Memory & Cognition
1974, Vol. 2 (4), 758-770

Stimulus intensity, catch trial effects, and the speed-accuracy tradeoff in reaction time: A variable criterion theory interptation*

G. ROBERT GRICE†, REED L. HUNT, BRUCE A. KUSHNER, and CHARLES MORROW
University of New Mexico, Albuquerque, New Mexico 87131

Ss responded to a 1,000-Hz tone of 50, 80, or 100 dB. Catch trial conditions were none, blank trials, a red light, a noise, and an 1,800-Hz tone. Auditory catch signals were of the same intensities. RT distributions in the first three conditions were well described by a family of exponential growth functions dependent upon stimulus intensity and by the parameters of normal criterion distributions dependent upon catch trial conditions and between-session variability. Performance in the auditory catch trial conditions was not dependent upon the same set of sensory growth functions. Performance in these conditions was described by a two-dimensional analysis of information transmitted as a function of time and interpreted in terms of variable criterion theory. The speed-accuracy tradeoff in this situation appears to depend upon differential rates of growth of intensity and associative information and the criterion used in responding to this information.

In the context of a decision theory approach to reaction time (RT), catch trial effects may be viewed as resulting from variation in the decision criterion. Grice (1972b) has supported this interpretation in a quantitative analysis of an experiment by LaBerge (1971). Grice's theory of response evocation assumes that, following stimulus onset, sensory information (V) grows in strength according to some continuous function, the rate of growth depending upon stimulus intensity. In conditioning, and presumably in choice situations, there is also an associative information component which grows in a somewhat slower fashion (Grice, 1972a). When this sensory (and, when applicable, associative) growth reaches the level of the S's decision criterion or reaction threshold (T), the response occurs. Thus, in RT the latency is determined by the rate of growth of V and by the level of the criterion. The criterion is assumed to depend upon the requirements of the task; set, attentional, and motivational factors; and individual differences. Under homogeneous conditions of performance, it is assumed to be a normally distributed random variable with mean, T, and standard deviation, σ. Both of these parameters are influenced by experimental conditions and individual differences. The value determining the probability of response at a particular latency is $E = V - T$. (In conditioning, it is $E = H + V - T$, where H is the associative component.) E corresponds to Spence's (1956) concept of suprathreshold excitatory strength or reaction potential. Since T is normally distributed, so is E, and scaling procedures based on the normal model apply.

*This research was supported by PHS Grant MH16400 from the National Institute of Mental Health.
†Requests for reprints should be sent to G. Robert Grice, Department of Psychology, University of New Mexico, Albuquerque, New Mexico 87131.

One application of the analytic procedures associated with this model is to the cumulative RT distribution. Any point on a cumulative latency distribution may be regarded as a measure of the average amount of information processed up to that time. This, in turn, depends upon the rate at which the information becomes available and upon the parameters of the criterion distribution applied to that information. In applying the scaling procedure, the normal deviate corresponding to the probability value at any time is a measure of the information gained to that time, measured from the mean of the criterion distribution on a scale with σ as the unit. Thus, if the entire distribution is converted to normal deviates, a function is obtained describing the growth of V, measured from T as the origin, in units of σ. Such functions are negatively accelerated in form and can generally be fitted by exponential growth functions.

In research such as that discussed here, in which the criterion parameters are manipulated, additional scaling procedures are necessary to evaluate the differences in T and σ. This task is accomplished by means of a response evocation characteristic (REC) (Grice, 1971, 1972a). In the present application of an REC, the normal deviate transformation of the cumulative distribution for one condition or S is plotted against that of another, the two elements of each point representing the same time interval. V as a function of time is assumed to be the same for both conditions. Thurstone (1925) pointed out the properties of such a plot which were subsequently utilized in signal detection theory. Linearity implies underlying normality of a theoretical variable in both conditions. The slope estimates the ratio of the two σs (σ_x/σ_y), and the two intercepts give the scale separation between the two conditions in units of the two respective σs. In the present model, this is the scale distance between the two criterion means. The spacing of the points along the axis of the linear REC itself provides a joint estimate of the growth of V. Since the ratio of the σ units is known, the values of one condition may be converted to the unit of the other. The estimate of V at each time is the mean of the two values in common units. The origin of the function is arbitrarily located at the mean of either criterion. It is interesting and proper that such a joint estimate of a function is almost invariably more regular than that of either condition alone. This procedure may be extended to any number of conditions by successively forming another REC between one of the first two conditions and a third, etc. In the end, all values are converted to a common unit, and the mean of any condition is arbitrarily chosen as the origin. More complete discussion of these matters have been given in previous papers by Grice (1968, 1971, 1972a, b).

In the experiment by LaBerge (1971), Ss in all conditions were to respond to a tone of 1,000 Hz. The conditions differed with respect to the events which occurred on catch trials. One condition employed conventional catch trials with blank trials or nothing occurring at the expected time of the tone. In a second condition, a stimulus in another modality, a red light, was presented at that time. In the third condition, the stimulus presented was a noise. The fourth condition involved a 1,200-Hz tone. In general, the results were that of increasing RT and variability in the order in which the conditions were given above. LaBerge's interpretation was in terms of different levels of processing required, emphasizing particularly the distinction between detection and discrimination. Grice's analysis, however, suggested that the entire result could be explained by differences in criterion effects induced by the different conditions. Employing the procedures described above, he obtained estimates, on a common scale, of the criterion parameters for each of the three Ss in each condition. He then obtained an estimated recruitment function for V to the 1,000-Hz tone which was common for all Ss and conditions. This turned out to be a remarkably smooth function and was fitted with an exponential growth equation. Using values of V calculated from this equation and the estimates of T and σ, he then obtained good fits to all 12 of the RT distributions.

Since the rate of sensory growth is assumed to depend upon stimulus intensity, a reasonable next step in this type of research is to introduce intensity variation. This is the first concern of the present research, and three levels of signal intensity have been studied. Following the reasoning of Grice (1968), it was deemed necessary that this manipulation be carried out within Ss as well as within conditions in order that comparable criterion values should apply to all intensities. Four conditions similar to those of LaBerge have been used, but, in addition, a no catch trial condition has been introduced in an effort to extend the range of criterion manipulation. Also, since the variable criterion model contains a built-in approach to analysis of the speed-accuracy tradeoff, the high tone catch trial condition was also replicated under instructions emphasizing speed rather than accuracy, as an initial attack on this problem.

While the present research is based upon the design used by LaBerge, it does not consist of a replication with added intensity manipulation. There is one particularly significant procedural difference. LaBerge used the 1,000-Hz RT signal also as the warning signal, with a constant 1-sec interval between the two presentations or prior to catch signals. Because of complications involved with the intensity differences, this was judged not to be a reasonable procedure here. We used a light warning signal with a variable warning signal interval. Thus, in the LaBerge procedure, the response decision was based on an immediate comparison or a same-different judgment, while in the present research it was based upon an absolute judgment. This may be one of the factors involved in the fact that the present research has only partially confirmed Grice's analysis of the LaBerge results. In the conditions with blank, visual, or no catch trials, the interpretation appears to be well confirmed. However, in the conditions with auditory catch trials, decision appears not to be based upon the same recruitment functions. In this respect, the results more nearly support LaBerge's distinction between detection and discrimination processes. For this reason, the findings from the two classes of conditions will be presented and analyzed separately.

METHOD

Subjects

The Ss were five young adults who were paid for participation. Four were psychology graduate students and one was a secretarial employee of the Psychology Department. Ss 1, 2, and 3 were men, and Ss 4 and 5 were women.

Procedure and Apparatus

The RT response was the depression of a conventional telegraph key. The warning signal was the onset of a 7-W bulb behind a 4 x 8 cm milk glass window in a small box on S's table. Its duration was .5 sec, and its onset preceded RT or catch signals by 1, 1.5, and 2 sec equally often in an irregular order. The RT signal was a 1,000-Hz tone of 50, 80, or 100 dB spl. All auditory signals were presented binaurally through Telephonics TDH-39 earphones. Switching with rise and decay times of 10 msec was by means of a Grason-Stadler electronic switch. RTs were registered in milliseconds on an electronic counter and were recorded manually by E. Programming and stimulus selection were by paper tape readers and solid state logic devices. Ss were seated in an Industrial Acoustics sound chamber.

The six experimental conditions were as follows: none–no catch trials; blank–no signal followed the warning signal on catch trials; red–catch trials consisted of the onset of a red light in a similar window, beneath that of the warning signal; Ss were instructed to watch for the red light, but not to respond to it; noise–the catch signal was a band of noise 200 to 4,000 Hz and its intensity was 50, 80, or 100 dB spl; high, accuracy–the catch signal was an 1,800-Hz tone of 50, 80, or 100 dB spl (pilot work indicated that the 1,200-Hz tone used by LaBerge was too difficult under the present experimental conditions); in this condition, as in the preceding conditions, Ss were instructed to respond as fast as possible, while avoiding errors; high, speed–this was the same as the previous condition with modified instructions–Ss were told now to react faster at a speed where they could not avoid making errors; it was suggested that if they made about 25% errors, they would be at about the desired speed.

Each experimental session began with a warm-up period of 21 RT trials, 7 with each signal intensity in an irregular order. In catch trial conditions, there were also 9 catch trials, 3 of each intensity if they were auditory signals. The session proper consisted of 180 RT trials, 60 with each intensity. There were 90 catch trials in such conditions, 30 with each intensity if they were auditory. Catch trials, RT trials, and intensities were presented in irregular orders. Signals remained on for 1.5 sec, and if S had not responded by that time the trial was terminated. The occasional missed responses to the RT signal were counted as long-latency responses. The intertrial interval was 6 sec. Two brief rest intervals were introduced in a session. Each S began with one practice session under the none condition. Following that, there were three sessions under each of the first five conditions, conducted in a different irregular order for each S. All orders were such that each of the five conditions occurred before one was repeated. Following completion of the first five conditions, each S was run for five sessions in the high, speed condition. In all, each S participated in 21 experimental sessions.

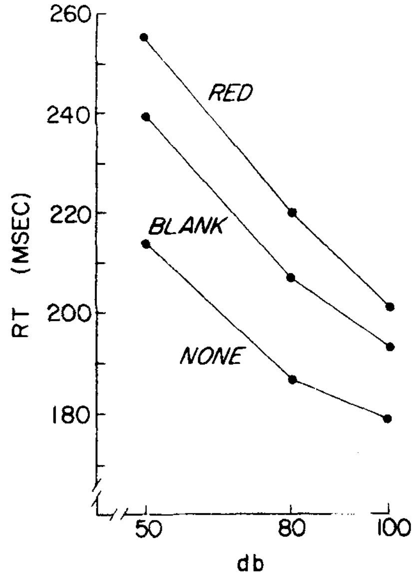

Fig. 1. Means of median RTs in the none, blank, and red conditions, plotted as a function of intensity.

RESULTS AND THEORETICAL ANALYSIS

Conditions None, Blank, and Red

As indicated above, the results in conditions with and without auditory catch trials were sufficiently different to preclude application of the same theoretical analysis to both, and they are considered separately. The medians and interquartile ranges, as a measure of variability, for the none, blank, and red conditions are presented in Table 1. S 3 made eight, obvious short-latency anticipatory responses to each intensity in the none condition. Since these are excluded, these entries are based on 172 responses. All others are based on 180. The means of the medians are presented in Fig. 1. While this research is regarded primarily as an investigation of individual Ss, these means do no serious violence to the individual data, since the Ss were quite consistent in this respect. The figure shows the typical picture of a strong, within-S stimulus intensity effect in combination with criterion manipulation (e.g., see Kohfeld, 1969; Murray, 1970; Henriksen, 1971; and Speiss, 1973). This includes the steeper intensity functions associated with stricter criteria. Or, conversely, it may be stated that the magnitude of a criterion effect varies inversely with stimulus intensity. An analysis of variance of these data merely confirms the visual picture. Its results were as follows: intensity, $F(2,8) = 57.01$, $p < .001$; conditions, $F(2,8) = 15.46$, $p < .005$; Intensity by Conditions, $F(4,16) = 3.99$, $p < .025$.

The variability also shows the typical inverse relation to signal intensity. It also generally tends to increase as the condition medians increase, but this is a bit less consistent over Ss. Means of the interquartile ranges are presented in Fig. 10 for purposes of a later comparison.

Table 1
Median RTs and Interquartile Ranges (Q) for Conditions None, Blank, and Red

	Condition	50 dB Md	50 dB Q	80 dB Md	80 dB Q	100 dB Md	100 dB Q
	None	179	23	167	19	158	17
S 1	Blank	194	41	171	30	161	24
	Red	212	68	181	54	169	43
	None	207	54	177	41	173	32
S 2	Blank	213	32	188	30	178	25
	Red	213	45	194	34	181	30
	None	186	36	164	23	150	18
S 3	Blank	230	64	187	47	168	41
	Red	269	66	215	63	187	43
	None	229	43	206	29	200	28
S 4	Blank	265	53	235	46	220	31
	Red	282	59	245	47	226	53
	None	268	39	229	36	213	30
S 5	Blank	293	54	255	49	238	38
	Red	300	40	264	54	244	42

For purposes of quantitative analysis in terms of the variable criterion model, it was necessary to treat individual Ss in different ways. It is an important boundary condition of the model that S perform in a consistent manner in the block of trials under analysis, since only under such circumstances may a normal criterion distribution be expected. The Ss differed considerably in their consistency from session to session. Ss 1 and 3 were sufficiently consistent in all sessions that it was reasonable to combine the three sessions for each condition. S 2 was quite consistent in the blank and red conditions, but not in the none condition. Thus, for this S, data were combined for blank and red, but only a single session was included for none. The second session, which was intermediate in performance between the other two, was arbitrarily selected. Ss 4 and 5 had considerable between-sessions variability in all conditions, and will be treated separately.

The first step in the analysis of the first three Ss was to obtain estimates of the differences between criterion means and the σ ratios between conditions. Since the value of T for one condition becomes the origin and its σ the unit of the scale, complete scaling requires one less REC than the number of conditions being scaled. Since RECs based on cumulative distributions are limited to the overlapping portions, it is advisable to pair those conditions in which the overlap is maximal. In the present instance, there are actually three potential RECs for each pair of conditions. Since each stimulus intensity has a different temporal recruitment function, there is one for each intensity. However, since the three intensities were presented in an irregular order within sessions and conditions, response to them should be based upon the same criterion distributions. This implies that, except for sampling error, their slope and intercept parameters should be equal. If, in fact, the three do have similar slopes and intercepts, considerable support is provided for the model. In the present instance, they have been plotted together on the same coordinates as a single REC, but the three sets of points have been distinguished. The two such hybrid RECs for each S, which were used in the scaling solutions, are presented in Fig. 2. Points are plotted at 10-msec intervals, except

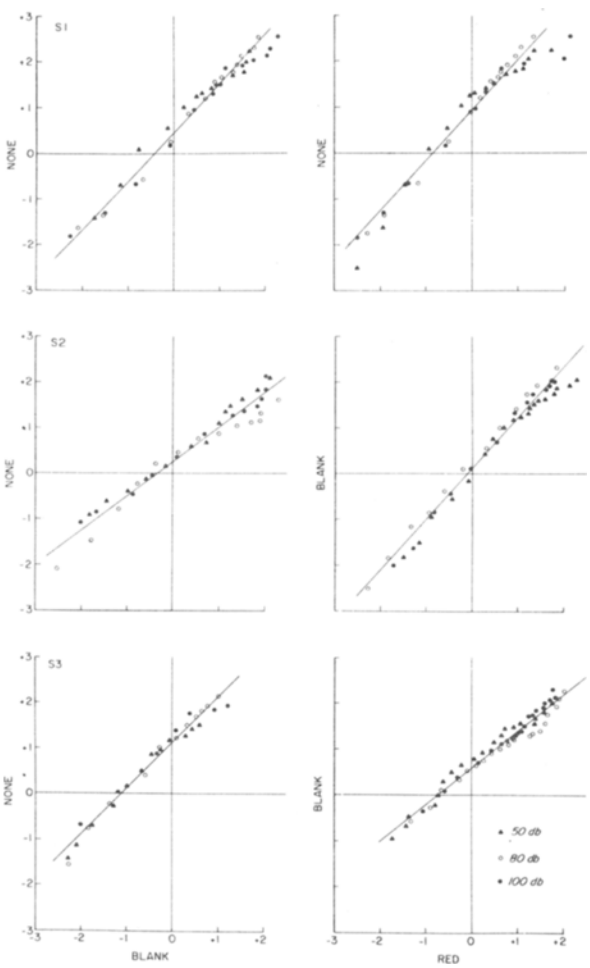

Fig. 2. Between-conditions RECs for Ss 1, 2, and 3 in the none, blank, and red conditions.

where no new information is available. The overall picture of linearity provides generally good support for the assumption of normality, and there certainly is no evidence of systematic slope and intercept differences dependent upon intensity. The tendency for tail points to become scattered is to be expected because of the low reliability of normal deviates based upon extreme proportions. Since we wished to obtain criterion estimates for each condition, independent of intensity, and in view of the consistent picture, each of the six sets of points was treated as a single REC and fitted with the straight line indicated.[1]

The scaling for each S used the intercepts as estimates of the differences in threshold means (d) and the slopes as estimate of the σ ratios. The first step is to convert the values of d into the units of a single condition σ by appropriately dividing or multiplying by the estimate of their ratio. The three conditions were then adjusted to a common origin at the mean of one condition by the use of the values of d. These scaled values of T and σ are presented in Table 2. As expected, the lowest criterion is in the none condition for all Ss. Difference between blank and red varied somewhat between Ss as did the effect of the conditions upon criterion variability. It should be noted that the units in Table 2 are comparable within Ss, but not between Ss. The values for Ss 4 and 5 were obtained in a somewhat different way to be described later.

In obtaining estimates of the recruitment functions for each intensity, the normal deviate transformations of the entire cumulative distributions were converted to the scale of the third by multiplying by the value of σ and adding T. Thus, there were now estimates of the three recruitment functions from each of the three conditions in common units. The joint estimates of these functions was a weighted average of the three, based on the Mueller-Urban weights, computed at 10-msec intervals. In the case of S 2, the estimate was based on the blank and red conditions only, since the none data were from a single session. The points so obtained described three regular, negatively accelerated functions for each S, which were fitted by exponential growth functions of the type previously used by Grice (1972b). Graphs of these equations are presented in the upper left panel for each S in Fig. 3. This quite consistent family of curves describes the scaled temporal growth of sensory information (V) as it depends upon stimulus intensity. The locations of the criterion means for each condition are also indicated on the ordinates. If normal probability

Table 2
Mean (T) and σ of Criterion for Each Condition

	None		Blank		Red	
	T	σ	T	σ	T	σ
S 1	.000	1.000	.460	1.064	.922	1.057
S 2	−.362	1.332	.000	1.000	.121	1.115
S 3	−1.108	1.002	.000	1.000	.584	.800
S 4	.000	1.000	.910	.805	.818	.793
S 5	.000	1.000	.571	1.163	.430	1.020

functions were located at these points, drawn vertically, with the σs of Table 2, they would illustrate the logic of the model. At any given time, t_i, the area of a normal function lying below the ordinate of a recruitment function estimates the cumulative probability of above-criterion response strength. In other words, it estimates the cumulative probability of response in a particular condition to a stimulus of particular intensity. From the three growth equations and the values of T and σ, all of the nine cumulative distributions may be calculated for each S. The results of these calculations are presented in the remaining portions of Fig. 3, together with points indicating the empirical data. These fits are regarded as rather adequate in describing the general form and relationships of the distributions. As an additional aid in evaluating the goodness of fit, we have obtained the calculated values of the medians and interquartile ranges (Q) for comparison with the 27 corresponding values of each in Table 1. In the case of the medians, a coefficient of determination (r^2) indicates that the calculated values account for 97% of the variance of empirical values. For the values of Q it was 87%. The constant error in predicting medians was −0.4 msec and +1.3 msec for Q. The mean absolute error for medians was 3.8 msec, the median error 3.0 msec, and the rms error 4.4 msec. The corresponding errors for Q were 4.7, 3.5, and 6.3 msec. The number of theoretical constants computed from the data to fit the nine distributions for each S was 13, 3 for each growth function, 2 σ ratios, and 2 values of d. One value of T and one of σ were arbitrarily assigned. In the case of S 2, the three distributions for the none condition were fitted with just two constants computed from the data.

A between-Ss solution is not presented here as it was in the analysis of the LaBerge data. This could have been done in the case of Ss 1 and 2 with practically no loss in goodness of fit. The relation between the two sets of computed recruitment functions was linear with only very slight separation of the three intensities. Either set of functions or a joint estimate of them would have provided essentially the same fit as obtained. The criterion parameters for these Ss may be converted to comparable units for comparison. For example, the values of T for S 1 on the scale of S 2 for none, blank, and red are −1.141, −.530, and +.089, and the values of σ are 1.330, 1.414, and 1.338. There was a different situation in the case of S 3. When these functions were plotted against Ss 1 or 2, the three intensities showed clear separation at short latencies and later convergence at longer latencies. This indicated a greater intensity difference in initial growth rate for S 3. Also, the three intensity RECs were each moderately curvilinear. Since normality was indicated by all within-S RECs, this indicates a later, more graduate approach to asymptote for S 3. Clearly, the recruitment functions were different for this S, and between-S scaling was not possible with the present methods.

The two Ss whose performance differed substantially from session to session provide the opportunity for an additional analysis. The obvious interpretation of the theory is that these Ss displayed differences between the sessions of each condition in the level and variability of their criteria. The first step in the analysis was to obtain estimates of the session parameters within each condition. When plotted, the appearance of these six RECs for each S was similar to the between-conditions relations of Fig. 2, and the parameter estimates were made in the same way. Using these estimates, the nine cumulative distributions within each condition were then transformed to the units and origin of a single session within the condition. Within-condition estimates of the three recruitment functions were then obtained by computing the weighted average of the three transformed session distributions for each intensity at

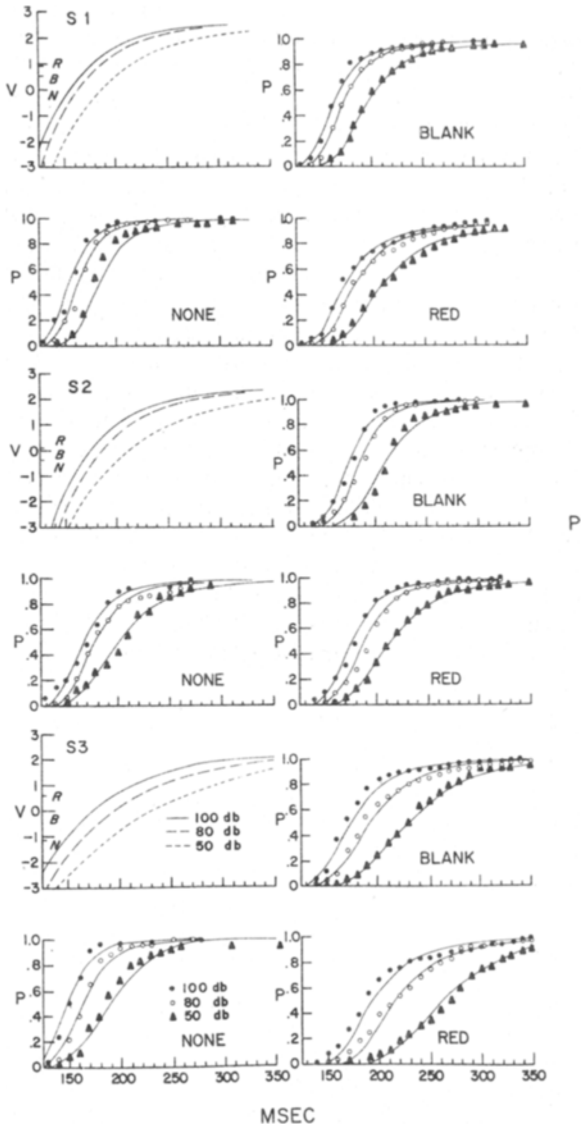

Fig. 3. Obtained and calculated cumulative distributions for Ss 1, 2, and 3 in the none, blank, and red conditions. Graphs of equations fitted to the three intensity recruitment functions are in upper left for each S.

10-msec intervals. The scale values so obtained provide the basis for estimating the average effect of conditions upon the criterion parameters. Thus, the scale values from the conditions were plotted against each other, again in the form of hybrid RECs containing all three intensity functions. These RECs were limited to the range in which all sessions overlapped and were thus contributing to the weighted means. These RECs for Ss 4 and 5 are presented in Fig. 4. Again, they are essentially linear, with no systematic departure of the three intensities from the common linear trends. The values of T and σ for each condition for Ss 4 and 5, presented in Table 2, were obtained from the lines fitted to these RECs. For each S the origin of the scale is the criterion mean of a single session in the none condition and the unit is the σ for that session. The three sets of scale values for each condition were next transformed to this scale. Final estimates of the three intensity recruitment functions were then obtained from weighted averages, over conditions, of these transformed scale values for each intensity. In this instance, the weights employed were sums of the weights used in obtaining the within-condition estimates from the original session data.

The points comprising the estimates of sensory recruitment formed regular sets of functions similar to those obtained in the previous analysis. Since a between-S solution was planned in this instance, only the data for S 4 were fitted with exponential growth functions. These scale values, together with curves calculated from the fitted equations, are presented in Fig. 5. The analysis appears to have been quite successful in displaying the underlying regularity of sensory

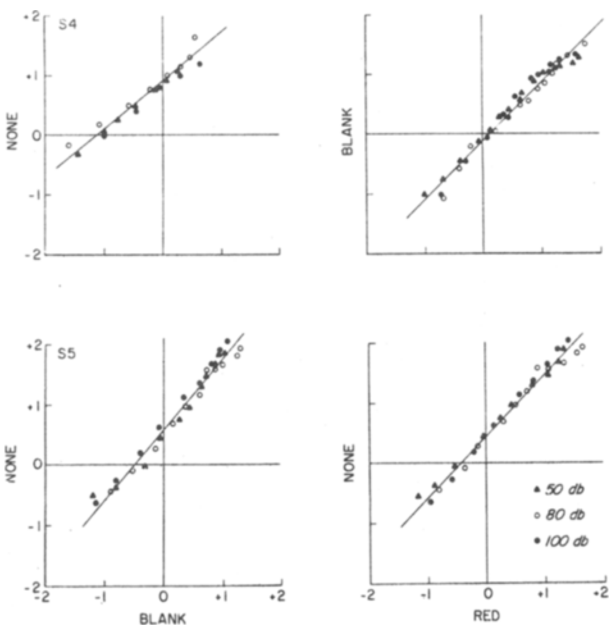

Fig. 4. Between-conditions RECs for Ss 4 and 5 in the none, blank, and red conditions. Points are scale values averaged over sessions within each condition.

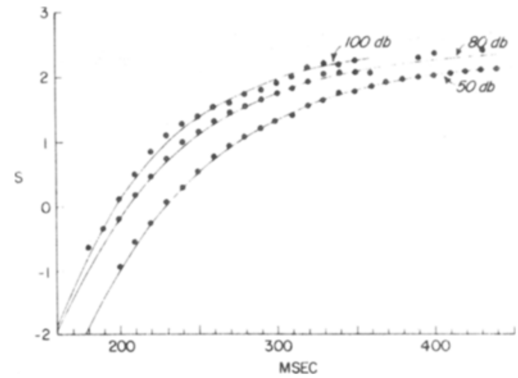

Fig. 5. Sensory recruitment or detection functions for S 4. Points are scale values averaged over conditions for each intensity. Curves are computed from exponential growth functions fitted to these points.

recruitment present in this quite variable set of RT data. This provides the basis for the three intensity distributions in each of the nine experimental sessions involved in the analysis. Here there would be nine criterion mean levels located on the ordinate. In order to complete the analysis, values of T and σ for sessions within the blank and red conditions were transformed to the final scale which was already common to the sessions of the none condition. The 27 calculated distributions for S 4 are presented in Fig. 6, with the empirical data plotted in 20-msec class intervals. The three sessions under each condition are arranged in order from left to right. In this instance, the distributions are not presented in cumulative form. The N of 60 responses with each intensity per session is too small to obtain smooth empirical distributions, but sessions of much greater length would not be practical. In spite of the irregularity of the data, however, the calculated distributions do a fairly good job of representing it. Over all the data in Fig. 6, the calculated proportions account for 86.3% of the variance in the empirical points, which is about as well as could be expected with any family of smooth curves. Also, the predicted and obtained medians and interquartile ranges have been calculated for each distribution. The predicted medians account for 95.1% of the variance of the obtained values, and for Q the percent of variance was 78.7. The constant errors of prediction were +1.0 msec for medians and +1.9 msec for Q The mean, median, and rms absolute errors for medians were 4.4, 3.0, and 6.0 msec. For Q they were 6.0, 5.0, and 8.0 msec. In all, there were 25 constants computed from the data in this analysis, 2 fewer than the number of distributions.

An attempt has been made to apply the recruitment functions obtained for S 4 to the data of S 5. It has been generally assumed in this approach that criterion effects constitute the major source of individual differences. The analysis began by relating the final estimate of the intensity functions for S 5 to those of S 4 obtained from the fitted equations. This relationship is presented in Fig. 7. The functions for 80 and 100 dB are linear, and

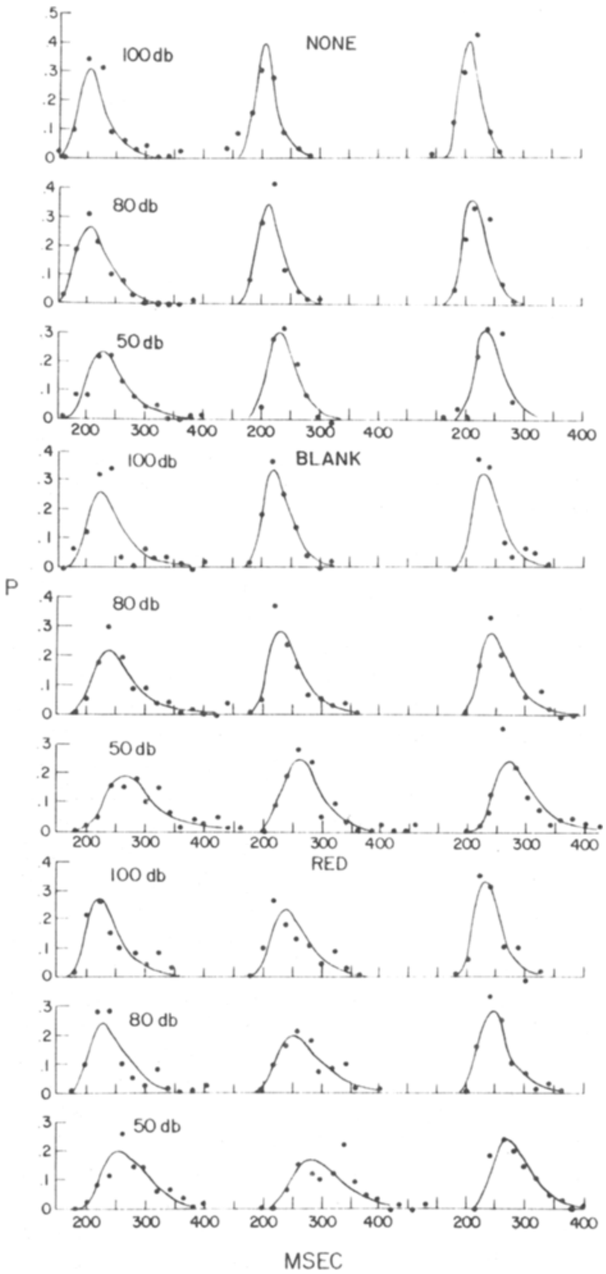

Fig. 6. Obtained and calculated RT distributions for S 4 for all sessions and intensities in the none, blank, and red conditions. Sessions within each condition are arranged in order from left to right.

while there is a slight separation between them, they can be fitted by a single line with only small error. It may be assumed, then, that the two Ss share the same recruitment functions for these intensities and that the fitted line estimates the scale transformation between the two Ss. The 50-dB points, however, show clear divergence, indicating less relative sensitivity for S 5 at short latencies. It has been our informal observation from pilot work in the laboratory that this is characteristic of Ss with minor hearing loss. Because of the evidence that the 50-dB functions were different, a separate exponential growth function was fitted to the data for S 5 for that intensity. Throughout the major range of the function, the quality of the fit was equivalent to that shown for S 4, although the few responses under 250 msec suggested an initial phase of positive acceleration. The fit of the model to S 5's data, then, was evaluated with this separate function and with the 80- and 100-dB functions of S 4, together with the scale transformation indicated by the line of Fig. 7.

The calculated distributions for S 5 together with the empirical data points are presented in Fig. 8. In this analysis there were just 21 constants, computed from the data, used in the fit of the 27 distributions. Again, the fits are generally adequate for such data. The calculated proportions account for 85.4% of the variance in the empirical points. For medians, calculated values account for 98.8% of the obtained variance, and for Q they account for 32.2%. Mean constant errors were +0.8 msec for both medians and Q. Mean, median, and rms absolute errors were: for medians, 3.2, 3.0, and 3.8 msec; and for Q, 7.1, 5.0, and 10.8 msec. It may be noted that the relatively poor showing in predicting the values of Q is largely due to the presence of a single atypical distribution where the fit was clearly poor—the 50-dB stimulus in the first session with the red condition. Of course, Q is also relatively unreliable since it contains the sampling error of two quartile points. If the analyses of Ss 4 and 5 are treated as one, with the individual difference factor included, the predicted medians account for 97.7% of the variance in the 54 obtained medians and for Q the percentage of variance is 77.5.

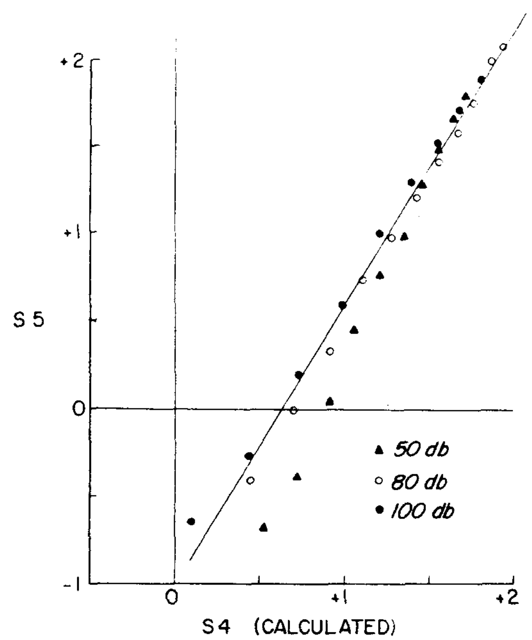

Fig. 7. REC relating final scale values for each intensity for S 5 to those for S 4. Values for S 4 are calculated from the fitted equations. The line is fitted to the 80- and 100-dB data only.

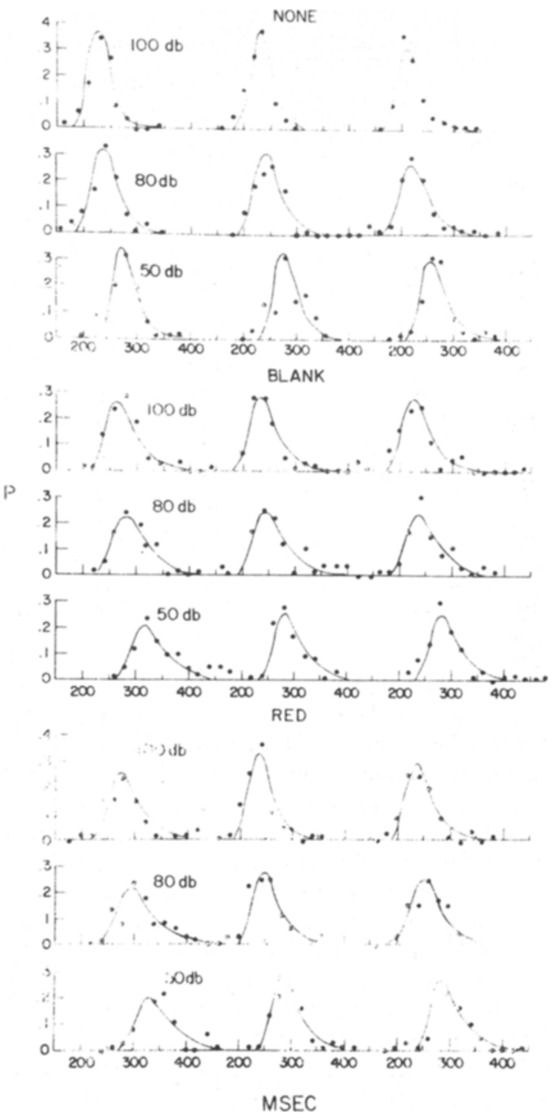

Fig. 8. Obtained and calculated RT distributions for S 5 for all sessions and intensities in the none, blank, and red conditions.

The availability of the estimates of T and σ, in common units, for the nine sessions of each of these two Ss makes possible an observation not previously available. While the means and σs of RT distributions are typically positively related, this is not necessarily true of means and σs of the underlying criterion distributions. In fact, the rank difference correlation between T and σ for each of these Ss is $-.55$. This at least hints of the possibility of a tradeoff between criterion level and variability. Presently, this is regarded merely as a suggestive possibility for future observation.

Auditory Catch Trial Conditions and the Speed-Accuracy Tradeoff

For the conditions with auditory catch trials, the median RTs, the values of Q, and the numbers of responses to catch signals (errors) are presented in Table 3. For the high, speed condition, the medians and Qs of the error distributions are presented as well. For noise and high, accuracy, there were 180 RT trials and 90 catch trials of each intensity. For the high, speed condition, there were 300 RT trials and 150 catch trials with each intensity. In general, as expected, the RTs are slower and more variable than in the first three conditions. While the individual patterns are not quite so consistent here, certain features of the data are clear from an examination of means. Means of the medians are presented in Fig. 9. None of these intensity functions is nearly as steep as would be expected by an upward extension of the family of functions in Fig. 1. There is even an inversion at 100 dB in the high, accuracy condition, and in some sessions response to 100 dB was actually the slowest. Clearly, there is more relative difficulty with the higher intensities than in the other conditions. The most obvious implication of the different patterns of Figs. 1 and 8 is that there are different processes involved other than mere elevation of the criteria.

An even more striking difference in the two sets of conditions may be seen in the means of the values of Q presented in Fig. 10. In the first three conditions, presented on the right, variability shows its typical decrease as a function of intensity. On the contrary, all of the auditory catch trial conditions, for correct responses, showed their greatest variability at 100 dB. Recalling our earlier point that a cumulative RT distribution is a continuous measure of information processed as a function of time, the value of Q may be regarded as an inverse measure of processing rate, i.e., it is the time required to process the middle half of the responses. Thus, processing rate was slowest for the loudest signal in these conditions.

In the high, speed condition, errors were consistently faster and of less variability than correct responses for

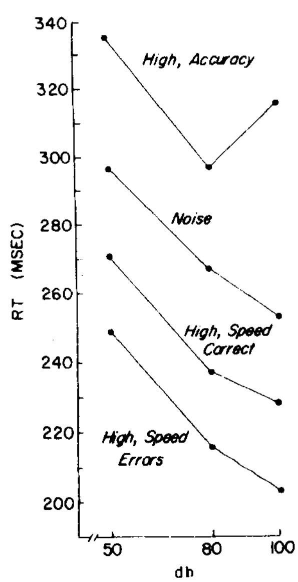

Fig. 9. Means of median RTs in the noise; high, accuracy; and high, speed conditions, plotted as a function of intensity.

Table 3
Median RTs, Interquartile Ranges (Q), and Numbers of Errors in Auditory Catch Trial Conditions

		50 dB			80 dB			100 dB		
S	Condition	Median	Q	Errors	Median	Q	Errors	Median	Q	Errors
S1	Noise	299	51	4	285	50	6	284	67	6
	High, Accuracy	360	60	6	286	60	6	298	75	6
	High, Speed, Correct	250	43		226	56		235	100	
	High, Speed, Errors	233	24	24	187	40	29	175	21	25
S2	Noise	261	54	2	235	45	15	224	64	8
	High, Accuracy	266	51	6	237	52	8	220	63	7
	High, Speed, Correct	223	50		194	32		184	26	
	High, Speed, Errors	226	42	55	199	48	89	199	33	88
S3	Noise	272	62	3	228	39	13	202	48	5
	High, Accuracy	314	62	1	272	71	7	301	148	1
	High, Speed, Correct	253	80		210	64		189	86	
	High, Speed, Errors	245	60	31	189	61	57	172	44	64
S4	Noise	308	57	1	289	68	3	266	75	4
	High, Accuracy	370	79	3	314	98	6	380	143	5
	High, Speed, Correct	300	63		259	50		253	73	
	High, Speed, Errors	284	39	35	248	42	50	226	32	49
S5	Noise	347	85	0	298	86	6	287	104	2
	High, Accuracy	423	123	2	377	84	1	382	94	1
	High, Speed, Correct	330	76		294	66		279	108	
	High, Speed, Errors	308	52	51	259	47	66	243	61	44

four of the five Ss. S 2 was an exception in this respect. It is not clear whether he performed the tradeoff on some different basis or whether this reflects his general low level of discrimination (52% overall error rate). The intensity function for errors more closely resembles those of the first three conditions than do any of the functions for correct responses.

Attempts have been made to apply the variable criterion model to the data of these conditions, but these were generally not sufficiently successful for full presentation here. However, the results will be summarized. Attempts to relate the auditory catch trial conditions to the earlier ones by means of RECs were hampered by limited overlap, but their outcome was still convincing. Generally, they tended to be curvilinear in the upper portions, indicating a later growth than could be predicted by the recruitment functions of the first analysis. In addition, the intensities tended to separate

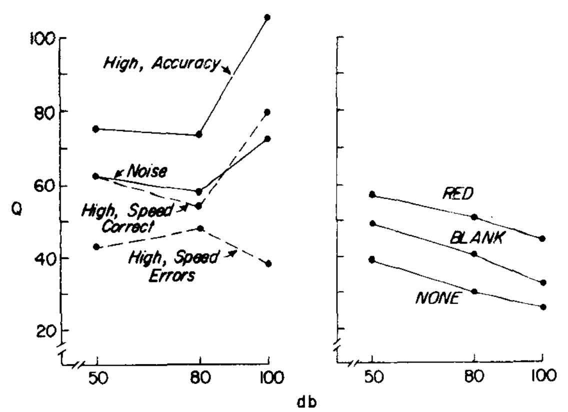

Fig. 10. Means of interquartile ranges (Q) as a function of intensity for all conditions.

with increasing latency, reflecting the slower relative growth of information from intense stimuli found in the present conditions. Thus, these attempts to apply the model were useful in demonstrating that the growth of the information required for response has a different time course in the two classes of conditions. This supports LaBerge's distinction between detection and discrimination. In his language, the sensory recruitment functions of the first analysis could be termed detection functions. They purport to represent the basic laws describing the temporal growth of detection information as a function of stimulus intensity.

Following the evidence of different underlying growth functions, attempts were made to apply the model separately to the auditory catch trial conditions and to extract from the data the underlying discrimination or associative growth functions. The basic conception was that the associative strength of the RT signal grows in a negatively accelerated fashion. The erroneous associative strength of catch signals grows in a similar fashion, but reaches a lower asymptotic strength, and the functions diverge with time. The difference between them could be termed a discrimination function. Responses would occur to either stimulus when the criterion was below the function of that stimulus. This conception provides a model for the speed-accuracy tradeoff in terms of adjustments in criterion level and variability, and it tends to predict that errors will be of shorter latency and less variability than correct responses. Attempts to apply the model were primarily with the high, speed condition where estimates of both functions were possible. The data did, in fact, tend to conform approximately to such a model, and there was even a suggestion that it should be possible to predict the accuracy data from the speed functions. The problem preventing a full and convincing application was the lack of sufficient stability in the data. Between-session variability was large. In some between-session RECs, there was curvilinearity indicating lack of within-session consistency. In other cases there was some separation of the different intensity RECs, indicating either shifts in strategy or changes in S's conception of the discrimination functions. The final decision was that an adequate evaluation of the model was not possible with the present data.

In the absence of a full theoretical account of these data, an analysis has been performed on the two high conditions in terms of information transmission in order to elucidate the nature of the problem facing the S and to provide an analytic description of the speed-accuracy tradeoff. This was a two-dimensional analysis of information transmitted (H_T) (Attneave, 1959). The analysis was computed from the cumulative RT distributions of both errors and correct responses. Response information was based on the probabilities that S had and had not responded by a particular time. The stimulus dimensions were frequency (1,000 and 1,800 Hz), the relevant dimension, and intensity (50, 80, and 100 dB), an irrelevant dimension. The probability of each intensity was .333, so there were 1.585 bits of information in this dimension. The probability of the 1,000-Hz RT signal was .667, and the probability of the 1,800-Hz catch signal was .333, making the information in this stimulus dimension .918 bit. This would be the final amount of transmission in perfect performance. The analysis was computed at 40-msec intervals in the range from 150 to 510 msec. It was computed for the high, speed and accuracy conditions separately and, for each S, was based on all sessions combined.

The results of this analysis are presented graphically in Fig. 11. Information transmitted in the speed condition is presented in the upper left panel for each S, and the accuracy condition is in the upper right. The three functions display the time course of total information transmitted by the two stimulus dimensions in combination and by frequency and intensity individually. For four of the Ss, the total transmission function in the speed condition shows an early hump before entering a phase of smooth increase with negative

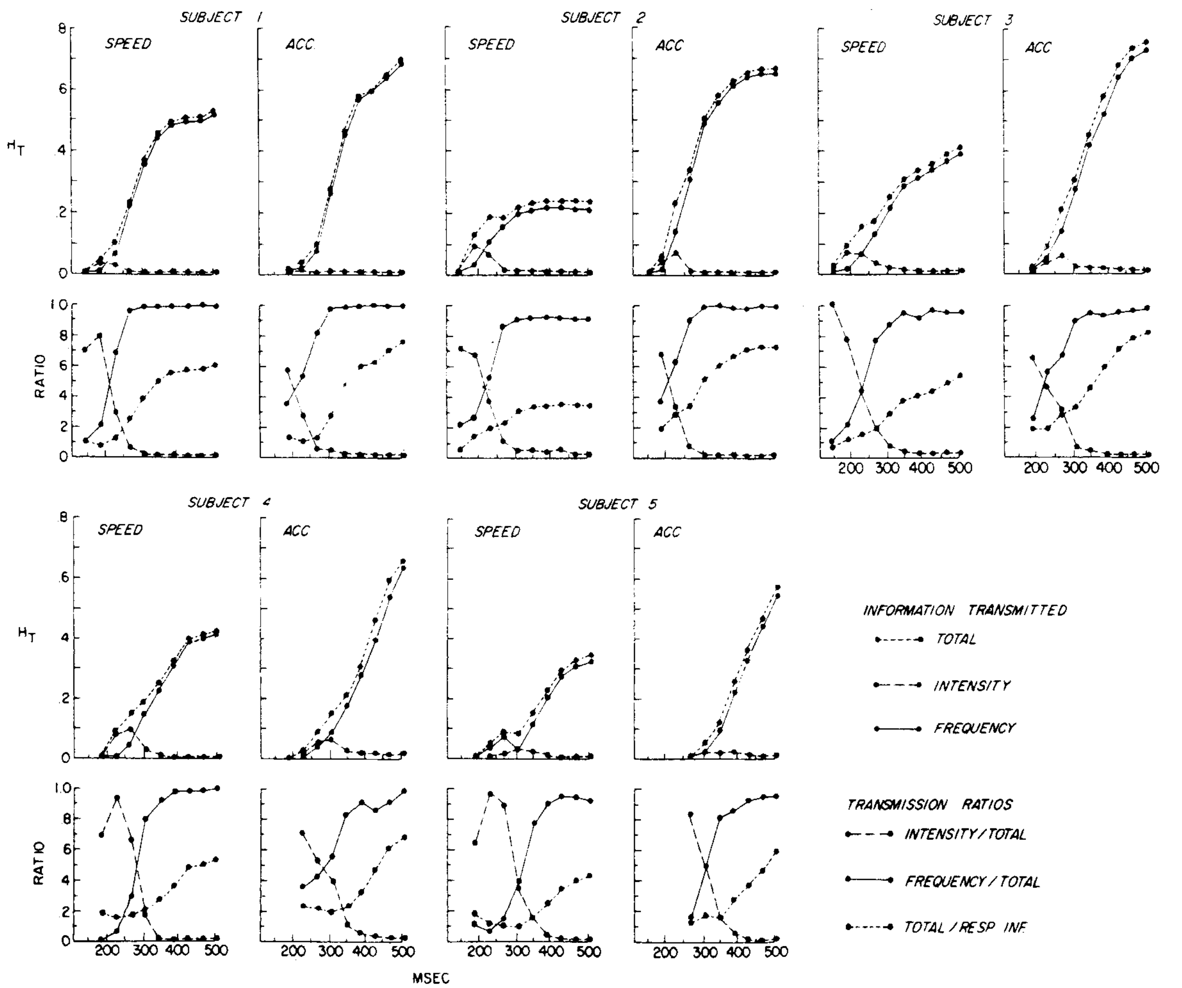

Fig. 11. Two-dimensional analysis of information transmitted (H_T) for all Ss in the high, speed and high, accuracy conditions. Upper figures for each S indicate total transmission (in bits) and transmission by frequency and intensity separately, all plotted as functions of time. Corresponding transmission ratio functions are immediately beneath.

acceleration. For Ss 2 and 5 the hump is actually a maximum. This would appear strange, were it not for full analysis which shows that, for all Ss, transmission by intensity reaches a maximum in this short-latency region before decreasing to near zero levels at longer latencies. S 1, who showed the least tradeoff, does not display the hump in total transmission simply because there was very little responding at the short latencies where transmission by intensity reaches its maximum. For the same reason, the early maximum for intensity transmission is either reduced or absent for all Ss in the accuracy condition. In both conditions, transmission by the relevant frequency dimension increased with initial positive and subsequent negative acceleration. Except for the different levels of transmission attained, they are quite similar in form and location. For each S, the maximum rate of gain in information transmitted occurs at about the same time in both conditions. It seems probable that these functions reflect an underlying function which would describe the increasing availability or growth of associative information and which limits performance independently of the degree of tradeoff. It may be noted that the several functions of Ss 4 and 5 appear to be displaced to longer latencies than the corresponding features for the other Ss. It seems possible that this may indicate slower movement times rather than true differences in the rate of processing the sensory and associative information.

In addition to the H_T measures themselves, three "transmission ratios" have been computed from them as a further analytic device. These were also plotted as functions of time and are presented in Fig. 11 below the corresponding H_T measures. Two of these are the proportions of the total transmission which are transmitted by intensity and frequency. It should be noted that these two ratios do not sum exactly to one, since this type of analysis has an interaction component, which is quantitatively negligible in this instance. The two functions are essentially complements of each other. The percentage of transmission by intensity begins a high level at short latencies and decreases rather sharply to a very low level. Conversely, the proportion of total transmission by frequency begins at a low level and increases quickly to a high level. These functions are quite similar in the speed and accuracy conditions, except that the first one or two points are missing for the accuracy condition where no responding occurred. The third ratio is the proportion of the response information which is transmitted by the total stimulus information. This is the so-called coefficient of constraint of stimulus on response. It indicates that portion of the response uncertainty which is dependent upon which stimulus was presented–in effect, the "nonnoise" proportion of the response information. These functions begin at a low level with an initial flat section or period of slow increase followed by a faster increase and end in a negatively accelerated increase approximately paralleling the frequency transmission function. The elevation of the early segment is due to the short-latency transmission by intensity. Had this index been computed for frequency alone, it would have started at near zero and roughly paralleled the frequency transmission function throughout. The functions for speed and accuracy are similar in form, with the accuracy functions starting slightly higher and reaching substantially higher levels.

This information analysis has produced quite a consistent picture which can be interpreted rather simply. If S's criterion is so low that the response occurs at a short latency when little stimulus information is available, the response information is largely noise, i.e., the response is independent of which stimulus occurred. However, since the intensity information grows more rapidly than the associative information based on frequency, that portion of the early response information which does depend upon the stimulus is constrained by intensity rather than by the relevant frequency dimension. The functions describing proportion of response information transmitted by intensity are similar in the speed and accuracy conditions. What the tradeoff involves with respect to intensity is the number of short-latency responses which occur in the region where constraint by intensity is predominant, and hence with respect to total amount of information so transmitted. The Ss are well aware of this tendency to respond on the basis of intensity, and this affects their conception of the discriminations required. The slower relative growth and higher variability of response to the more intense stimuli no doubt reflect their general wariness in this respect. The major quantitative factor in the tradeoff, however, is not intensity, but the proportion of noise in the response information. In variable criterion theory, noise, or lack of discrimination, is introduced by the two interacting factors of criterion level and criterion variability. With a very low criterion resulting in short-latency responses when little discriminative information is yet available, even low variability will result in errors when the catch signal satisfies the criterion. However, with high mean criterion levels and sufficiently high variability, there will also be frequent errors when error signals exceed the lower criterion values. Error-free performance results from an appropriate combination of criterion level and reduced variability such that the criterion exceeds the erroneous information on all catch trials. The relative efficiency of such combinations may be evaluated by the transmission rate. For example, in the 510 msec of this analysis, the Ss varied from 1.08 to 1.42 bits/sec in the accuracy condition and from .41 to 1.01 bits/sec in the speed condition.

One additional analysis has been performed which further illustrates the effect of stimulus intensity in this situation. The relevant frequency transmission has been computed for RT and catch signals of each intensity considered separately. This has also been obtained at 40-msec intervals in the speed and accuracy conditions and is presented in Fig. 12. The computations were for all trials, for all Ss combined. In both conditions,

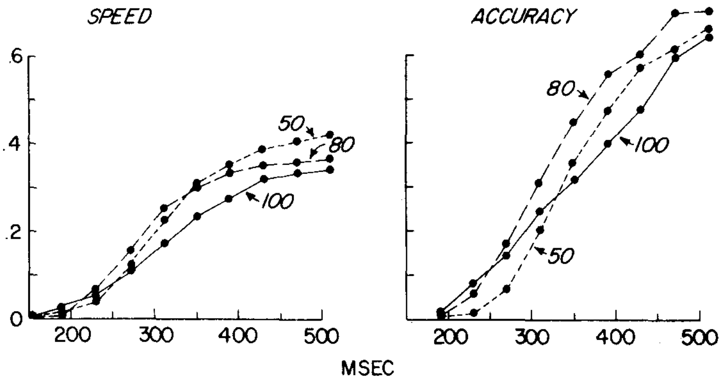

Fig. 2. Functions, based on all Ss combined, indicating frequency transmission for each intensity separately in the high, speed and high, accuracy conditions.

transmission by the 100-dB signal gained at the slowest rate and reached the lowest terminal level. In the speed condition, there was a simple inverse relation between intensity and transmission gain. In the accuracy condition, the 50- and 80-dB functions gained at about the same rate, but the 80-dB function reached a higher final level. This represents quite a different picture than would be obtained from an analogous analysis of the blank and red conditions where there is a simple positive relation between intensity and transmission rate. The analysis is a further illustration of the difficulty experienced by S in inhibiting false responses on the basis of intensity.

DISCUSSION

The theoretical analyses of the individual S data in the none, blank, and red conditions are regarded as lending further support to the variable criterion theory of simple RT. The fact that it was possible to extract such a consistent family of sensory recruitment or detection functions from the different conditions and heterogeneous performances between sessions suggests that they do, in fact, represent a common principle underlying this variable behavior. In combination with the notions of criterion level and variability, the system appears to make sense as a way of analyzing the effects of the experimental conditions, the inconsistency of performance between experimental sessions, and the role of individual differences. Particularly encouraging was the consistency with which the different intensities led to similar parameter estimates in between-conditions and between-sessions RECs. This was convincing evidence that properties of the conditions or sessions were being appropriately measured.

Evidence was convincing that performance in the conditions with and without auditory catch trials was based on different underlying growth functions of the required information and not merely criterion differences. This obviously raises the question as to why the criterion interpretation worked so well with the LaBerge data. The explanation must lie in one or both of the two major procedural differences—the different type of judgment required and the presence of the intensity variable in the present research. The most obvious possibility is that detection information and that required for the simple same-different judgment grow according to the same function, and all that is required is a higher criterion in order to avoid response to the erroneous information. On the other hand, the present task required absolute recognition, and recognition functions appear to have a more graduate rate of growth. Another possibility is that the detection and same-different functions are exactly parallel, but vertically displaced, in their region of overlap—roughly 200 to 300 msec. In this instance, Grice's analysis would not have been sensitive to making the distinction between the vertical separation and a criterion difference. While this may not be highly probable, this limitation of the method should be recognized. One advantage of the presence of the intensity variable here is that it increases the sensitivity of the model to such discriminations. The intensity variable was a complicating factor in the present task. It appears probable that it was responsible for the instability in the auditory catch trial conditions as the Ss attempted varying strategies in their effort to avoid response purely on the basis of intensity. It is highly desirable that ways be found to apply fully the variable criterion model to the present type of data, largely because of its potential power in analyzing the speed-accuracy tradeoff. The present attempts were sufficiently suggestive and close to success that ultimate success seems likely. Various approaches to the obtaining of more stable data are available, which will be attempted in future research.

As a final methodological note, attention is called to the analytic power of repeatedly recalculating the information transmitted analysis as a function of time. For example, if it had been based only on total responses, it would have merely indicated that there was very little information transmitted by intensity and that transmission was greater in the accuracy condition. The interesting dynamics of the situation would not have been displayed.

REFERENCES

Attneave, F. *Applications of information theory to psychology.* New York: Holt, Rinehart & Winston, 1959.

Bock, R. D., & Jones, L. V. *The measurement and prediction of judgment and choice.* San Francisco: Holden-Day, 1968.

Dorfman, D. D., & Alf, E., Jr. Maximum likelihood estimation of parameters of signal detection theory—a direct solution. Psychometrika, 1968, 33, 117-124.

Grice, G. R. Stimulus intensity and response evocation. Psychological Review, 1968, 75, 59-373.

Grice, G. R. A threshold model for drive. In H. H. Kendler and J. T. Spence (Eds.), *Essays in neobehaviorism: A memorial volume to Kenneth W. Spence.* New York: Appleton-Century-Crofts, 1971. Pp. 285-310.

Grice, G. R. Conditioning and a decision theory of response evocation. In G. H. Bower (Ed.), *The psychology of learning and motivation.* Vol. 5. New York: Academic Press, 1972a. Pp. 1-65.

Grice, G. R. Application of a variable criterion model to auditory reaction time as a function of the type of catch trial. Perception & Psychophysics, 1972b, 12, 103-107.

Henriksen, K. Effects of false feedback and stimulus intensity on simple reaction time. Journal of Experimental Psychology, 1971, 90, 287-292.

Kohfeld, D. L. Effects of the intensity of auditory and visual ready signals on simple reaction time. Journal of Experimental Psychology, 1969, 82, 88-95.

LaBerge, D. On the processing of simple visual and auditory stimuli at distinct levels. Perception & Psychophysics, 1971, 9, 331-334.

Murray, H. G. Stimulus intensity and reaction time: Evaluation of a decision theory model. Journal of Experimental Psychology, 1970, 84, 383-391.

Speiss, J. M. Effects of preknowledge and stimulus intensity upon simple reaction time. Journal of Experimental Psychology, 1973, 101, 109-115.

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale University Press, 1956.

Thurstone, L. L. A method of scaling psychological and educational tests. Journal of Educational Psychology, 1925, 16, 433-451.

Thurstone, L. L. Scale construction with weighted observations. Journal of Educational Psychology, 1928, 19, 441-453.

NOTE

1. The method of fitting the lines here was somewhat more refined than the procedures used by Grice in previous analyses. The slope was obtained by the expression:

$$\text{Slope} = \frac{\sigma_y^2 - \sigma_x^2 + \sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4\text{cov}_{xy}^2}}{2\text{cov}_{xy}}.$$

When this line is passed through the point formed by the means, it minimizes the squared perpendicular residuals and the sum of the squared residuals in x and y. Here the slope was calculated with the Mueller-Urban weights, and the intercept was determined by the weighted means. The two weights for each point were combined by the expression derived by Thurstone (1928):

$$W = \frac{1}{\frac{1}{w_x} + \frac{1}{w_y}}.$$

This is an approximate least-squares solution for the bivariate case and is analogous to the weighted "minimum normit chi-square" solutions described by Bock and Jones (1968), applied to single sets of proportions. It is similar to Thurstone's (1928) solution without assuming unit correlation as he did. It yields values close to those of the maximum likelihood solution of Dorfman and Alf (1968) and is suitable for desk calculation.

(Received for publication March 12, 1974; accepted April 18, 1974.)