To compare this more precisely with the variability across days, we examine the $\lambda = 0.09$ conditions with 60%, 75%, and 90% signal levels, which are shown in Table 3. The within day variability, although it is somewhat less than the across day variability, especially for SW, is still considerably greater than one would like. The lack of control of the criterion is unsatisfactory and is a major flaw in our experimental procedure.

Exponential Tails

Evidence from other experiments (Egan et al, 1961; Watson & Nichols, 1966) as well as internal evidence from this experiment suggests that we will not err seriously if we assume that all response times are less than 2 sec. Although by the standard of a reaction time experiment, response times beyond 0.5 sec are excessively long, the fact that S was focused upon

accuracy of detection and was not pressed for rapid responses suggests that 2 sec may not be unduly conservative. Using $\tau = 2$, we calculated maximum likelihood estimates ($1/\text{mean time}$) of exponential time constants for the tails of Distributions 1, 2, and 4 (Table 1) in each experimental condition. For each theoretical distribution, the time axis was subdivided so as to produce 20 intervals of theoretically equal (0.05) probability, and the corresponding χ^2 s were computed. The resulting empirical distributions of χ^2 , along with the expected value and the upper 2σ point, are shown in Fig. 1. Somewhat too many χ^2 s exceed the 2σ value; this is especially true for SW.

Some insight into the discrepant cases is gained by examining the hazard function, i.e., the conditional density, $h(t)$, that a response occurs t sec after some prescribed event given that it has not already occurred:

$$h(t) = \frac{f(t)}{1 - \int_0^t f(x) dx} \quad (3)$$

where f is the ordinary density function. Conversely, f and h are also related by

$$f(t) = h(t) e^{-\int_0^t h(x) dx} \quad (4)$$

Observe that $f(t) = \alpha e^{-\alpha t}$ if and only if $h(t) = \alpha$, i.e., the only density with a constant hazard function is the exponential one.

Using the 10 intervals that subdivide the best fitting exponential distributions into regions of equal probability (0.1), the empirical hazard functions were estimated in the obvious way for all IRTs greater than 2 sec. Those for the liberal and conservative conditions are shown in Figs. 2 and 3. For RT (Fig. 2) they are

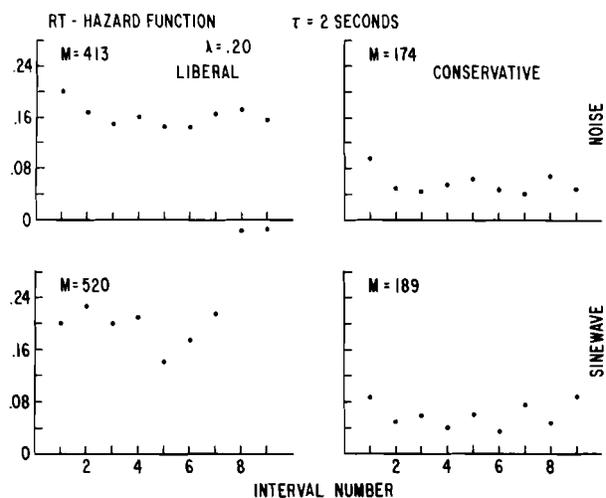


Fig. 2. Empirical hazard functions (see text) for liberal and conservative instructions with noise and sine wave backgrounds. Subject RT.

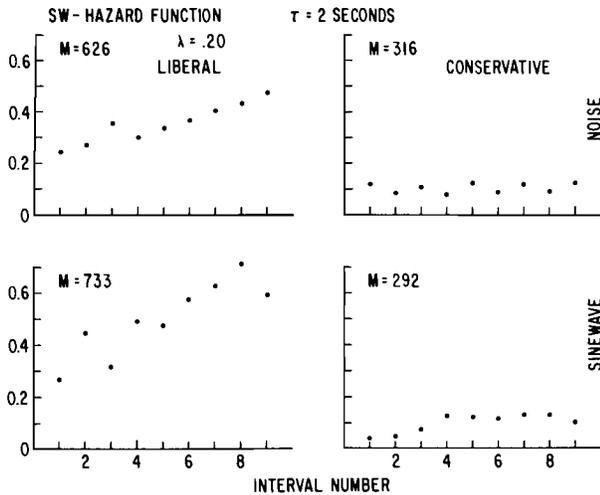


Fig. 3. See Fig. 2. Subject SW.

all approximately constant, as we might expect from Fig. 1; the only effect of liberalizing instructions was to raise the level above that for the conservative instructions by a factor of from 2 to 3. For SW (Fig. 3) the hazard functions show a different pattern. Under the conservative instructions they are approximately constant; however, the liberal instructions not only raised the initial level by a factor of from 2 to 3, but they also rotated it to a positive slope. That is to say, not only did SW have a greater tendency to respond under the liberal instructions, but that tendency increased with increasing time since the preceding response.

If a lax response criterion is the main reason why the tails of SW's distributions are not exponential, then we might expect to find a moderate positive correlation between her χ^2 values and the ratios in Table 2. Little or none exists, which suggests that other factors

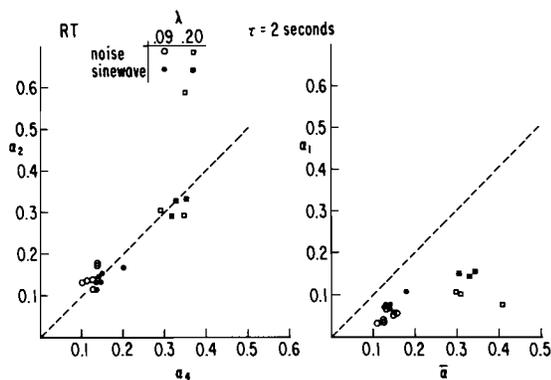


Fig. 4. Comparison of estimated exponential time constants for distributions 2 and 4 and for 1 and the weighted average of 2 and 4 (Table 1). If the theory is true and in the absence of variability, the α_2 vs α_4 points should fall on the dotted line and the α_1 vs $\bar{\alpha}$ points should lie below the dotted line. Subject RT.

must be at work. Indeed, if we average her hazard functions over signal levels, we find that the slope is greater for the sinewave than for the noise background at both rates of presentation, and that it is greater for the rapid than for the slow presentation rate under both background conditions. In fact, the curves for the fast rate under the standard instructions are very similar to those under the liberal instructions (and fast rate).

Consistency of Parameters

Although some of the distributions do not have exponential tails, let us proceed as if all were well. According to Table 1, the time constants for Distributions 2 and 4 should be identical, even though 2 is an IRT and 4 is an SRT distribution. (A similar observation is true for Distributions 1 and 3, but computations for the latter distribution were not made.) Moreover, if the model is correct, the time constant for Distribution 1 should be strictly less than those for 2 and 4 since $bq < 1$ and so $bv + bq\lambda < bv + \lambda$. Let

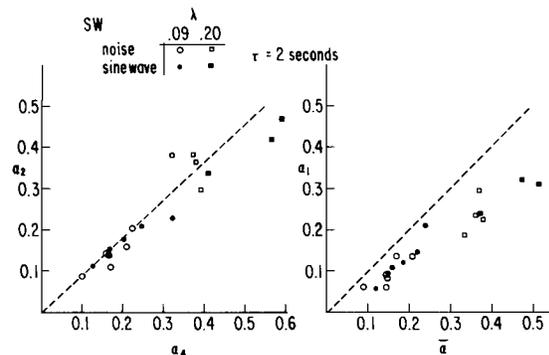


Fig. 5. See Fig. 4. Subject SW.

α_1 denote the estimate of the time constant for Distribution 1 and let $\bar{\alpha}$ denote the weighted average of α_2 and α_4 . Figures 4 and 5 show the comparisons of α_2 with α_4 and of α_1 with $\bar{\alpha}$. Neither S seems to exhibit a systematic departure from the hypothesis that $\alpha_2 = \alpha_4$, although, as would be expected, SW's plot is the more variable. The hypothesis $\alpha_1 < \bar{\alpha}$ is confirmed without exception.

Detectability

Since we really are not sure how signals close to other signals or close to responses cause S to react, perhaps the safest crude measure of detectability (but ignoring false alarm rate) is the proportion of responses that occur within $\tau (=2)$ sec after a signal (conditional on there being no other observable event signal or response) from τ sec prior to the signal until the response (this is Distribution 4 of Table 1). These proportions are shown in Table 4. It is evident that this measure correlates well with the detectability of the signal as measured by the forced choice pro-

Table 4. Proportion of responses within 2 sec of a signal given that no other observable event occurred from 2 sec prior to that signal until the response.

	λ	Subject SW				Subject RT			
		60	75	90	100	60	75	90	100
Noise	.09	0.47	0.60	0.90	1.00	0.35	0.44	0.80	1.00
	.09	0.44	0.48	0.68	1.00	0.43	0.37	0.67	1.00
	.20	0.71	0.78	0.91	1.00	0.61	0.73	0.94	1.00
Liberal	.20					0.82			
Conservative	.20					0.63			
Sine Wave	.09	0.49	0.39	0.62	1.00	0.31	0.42	0.55	1.00
	.09	0.38	0.50	0.75	1.00	0.17	0.41	0.53	1.00
	.20	0.68	0.70	0.77	1.00	0.50	0.61	0.72	1.00
Liberal	.20					0.82			
Conservative	.20					0.54			

cedure, that it is uniformly higher (and roughly the same as percent correct) for the faster presentation rate, and that differences in this measure for replications do not correlate well with the corresponding differences in the ratio of the number of responses to the number of stimuli given in Table 2.

As we noted earlier, we may make a parameter-free prediction of these proportions by estimating b_v and b_q from the tails of Distributions 1, 2, and 4, λ from the signal presentation distribution, and $R_{Dv}/R_{-\lambda}$ from the 0 to τ portion of Distribution 2. These predictions are shown in Fig. 6, and clearly they are poor. Aside from too much scatter, we predict, especially for RT, too few fast responses.

To see in another way that there are more fast responses than is conceivably consistent with the theory, consider the probability of an IRT less than τ sec (Distribution 1 of Table 1). Since $\eta \geq 0$, Eqs. (1) and (2) imply that $R_{Dv} \geq 1$, hence the probability of an IRT less than τ cannot possibly exceed $1 - e^{-b\eta\tau}$. With $\tau=2$ and the estimate α_1 of $b\eta$ from the tail of the IRT distribution, we obtain the comparison shown in Fig. 7. All of the points should be below the diagonal, yet less than half of them are.

Very Short IRTs

These last difficulties suggest that we had better examine the histograms of short IRTs conditional on no intervening signal. For times less than 1 sec, these are shown in Figs. 8 and 9 for the liberal and conservative conditions. The data from the other experimental conditions are similar, including the existence of a second mode at about 0.4 sec in SW's data. A sizable fraction of these IRTs lie in the interval from 100 to 150 msec, i.e., in the typical range of simple reaction times. We have verified that these doublets are real responses, not relay chatter or other recording difficulties. They are not due to relay chatter because the key was locked for a period of about 70 msec after a response. They are not due to hand tremor because of the hysteresis of the key; it had to be released before another depres-

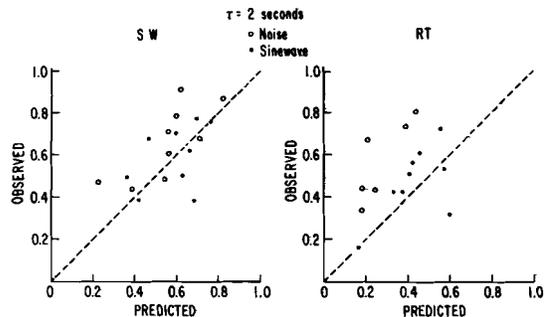


Fig. 6. Observed vs predicted values of the probability of a response within 2 sec of a signal given that neither another signal nor response occurred from 2 sec prior to the signal until the response in question.

sion would activate the switch. It is extremely doubtful that they are due to rebound of the key and hand as a single mechanical system, although the key spring coupled with certain dead weights would generate a decaying oscillation having a period of about 100 to 150 msec when excited by an impulse.

We feel that the most likely explanation is that S did detect signals during the course of response processes, and she responded to them. The argument depends, first, upon the probability of occurrence during an incompleting response process of an event that is capable of inducing another response, and second, upon what S does about it.

Consider, first, the probability. Judging by the reaction time distributions (discussed in the next section), the mean latency is sizable, of the order of .5 sec. Since the effective detections are Poisson with intensity $b\eta = b_v + b_q\lambda$, we may use the parameter α_1 to estimate $1 - e^{-b\eta/2}$, which is the probability of at least one event within .5 sec. Since α_1 lies

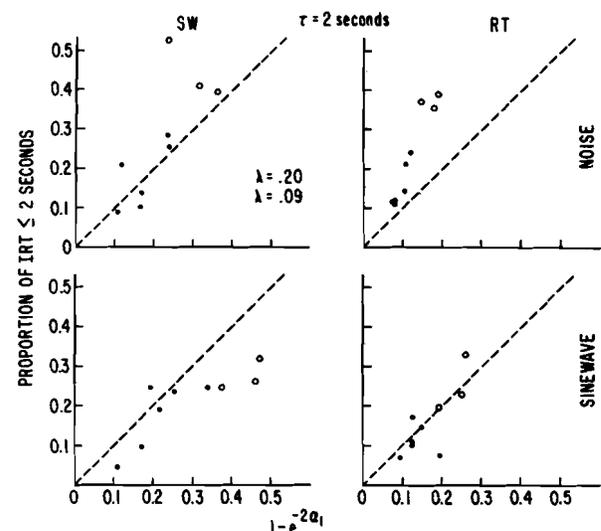


Fig. 7. The proportion of IRT ≤ 2 sec vs $1 - e^{-2\alpha_1}$. If the theory is true and in the absence of variability, the points should all lie below the dotted line.

Table 5. Mean reaction time to a signal of those responses within 1 sec of that signal given that no other observable event occurred from 1 sec prior to that signal until the response. The data are averaged over presentation rates.

		Signal			
		60	75	90	100
RT	Noise	0.60	0.60	0.59	0.36
	Sine Wave	0.52	0.58	0.63	0.37
SW	Noise	0.50	0.52	0.50	0.32
	Sine Wave	0.55	0.56	0.56	0.32

between 0.03 and 0.50, depending upon the condition, the probability of an effective second detection lies somewhere between 0.015 and 0.22. This is sufficiently large to matter, at least in some conditions.

Second, what did S do? That is more difficult to say; however, as we mentioned, when they asked what they should do, we told them to "respond each time you hear the signal." Presumably, this meant that they responded twice whenever they heard a second signal before they had responded to the first. If so, the behavior is clearly incompatible with the postulate of the theory that $b(t)=0$ when a response process is underway. Our only defense for our instructions is that we assumed at that time that the response latencies were of the order of 1/10 sec, in which case only about 1% of the responses would have been affected. The large proportion of double responses was not discovered until later when detailed analyses were performed.

Reaction Times

For the same reasons that we used Distribution 4 to get a measure of detectability, and now with the added reason that we know of the existence of doublet responses, we again use that distribution to examine reaction times. Individual histograms suggest that we may average over the presentation rates and use the mean over 0 to 1 sec as a suitable measure of reaction time. These means are shown in Table 5. Two facts

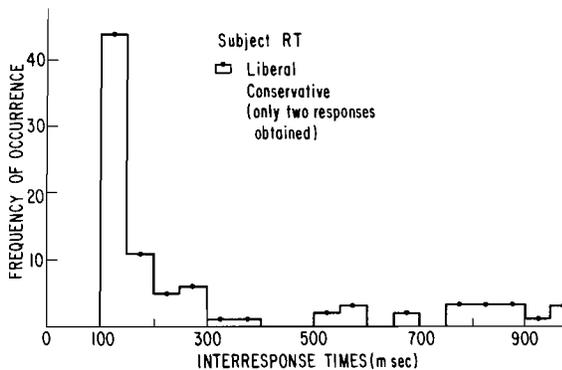


Fig. 8. Histogram of IRT's < 1 sec conditional on no observable event preceding the first response by more than 1 sec for the liberal and conservative instructions. Subject RT.

emerge. First, the mean reaction time to the imperfectly detectable signals is independent of their level of detectability. Second, the reaction time to the 100% detectable signal is about 40% faster than to the other signals. Since we do not know how to analyze the theoretical distributions for this region, we do not know if this difference is compatible with the assumption of a response time distribution r that is independent of signal level. Intuitively, it seems unlikely.

DISCUSSION AND SUMMARY

Perhaps our most important conclusion is methodological: Detection experiments in which the signals occur at random times and S is free to respond at any time can be usefully analyzed in terms of several observable IRT and SRT distributions. This approach, and the motivation behind it, differ somewhat from earlier work on the problem. In their pioneering work on what they called the free-response task, Egan et al (1961) attempted to estimate quantities analogous to the hit and false alarm rates of the fixed interval, Yes-No task. To do this, they simply partitioned time into equal intervals and treated each interval as such a task. Although the choice of these intervals is completely arbitrary, Watson and Nichols (1966) established that within reasonable changes in their lengths, the results are relatively invariant. Both studies demonstrated that the false response and detection rates covary, much as they do in the fixed interval task, and that appreciable changes in performance are exhibited without any appreciable change in the apparent sensory sensitivity of S. The pertinence of these analyses to the evaluation of performance in vigilance and similar realistic detection situations is clear.

The fundamental difference between their approach and ours is that they used SRT distributions to estimate two parameters—the hit and false alarm rates—whereas we treated signals and responses simply as temporal events and we attempted to study the entire distribution of certain classes of inter-event intervals. The only

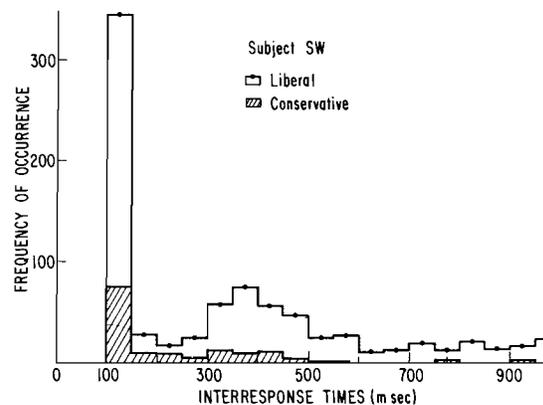


Fig. 9. See Fig. 8. Subject SW.

partition that we found useful—nay, essential—was into those longer times whose distribution (aside from multiplicative constants) could not conceivably have been affected by the unknown response time distribution and into the shorter times whose distribution most certainly was affected by it.

The longer times—the tails of the IRT and SRT distributions—were tested against an exponential hypothesis by χ^2 and were studied in greater detail by examining their corresponding hazard functions, i.e., the conditional density of the event given that it has not already occurred. It is our impression that this is the most illuminating way to discover the sorts of process that generate the tails. The shorter times were studied in three ways, none of which is wholly satisfactory. We examined the proportion they form of all times of a given class, their means, and their histograms. Once we have evolved a more satisfactory experimental procedure and have developed a model that has a better chance of being approximately correct, then it will surely be desirable to derive and work with explicit expressions for the distributions of short times.

The analysis uncovered various difficulties, at least one of which is experimental. The others may be also, although we are inclined to think that they simply reflect adversely on the model. The fact that, when a condition is replicated, the response/signal ratio varied by a factor as large as 3 makes it clear that our instructions failed to induce a stable response criterion. This is sufficient to render uncertain other conclusions; however, the relatively low within day variability reinstates some confidence in them.

There is no doubt that the experiment should be redesigned so as to increase the stability of the response criterion without, however, destroying the general spirit of the current design. One could, of course, establish an experimental criterion, such as 1 sec, and define a "hit" to be any response that follows a signal in less time than the criterion. Probably this should be tried in spite of the uncomfortable arbitrariness of the criterion and of the possibility, which we fear, that the very act of feedback (presumably in another modality) may disrupt S's attention. Another possibility is to mimic more accurately those real life detection situations, such as detecting obstacles on a highway at night, in which the signal level increases steadily until finally (and, occasionally, fatally) it becomes perfectly detectable and so serves as its own information feedback. Among the possible payoff schemes, the most natural is to reward responses as a diminishing function of time from the onset of the signal and to punish those that occur when no signal is present. In this way, we should be able both to manipulate and to stabilize S's response criterion. The major drawback to this idea is that it is not clear how to relate such data to

those obtained using fixed signal levels in well defined signal intervals.

A second major difficulty was the occurrence of responses separated by little more than S's simple reaction time. Although such doublets may have been caused by some mechanical artifact, we doubt it; rather, we suspect that the second responses were due to detections that occurred while the first response process was underway. If so, a major assumption of the model, namely, that $b(t)=0$ during a response process, is wrong, which is unfortunate since that assumption permitted a very simple analysis of the temporal process. Either it must be relaxed to accord with the facts, which will complicate greatly the mathematical analysis, or the experiment will have to be modified to make the assumption valid. One might simply instruct S to disregard any apparent signals if they occur while a response process is underway. Whether S can do so and whether we, as experimenters, can verify that he has done so is not easy to say. Another possibility would be to modify the distribution of signals from the Poisson so as to avoid short interpresentation times, but this would make the theoretical analysis vastly more complicated. Moreover, it would not prevent the occurrence of noise induced D states near signal occurrences and, hence, does not completely alleviate the problem.

A third difficulty for the theory is the fact that one S (SW) had, in some conditions, rising hazard functions. The most obvious way to account for this is to assume that $b(t)$ increases with time since the preceding response until the next response occurs, at which point it is reset to some base level. Another possibility is that $b(t)$ remains constant but that the definition of the states D and \bar{D} changes in such a way that both q and ν increase over time. This possibility raises the whole issue of whether a two state model is appropriate in the first place. There is ample evidence, both formal and informal, that there are many more than two sensory states—the growth of loudness with intensity makes it abundantly clear. Indeed, for the sake of the argument, let us suppose that the sensory states form a continuum. Then changes in the sensory criterion that defines what shall be classed as a detection will alter q and ν simultaneously and continuously. This corresponds to the generation of a bowed ROC curve in the fixed interval experiment. The exact details of these changes depend upon the assumed distribution of sensory effects over the set of states. As a matter of fact, however, for any reasonable assumptions, the effect of changing sensory criterion can be approximated by assuming that q and ν are fixed and that b is changing. Thus, for a while, it is probably simplest to study how changes in b affect the hazard function. Eventually, of course, once some of the current difficulties are overcome, it may be desirable to study if, and how, q and ν vary with instructions and payoffs

to the Ss. But since agreement as to the exact nature of these changes is yet to be reached in the fixed interval task, it does not strike us as the most promising initial line of research in the less well understood procedure using random signal presentations.

References

- Broadbent, D. E., & Gregory, Margaret. Vigilance considered as a statistical decision. *Brit. J. Psychol.*, 1963, 54, 309-323.
- Egan, J. P., Greenberg, G. Z., & Schulman, A. I. Operating characteristics, signal detectability, and the method of free response. *J. Acoust. Soc. Amer.*, 1961, 33, 993-1007.
- Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics*. New York: John Wiley & Sons, 1966.
- Lucas, P. A. Human performance in low signal probability tasks. Technical Report, Sensory Intelligence Laboratory, University of Michigan, 1966.
- Luce, R. D. A model for detection in temporally unstructured experiments with a Poisson distribution of signal presentations. *J. math. Psychol.*, 1966, 3, 48-64.
- Swets, J. A. (Ed.) *Signal detection and recognition by human observers*. New York: John Wiley & Sons, 1964.
- Watson, C. S., & Nichols, T. L. Replication and revisions of Egan's method of free response. (Abstract) *J. Acoust. Soc. Amer.*, 1966, 36, 1247.

Note

1. The bulk of this work was supported by National Science Foundation Grant GB-3689 to the University of Pennsylvania. Computations were performed at the Computer Centers of both the University of Pennsylvania and the University of California, San Diego. During the final stages of writing, the second author was an NSF Senior Postdoctoral Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, California. We are deeply indebted to Miss Virginia L. Maier who ran the Ss and who prepared the several computer programs needed to analyze the data. We also thank Dr. Joseph Markowitz, Dr. G. Bruce Henning and Mr. David Weiss for their assistance in the design, construction and maintenance of some of the test apparatus.

(Accepted for publication June 12, 1967.)