Application of variable criterion theory to choice reaction time

G. ROBERT GRICE, ROBERT NULLMEYER, and V. ALLAN SPIKER
University of New Mexico, Albuquerque, New Mexico 87131

A theory of choice reaction time (CRT) is presented which is based on a variable criterion model assuming the decision criterion to be normally distributed. The theory provides functions of time describing the growth of component processes following stimulus onset. For correct responses, these processes are sensory and associative strength. For errors, the processes are sensory strength, generalized associative strength, and associative inhibition. A model for separating the effects of these processes from that of response competition is presented. The theory describes, with great accuracy, the CRT distributions of correct responses and errors, not only for the experiment for which it was derived, but also for a second experiment in which the criterion was experimentally manipulated. It accounts for the basic facts of the speed-accuracy tradeoff, not only with respect to the two experiments with different speed requirements, but also with respect to the variation of speed over trials within each experiment. While derived from group data, the theory also describes the performance of individual subjects. The mean and variability of the criterion distribution and the ability to utilize associative inhibition are identified as the major sources of individual differences. Methodological problems involved in the scaling analysis of group data are also discussed.

The many years of continued interest in reaction time (RT) research lies largely in the fact that RT measures the time required for the processes involved in response evocation. Through experimental and theoretical analysis, it has been the goal to identify these processes and to specify their temporal and other properties. One approach, instituted by Donders and pursued more recently with increasing sophistication, is that RT is merely the sum of the times required for each of the necessary component processes. An alternative is that each of the processes involved may develop as a separate function of time following stimulus onset, and that these functions may, at least in part, grow simultaneously. It is this type of interpretation which has resulted when variable criterion theory (Grice, 1968) has been applied to the analysis of latency phenomena. According to this view, response evocation will result when the combined strength of the processes satisfies a decision criterion.

In analysis using the variable criterion principle, we assume the criterion to be a normally distributed random variable, under homogeneous conditions of performance, and treat it as the only source of random variability. This is in contrast to other models

of latency phenomena which have typically attributed variability to assumed stochastic properties of sensory, perceptual, or associative processes, and have assumed the criterion to be constant. Grice (1968) argued that this latter assumption is implausible on the basis of evidence that the criterion is so easily manipulated that it must be extremely labile and the prime candidate to be the major source of variability. Since then, evidence has accumulated that the criterion is readily affected by a wide variety of experimental variables. The major classes of processes involved include motivation, instructions, attention, set, adaptation, and individual differences. The assumption of normality stems from the principle that a complex process, determined by a large number of independent factors, is normally distributed. As Grice (1968, 1972a) has pointed out, the treatment of the criterion as the sole source of RT variability is a simplification. Presumably, no biological process is free from variability, and our variability estimates may properly be regarded as the sum of variability from all sources. However, it is our assumption that criterion variability is large with respect to that from other sources (Grice, 1968). The assumption of normality makes possible the applications of Thurstonian scaling procedures which provide the metric for theoretical analysis.

Our analysis of latency data, from simple reaction time (SRT), disjunctive reaction time involving a single response (DRT), and conditioning, has been based on the cumulative probability distributions of latency. This is in essential agreement with the

approaches of Luce and Green (1972) and McGill (1963) in viewing the distributions as reflecting the development of stimulus-produced information. That is, any point on a cumulative distribution may be taken as an index of the average amount of information processed by that time following stimulus onset. In terms of the variable criterion model, the cumulative distribution indicates the proportion of trials upon which the excitatory strength of the stimulus information was equal to or greater than the criterion. The probability density at any time reflects the rate of gain of information at that time. Other that this similarity, the present approach is quite different from that of the above authors. Rather than viewing the distributions as reflecting directly the probabilistic properties of stimulus input, we consider them as reflecting orderly functions indicating the growth of excitatory strength acting with respect to a normally distributed criterion distribution.

It should also be clear that there is a fundamental difference in philosophy between the present approach and many current mathematical models based upon probability theory. The most common approach is to begin with a priori assumptions concerning the probabilistic nature of the processes, and to formulate a model with sufficient precision that it leads immediately to, perhaps limited but at least directly testable, implications. Our goal, which is long term, is the development of a general theory of response evocation which will include quantitative descriptions of all of the underlying processes involved and their relations to the experimental variables which affect response latency, error production, and such phenomena as the speed-accuracy tradeoff. We believe that the development of such a theory depends more upon discovery than upon the formulation of elegant but limited models with immediately testable implications. Such discovery comes about through judicious experimentation in combination with thorough analysis of the data at a theoretical level. The variable criterion principle, with its limited assumptions, provides the rationale and the methods by means of which we propose to conduct such theoretical investigations. Thus, we regard theory development as a more gradual and inductive process than appears to be the present vogue. Certainly, there is a basis for legitimate differences of opinion concerning choices of scientific strategy, and only time will be the judge. The present approach does not imply that experimental tests of theory should never accompany the longer term process of theory development. The present investigation, in fact, includes a test of the theory derived here, based on an experimental manipulation designed to affect criterion parameters.

The Variable Criterion Model and Its Scaling Implications

Variable criterion theory has so far been applied only to situations involving a single response. While the present paper is concerned with choice reaction time (CRT), it is desirable first to present the model as it applies to the single response case. The probability of response by any time following stimulus onset is equal to the probability that excitatory strength is equal to or greater than the criterion. Excitatory strength, expressed in scale units, is a deterministic, increasing function of time. Over trials, the criterion is a normally distributed, random variable. The mean and standard deviation of a criterion distribution depend upon the nature of the task, the nature of the subject, and a variety of experimental manipulations. The model, as illustrated in Figure 1, represents a case in which experimental conditions have produced differences in both the mean and variability. The two normal criterion distributions, with means $\overline{C}_1$ and $\overline{C}_2$ and standard deviations σ_1 and σ_2, are plotted on the theoretical dimension of excitatory strength (E). Also indicated on this dimension are selected values of a function indicating the growth of excitatory strength as a function of time since stimulus onset. These values
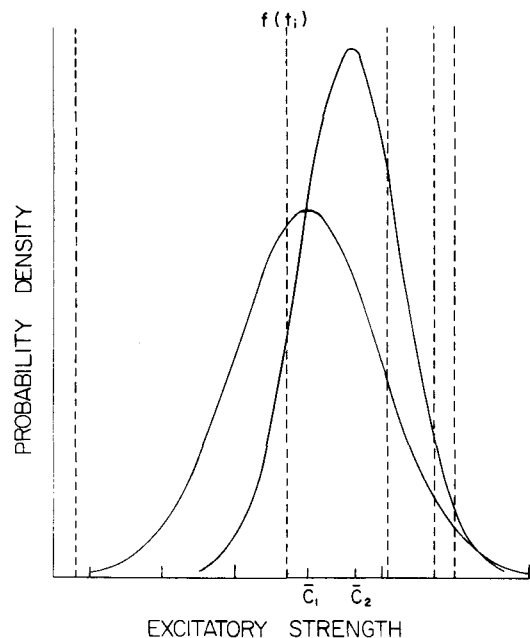

Figure 1. Graphic illustration of the variable criterion model applied to the evocation of a single response. The two normal functions, with means $\overline{C}_1$ and $\overline{C}_2$, represent the theoretical distributions of the criterion. The vertical lines represent values, equally spaced in time, of a function describing the increase of excitatory strength with time following stimulus onset. The cumulative probabilities of response by t_i are given by the areas of the normal functions to the left of $f(t_i)$.

are indicated by the vertical lines, one of which is labeled $f(t_i)$. For SRT, such functions are negatively accelerated and the values of the function indicated here are equally spaced in time. For each distribution, the proportion of trials on which excitatory strength equals or exceeds the criterion by t_i is given by the proportion of the normal function below (to the left of) $f(t_i)$. This also estimates the cumulative probability of response at t_i. In other words, cumulative probability of response at t_i is given by the integral of the normal function from minus infinity to $f(t_i)$. The empirical data are cumulative RT distributions which give the proportions of response by each t_i. The scaling task is to obtain from the empirical data values on the theoretical excitatory strength dimension for $\overline{C}_1$, $\overline{C}_2$, σ_1, σ_2, and also to obtain sufficient values of $f(t)$ that the form of the entire function may be determined. The first step is to obtain the inverse normal function (normal deviate) of each proportion on the two cumulative latency distributions. These give the values of $f(t)$ expressed as distances from $\overline{C}_1$ and $\overline{C}_2$ in units of σ_1 and σ_2. Thus, within the overlap of the distributions, there are two estimates of each value of $f(t)$, one on the scale of each distribution. These may be expressed as follows:

$$E_{1i} = \frac{f(t_i) - \overline{C}_1}{\sigma_1}$$

$$E_{2i} = \frac{f(t_i) - \overline{C}_2}{\sigma_2},$$

where E_{1i} and E_{2i} are the normal deviates of the two cumulative proportions at t_i. Solving both expressions for $f(t_i)$, we obtain,

$$f(t_i) = \sigma_1 E_{1i} + \overline{C}_1$$

and

$$f(t_i) = \sigma_2 E_{2i} + \overline{C}_2.$$

Setting the two resulting expressions equal and solving for E_{1i}, the result is:

$$E_{1i} = \frac{\sigma_2}{\sigma_1} E_{2i} + \frac{\overline{C}_2 - \overline{C}_1}{\sigma_1}. \tag{1}$$

Thus, if the E values for the pairs of proportions at all values of t are plotted against each other, the result is a linear relation with the slope estimating σ_2/σ_1 and the intercept estimating $(\overline{C}_2 - \overline{C}_1)/\sigma_1$. In previous papers in this series (Grice, 1972a), we have termed such a plot a response evocation characteristic (REC). Parameter estimates are made by fitting a mutual regression line, recognizing the presence of error in both sets of measures, e.g., Grice, Hunt, Kushner, and Morrow (1974). If, in fact, a linear relation is obtained, it implies that the data conform to the model of Figure 1, although it does not necessarily imply that they may not conform to other models as well. If $\overline{C}_1$ is taken as the origin of the scale and σ_1 as its unit, Equation 1 reduces to

$$E_{1i} = \sigma_2 E_{2i} + \overline{C}_2. \tag{2}$$

On this scale, the slope and intercept directly estimate σ_2 and $\overline{C}_2$.

Estimates of $f(t_i)$ are given directly by the deviates of the original cumulative proportions, and the form of the function may be described by curve fitting procedures. One approach is to convert the estimates from C_2 to the scale of C_1 by the transformation of Equation 2 and use the mean at each value of t as the estimate. On the other hand, if a model has been derived for one condition only, the resulting function may be converted to the scale of another by Equation 2, or its inverse if appropriate, and goodness of fit to the new condition evaluated. This procedure will be used later in this paper. When an REC is used to relate a new condition to a computed function derived from another, conventional unidirectional linear regression is appropriate, treating the computed function as error free.

This model is based primarily on Thurstone's (1925) original model for the scaling of abilities. Its conceptual as well as mathematical relationships to signal detection are also apparent. However, there are no implications of a metric or decision process based on likelihood ratio.

The function, $f(t)$, obtained from the cumulative distribution describes only total excitatory strength. However, with appropriate experimental designs and additional scaling procedures, it is possible to analyze it further into functions which may be taken as descriptive of the component psychological processes. In general agreement with Luce and Green (1972), and McGill (1963), SRT is conceptualized as dependent upon stimulus detection. In our analyses, the growth of excitatory strength is described by a single negatively accelerated function indicating the growth of sensory strength (V). When stimulus intensity is varied, there results a family of such functions, the rate of growth depending upon intensity (Grice et al., 1974). The latency distributions of human eyelid conditioning may be described by two component functions. The first of these describes the growth of sensory strength, the rate depending upon CS intensity. The second function, which describes the growth of associative strength, begins at a later time and tends to be inflected. The amount of growth depends upon the number of conditioning trials, and

the rate is inversely proportional to the length of the interstimulus interval (Grice 1972a, 1977).

The situation seems to be somewhat more complicated for DRT of the single response or c-reaction type. Here there is evidence for two kinds of associative processes, positive and inhibitory, which follow different time functions. Short-latency correct responses appear to be entirely under the control of the sensory strength function, yet false alarms may be associatively inhibited with increasing effectiveness as time passes. Thus, a low false alarm rate may be achieved by responding on the basis of detection if the criterion is sufficiently high to permit the growth of associative inhibition. However, there is also a later function indicating the growth of associative strength to the positive stimulus. The situation is further complicated by the fact that subjects may adopt strategies which emphasize the use of either of these associative processes. In any case, the speed-accuracy tradeoff is accomplished by adjusting the level of the criterion (Grice, Hunt, Kushner, & Nullmeyer, 1976). The lower the criterion and the faster the response, the greater is the probability of error. Both the ability to inhibit errors below the level of pure detection and positive associative strength increase with time since stimulus onset. Thus, both processes result in improved accuracy with higher criterion levels. Variable criterion theory has not yet been applied to CRT, and that is the task of the present paper.

Application of the Variable Criterion Principle to CRT

The model as presented in Figure 1 does not apply directly to the cumulative distributions of CRT in other than error-free performance. We shall limit our treatment here to the two-choice situation. In this situation, there are two responses, but only one response per trial. The trial is terminated by the first response to occur, whether it is a correct response or an error. It is our conception that there is a separate function of time describing the growth of excitatory strength to each response contingent upon the presentation of a given stimulus. These functions are continuous, and when the trial is terminated, it is possible that the alternative response may have reached criterion later had the competing response not occurred. Thus, the cumulative CRT distributions depend upon two factors: the growth of excitatory strenth for each response separately and the effect of response competition. Our objective is to obtain estimates of the growth of the excitatory strength of each response with the effect of competition excluded. A trial is then terminated when the first response reaches criterion.

The basic specific assumption for CRT is that there is a separate criterion distribution for each response, and that these distributions are inde-

pendent. This leads to the rather simple conclusion that the probability density of the correct response reaching criterion at t_i, independently of response competition, is equal to the conditional probability of reaching criterion given that an error has not previously occurred. This latter probability may be estimated directly from the empirical RT distributions. More formally, we now consider the situation in which a stimulus is presented and define the theoretical values relevant to the evocation of the correct response. Let $y(c_t)$ be the theoretical probability density of the correct response reaching criterion at t_i, depending purely on the stimulus input and the criterion distribution of that response. This would also be the probability density of actually reaching criterion in error-free performance. The integral of $y(c_t)$ estimates $f(t)$ according to the model of Figure 1, but the cumulative latency distribution of correct responses does not provide these estimates if errors are present. Let $y(c_t \cap \overline{e}_{<t})$ represent the theoretical probability density of the correct response reaching criterion at t_i in the absence of an error with latency less than t_i. This is also the density of actually reaching criterion in performance with errors present, and is estimated directly from the RT distribution for correct responses. Let $p(e_{<t})$ and $p(\overline{e}_{<t})$ represent, respectively, the probabilities that an error has or has not occurred prior to t_i. These values are estimated directly from the cumulative distribution of errors. The principle is that the correct response will occur at t_i with probability density $y(c_t)$ on trials when the opportunity exists. Under the assumption of independence of the two response systems:

$$y(c_t \cap \overline{e}_{<t}) = y(c_t)[1 - p(e_{<t})] \qquad (3)$$

and

$$y(c_t) = \frac{y(c_t \cap \overline{e}_{<t})}{1 - p(e_{<t})}, \qquad (4)$$

where $1 - p(e_{<t}) = p(\overline{e}_{<t})$. The numerator of this expression is the probability of the joint occurrence of the correct response reaching criterion at t_i and the absence of a prior error, and the denominator is the probability of the absence of a prior error. Thus, Equation 4 is also the conditional probability of the response reaching criterion at t_i, given that an error has not previously occurred. Under the assumption of independence, this is the proper estimate of $y(c_t)$. Substituting the empirical estimates of the terms of Equation 4, the estimation expression is as follows:

$$\hat{y}(c_t) = \frac{P(c_t)}{1 - P(e_{<t})}, \qquad (5)$$

where $P(c_t)$ is the proportion of trials on which the correct response occurred at t_i, and $P(e_{<t})$ is the proportion of trials on which errors occurred before t_i.

The cumulative probability that the strength of the correct response is equal to or greater than criterion, dependent purely on stimulus input and the criterion distribution and independent of competing errors, is given by

$$p_{cum} = \int_0^t y(c_t)dt. \qquad (6)$$

Normal deviates of p_{cum} are the values of $E_c = f(t)$, which are required for the model of Figure 1. In obtaining estimates from experimental data, calculations are based on proportions for brief, finite intervals and summated over these intervals. It is therefore desirable to include in Equation 5 a correction for the small effect of competition within the interval under the condition of independence. With this correction included, the finite expression for Equations 5 and 6 combined becomes

$$\hat{p}_{cum} = \Sigma_{t=0} \left\{ \frac{P(c_t)}{[1 - P(e_{<t})] - .5 P(e_t)} \right\}, \qquad (7)$$

where $\hat{p}_{cum}$ is the estimate of cumulative probability, $P(e_t)$ is the proportion of trials on which an error occurred within the interval, and the other terms are as previously defined. Similar computations are also made in order to obtain $E_e = f(t)$ for errors. These are made in exactly the same way, by interchanging the symbols for correct responses (c) and errors (e) in Equations 3 through 7. Equation 7 is applied to each of the two responses, correct responses and errors, beginning with the first interval with nonzero frequencies. It is then applied successively to each interval throughout the distributions, and $\hat{p}_{cum}$ is successively summated for each response. Normal deviates of the successive values of $\hat{p}_{cum}$ provide the estimates of $E_c = f_c(t)$ and $E_e = f_e(t)$. Strictly, Equation 7 applies separately to each stimulus-response combination. However, if the RT distributions are essentially the same, data for the two stimuli may be combined before applying Equation 7. In this case, the two responses are identified simply as "correct" and "error." This is the situation which obtains in the data of the present analysis.

In order to evaluate the fit to empirical data of a model derived in this way, calculated distributions must be obtained from the theoretical E functions. For correct responses, the term $y(c_t \cap \overline{e}_{<t})$ directly predicts the probability densities of the RT distribution, and the calculations are based on Equation 3. First, the theoretical cumulative proportions of Equation 6 are obtained by taking the normal function of values of the theoretical functions for $E_c = f_c(t)$ and $E_e = f_e(t)$ at successive brief intervals.

Values of $y(c_t)$ and $y(e_t)$ are then obtained for each interval by successive subtraction of the cumulative probabilities. Equation 3 for correct responses, including the correction for competition within the finite interval, becomes

$$y(c_t \cap \overline{e}_{<t}) = y(c_t)[1 - p(e_{<t})] - .5$$

$$\{y(c_t)[1 - p(e_{<t})]\} \qquad (8)$$

$$\{y(e_t)[1 - p(c_{<t})]\}$$

The first term in Equation 8 is simply Equation 3. The correction term is one-half the probability that both responses will reach criterion in the interval. The equation is adapted to the calculation of the error distribution by interchanging the c and e symbols. The correction term is the same. The predicted cumulative CRT distributions are the summations of the values of Equation 8 over all intervals. It is first applied to both responses for the earliest interval in which the values of $y(c_t)$ or $y(e_t)$ are nonzero in the number of decimal places desired. It is then computed successively for all intervals. Summations are made after each interval in order that the values of $p(e_{<t})$ and $p(c_{<t})$ may be continuously available for the next calculation.

Current RT theories are generally theories of the behavior of the individual subject, and analyses are limited in this way. Under the assumption of a normally distributed criterion from trial to trial, the present model is also a theory of the individual subject. However, we believe that a more adequate theory should be able to deal not only with individual data, but also group data and individual differences. Based on evidence that the criterion parameters constitute the major source of individual differences (Grice, 1968, 1972a), we believe that variable criterion theory has this potential. There are alternative ways of approaching this problem, but the most simple is to apply the analysis to the mixed distribution of all subjects and trials combined. While this approach is unconventional, there are clear advantages. First, there is the greater generality of theoretical functions or laws based on group as opposed to individual data. Second, very reliable data and regular functions are obtainable in this way. Because of the large N readily obtainable, the data are substantially more regular than those obtained from even highly practiced individual subjects. Third, it provides a method for identifying and measuring the sources of individual differences among the subjects who comprise the group. Fourth, it is possible to obtain a substantial body of stable data all based on the same level of practice. Fifth, it is possible to compare two or more experimental conditions in comparable groups without the complication of using subjects who experience more than

one condition. Finally, but of course not decisive factors, are the matters of relative simplicity of analysis and economy of data collection procedures. There is the limitation that parameter estimation or goodness of fit procedures involving independence are not applicable. Also, because of its unconventional and relatively unprecedented nature, this use of the mixed distribution requires further technical justification.

It is our assumption that the mixed criterion distribution will sufficiently approximate normality that distortions will not be introduced by the use of normal scaling procedures. Within the scaling literature, Bock and Jones (1968) have presented a model for multiple judgments by each subject in paired comparisons, and have shown the mixed distribution of the discriminal process to be normal under conditions of normality both within and between subjects. However, their model treats the within-subjects error distribution as a single distribution, which is equivalent to assuming homogeneous variance across subjects. This almost certainly does not apply in the case of individual criterion distributions in RT, so the analysis must be carried further. Following Bock and Jones, we apply a three-component model in which a value of the criterion consists of a value for the experimental condition, the general mean; a value for the subject, the subject mean; and a value specific to the trial. In deviate form, this becomes

$$x_{ki} = s_k + \varepsilon_{ki}, \tag{9}$$

where x_{ki} is the value of the criterion of the kth subject on trial i, the general mean is zero, s_k is the deviation of the mean of the distribution of the kth subject from the general mean, and ε_{ki} is the deviation from s_k on trial i. From this it may be shown that the variance of the mixed distribution is

$$\sigma_x^2 = \sigma_s^2 + \frac{\sum_1^N \sigma_{\varepsilon k}^2}{N}, \tag{10}$$

where σ_x^2 is the variance of the mixed distribution, σ_s^2 is the variance of the subject means, $\sigma_{\varepsilon k}^2$ is the variance of the distribution of the kth subject, and N is the number of subjects. Under the assumption of normality within subjects, it may be shown further that the third moment of the mixed distribution is

$$\mu_3 = \frac{\sum_1^N s_k^3}{N} + \frac{3\sum_1^N (s_k \sigma_{\varepsilon k}^2)}{N}. \tag{11}$$

Thus, if estimates of the subject means and variances are available, estimates of skewness, which involve only the assumption of normality within subjects,

may be computed from Equations 10 and 11. If normality is further assumed between subjects, the third moment reduces to

$$\mu_3 = \frac{3\sum_1^N (s_k \sigma_{\varepsilon k}^2)}{N}. \tag{12}$$

This is three times the covariance of the subject means and variances, and, in the absence of such correlation, the mixed distribution will be symmetrical. However, even if a moderate correlation does exist, the covariance will be small in relation to the variance of the distribution and only minimal skewness may be expected.

Also from Equation 9 and under the assumption of normality within subjects only, the fourth moment of the mixed distribution may be shown to be

$$\mu_4 = \frac{\sum_1^N s_k^4}{N} + \frac{6\sum_1^N (s_k^2 \sigma_{\varepsilon k}^2)}{N} + \frac{3\sum(\sigma_{\varepsilon k}^2)^2}{N}. \tag{13}$$

Thus, if estimates of the subject means and variances are available, an estimate of μ_4 may be obtained and a measure of kurtosis computed from Equations 10 and 13. With the additional assumption of normality between subjects, the fourth moment may be further reduced to

$$\mu_4 = 3(\sigma_s^2 + \frac{\sum_1^N \sigma_{\varepsilon k}^2}{N})^2 + 6cov_{s^2 \sigma_{\varepsilon}^2} + 3\sigma^2_{\sigma_{\varepsilon k}^2}. \tag{14}$$

The measure of kurtosis, β_2, is

$$\beta_2 = 3 + \frac{6cov_{s^2 \sigma_{\varepsilon}^2} + 3\sigma^2_{\sigma_{\varepsilon k}^2}}{(\sigma_s^2 + \frac{\sum_1^N \sigma_{\varepsilon k}^2}{N})^2}. \tag{15}$$

The covariance is between the subject variances and the squared deviations of the subject means from the general mean. It seems quite improbable that this relation would be other than of zero order. The term

$$\sigma^2_{\sigma_{\varepsilon k}^2}$$

is the variance of the subject variances, and will be nonzero except in the case of homogeneous variance. Since $\beta_2 = 3$ for normal kurtosis, it may be inferred that the mixed distribution will tend to be somewhat leptokurtic. However, two circumstances mitigate the importance of this conclusion. First, the variance of the subject variances is likely to be small in relation to the squared variance of the mixed distribu-

tion, so that the departure from normality will be small. Second, a moderate degree of leptokurtosis will introduce negligible distortion in a scaling solution based on the normal model. Thus, the major analyses of the present investigation are based on the mixed distributions, and the degree of normality is evaluated by means of Equations 10, 11, and 13. However, an additional solution which is not based on the mixed model is also provided in order that the two approaches may be compared. It should also be noted that, based on Equation 1, the presence of linear RECs between conditions provide an additional way of evaluating the applicability of the normal model of Figure 1.

The following theoretical investigation is based upon four experiments, all using the same auditory stimuli. The primary theoretical analysis is applied to a speeded CRT experiment in which a substantial error rate and the speed-accuracy tradeoff were obtained. It is our belief that such an experiment yields the most information about information processing dynamics. A second CRT experiment, with accuracy instructions, serves as a test of the theory derived in a condition which should differ only with respect to the criterion parameters. An SRT experiment is included solely to obtain the function describing the growth of sensory strength that is required for the CRT analysis. A DRT, c-reaction, experiment is also included as the source of a function describing the growth of associative inhibition, which is also used in the CRT analysis.

METHOD

Subjects
The subjects were 95 female students from a university course in introductory psychology, and received course credit for their participation. There were 24 subjects in each of the CRT experiments, 22 in the DRT experiment, and 25 in the SRT experiment.

Apparatus and Procedure
RT responses were performed by depressing, with the index finger of each hand, two conventional telegraph keys located conveniently on a small table before which the subject was seated. Also before the subject were three 3.5 × 7.5 cm milk-glass windows, arranged vertically, which could be illuminated by 7-W incandescent bulbs. The top window was illuminated in white and served as the warning signal. This signal was .5 sec in duration and appeared in an irregular order, 1.0, 1.5, or 2.0 sec before the RT signal in all conditions. The second window, illuminated in red, was lighted immediately following each incorrect response. The third window contained the word SLOW and was illuminated in yellow. This was used in the deadline experiments only, and was lighted at the conclusion of each response greater than the deadline. The RT signals were 100-dB SPL tones of 1,000 and 1,300 Hz, presented in headphones. They were switched with rise and decay times of 10 msec and were response terminated. Programming and timing were controlled by paper-tape readers in combination with solid state logic devices driven by a precision time base. The intertrial interval in all conditions was 6 sec. Responses were recorded to an accuracy of 1 msec by a digital counter.

The four experiments were as follows:

CRT, accuracy. Subjects were instructed to press the right key to the high tone and the left key to the low tone as fast as possible while avoiding errors. There were 264 trials. The first 4 trials were used to introduce the tones and the required responses, and the following 20 were treated as practice. Data analyzed were 120 trials with each stimulus, presented in an irregular order. This experiment was conducted first in order to provide a basis for selecting the deadline in the speeded condition.

CRT, deadline. This experiment was identical to the accuracy experiment except for the addition of a deadline contingency. It was the intention to produce sufficient speeding to induce a moderate level of the speed-accuracy tradeoff. The deadline selected was the general third quartile of the accuracy condition. Subjects were instructed to be as accurate as possible while meeting the deadline. The deadline signal was presented at the conclusion of each response which exceeded it.

SRT. There were 264 trials, half with each of the two tones in an irregular order. The first 24 trials were treated as practice only. In addition, there were catch trials, amounting to 10% of the RT trials, on which no tone was presented. Subjects were instructed to respond as fast as possible, but to avoid responses on catch trials. Only the right key was used.

DRT. Only the right key was present, and subjects were instructed to respond to the high tone only. Following the 24 practice trials, there were 120 presentations of each tone. The duration of the low tone was 1.5 sec if a false alarm response did not occur. A deadline was also employed which was the general third quartile of a comparable experiment conducted under accuracy instructions.

RESULTS AND THEORETICAL ANALYSIS

Summary Statistics

The distributions for the two stimulus-response combinations in both of the CRT experiments turned out to be essentially the same. The median latencies for correct responses differed by just 4 msec in the deadline experiment and by 5 msec in the accuracy experiment. None of the corresponding values of the first and third quartiles differed by more than 4 msec. For this reason, data for the two stimuli have been combined and summary statistics are presented for correct responses and errors. For the deadline experiment, the median latency for correct responses was 351 msec and the interquartile range (Q) was 116 msec. For errors, the median was 302 and Q was 130. For correct responses in the accuracy experiment, the median was 383 and Q was 119. For errors, the median was 344 and Q was 108. Thus, in both experiments, we obtained the common finding that errors are faster than correct responses. The error rates were .098 in the deadline experiment and .035 in the accuracy experiment. In the SRT experiment, the median latency was 236 and Q was 80. In the DRT experiment, the median for correct responses was 275 and Q was 78. For errors (false alarms) the median was 224 and Q was 65. The false alarm rate was .049.

Analysis of the DRT Experiment
An analysis of the DRT experiment is necessary

since the associative inhibition function obtained from this analysis will be used in the CRT analysis. The first step was to obtain from the simple RT data a function describing the temporal growth of sensory strength (V). Both stimuli were included in this experiment in case it was necessary to obtain a separate function for each stimulus. Since the two distributions were essentially the same, they were combined and a single function was determined. The distribution was first expressed as a cumulative probability distribution at 10-msec intervals, and these proportions were then transformed to normal deviates. As in previous analyses, this produced a smooth, negatively accelerated function. Since SRT is assumed to depend only upon sensory detection in relation to the criterion distribution, this function gives the expected values of $E_t = (V_t - \overline{C})/\sigma$, where $\overline{C}$ is the mean of the criterion distribution and is its standard deviation. Thus, these values estimate V as a function of time, measured from the mean of the criterion distribution, in units of σ. This function was quite precisely described by the relation:

$$V = 2.283 - 13.583e^{-.00758t}, \qquad (16)$$

with t in milliseconds. This function is on the scale of the SRT experiment, where the unit is the σ of the criterion distribution and zero is the mean, $\overline{C}$. Thus, the function has a value of zero at the time of the median RT when it has exceeded the criterion on half of the trials. The asymptote, 2.283, is an empirical value obtained from the curve fitting procedure. The intercept, -11.3, also derives from the curve fitting procedure, but is so long before the occurrence of any response that it is not psychologically meaningful. This problem and the possibility of an irreducible minimum RT have been discussed elsewhere (Grice et al., 1976, p. 445). The three-decimal-place accuracy used in these analyses is useful in the description of the present sets of regular data. It is not implied that this degree of accuracy applies to the statement of general laws. ·

The cumulative DRT distributions for both correct responses and false alarms were also converted to normal deviates at 10-msec intervals. The resulting function for correct responses was then plotted, at 10-msec intervals, as an REC against corresponding values of V computed from Equation 16. This REC relating the growth of E in DRT to that in SRT is presented in Figure 2. This extremely linear relation is described by the following least squares regression line:

$$E_D = 1.442 \, V_S - .870. \qquad (17)$$

The value of r^2 is .9996. This relation has an obvious implication in terms of the model of Figure 1 and

Equation 1. The DRT group was responding on the basis of the same function describing the growth of E as used in SRT, but with an increased mean criterion level $(\overline{C})$ and decreased variability (σ). The values of σ and $\overline{C}$ of SRT on the DRT scale are given by Equation 17. The values of σ and $\overline{C}$ of DRT on the SRT scale are given by the inverse function, i.e., $\sigma_{Disj} = .693$ and $\overline{C}_{Disj} = + .603$ when σ and $\overline{C}$ for SRT are one and zero. Figure 1 was actually constructed from these data. C_1 is the SRT criterion distribution and C_2 is the DRT distribution. Locations of the lines indicating f(t) were computed from Equation 16 at equally spaced intervals of time. The scale is that of SRT. The fit of Equation 17, which corresponds to the inverse of Equation 1, indicates that the correspondence of the data to the model is nearly complete. The spacing of the points along the fitted line indicate the growth of excitatory strength with respect to time at 10-msec intervals. When read on the abscissa, they are the calculated values of Equation 16 for SRT. When read on the ordinate, they are the values for DRT. The degree of linearity indicates that the two functions are the same except for the linear transformation representing the different parameters of the criterion distributions in the two experiments. These scale values are directly related to the empirical response probabilities by the normal function.

These DRT data represent a pure case of what Grice et al. (1976) termed the inhibitory strategy.
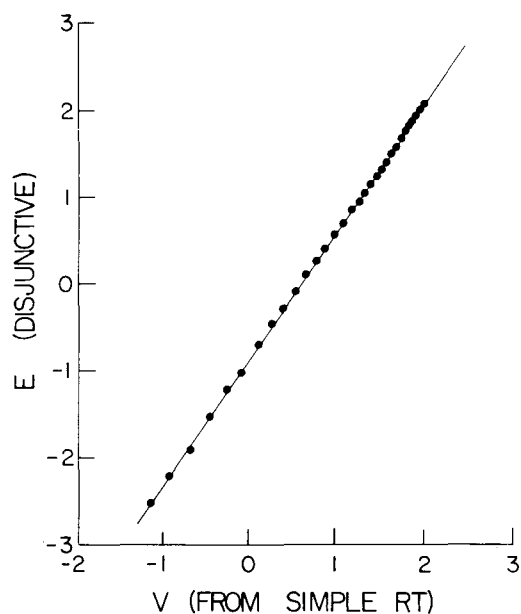

Figure 2. REC relating the growth of excitatory strength of correct responses in the DRT experiment to calculated values of the function describing the growth of sensory strength (V) derived from the SRT experiment. Zero points are at the means of the two criterion distributions. Points are at 10-msec intervals.

Correct responses are based entirely on the growth of sensory strength (detection information) and the criterion distribution. Excitatory strength for correct responses at any time following stimulus onset is simply

$$E_c = (V_t - C)/\sigma, \qquad (18)$$

where V_t is the value of Equation 16 at time t, and the criterion, C, is a normally distributed random variable with mean $\overline{C}$ and standard deviation, σ. The expected value at any time is

$$E_c = (V_t - \overline{C})/\sigma. \qquad (19)$$

When V_t is transformed to the DRT scale by Equation 17, $\overline{C} = 0$ and $\sigma = 1$. Thus, the expectation of E_c is equal to V_t. On any given trial, the value of E_c is equal to V_t minus a value of the normal random variable with mean zero and standard deviation one. Since, at any time, V_t is a constant, E_c is also normally distributed. The cumulative probability of response by any time is given by

$$P_{ct} = \int_{-\infty}^{E_{ct}} \Phi, \qquad (20)$$

where P_{ct} is the cumulative response probability, E_{ct} is the value of excitatory strength of the correct response at time t, and Φ is the unit normal function with mean zero and standard deviation one. This reasoning applies to all subsequent functions for the composition of E which are presented in terms of their expectations. As mentioned above, this model attributes all empirical RT variability to criterion variation. It may also be noted that for a fixed value of the theoretical parameter, σ, RT variability will increase with increasing levels of $\overline{C}$ because of the negatively accelerated growth of E. The form of the RT distribution is determined jointly by the growth of E and the normal criterion distribution.

The ability to reduce false alarms below the rate predictable on the basis of detection depends upon the temporal growth of associative inhibition. The expectation of excitatory strength for false alarms at any given time after stimulus onset is given by

$$E_{FA} = (V_t - I_t - \overline{C})/\sigma. \qquad (21)$$

V_t is from the transformation of Equation 16, and I_t is the strength of associative inhibition. Since $\overline{C}$ is the zero point of this scale and $\sigma = 1$, I_t may be estimated by

$$\hat{I}_t = V_t - E_t, \qquad (22)$$

where $\hat{I}_t$ is the estimate of associative inhibition at each t_i, V_t are calculated values, and E_t are normal

deviates of the proportions from the empirical cumulative distribution of false alarms. These estimates were obtained at 10-msec intervals and formed a smooth inflected function describing the growth of I. It was fitted with the following Gompertz function which describes the growth of I:

$$I = 3.700\,(7.9286 \times 10^{-16})^{.9863^t}, \qquad (23)$$

with t in milliseconds. (For those unfamiliar with the Gompertz growth function,[1] the very small values of the term in parenthesis which appear in this and subsequent functions may require explanation. This term is the intercept at $t = 0$ and is technically nonzero. However, the function does not begin to reach noticeable values until much later. For practical purposes, it may be regarded as asymptotic to zero in this region.) This function and the V function on the DRT scale are presented graphically in Figure 3.

For correct responses, the growth of E is given directly by the V function. For false alarms, it is the V function minus the I function according to Equation 18. The two cumulative distributions are then given by the integral of the normal function from minus infinity to $f(t_i)$ at each point on each of the functions. The resulting calculated distributions
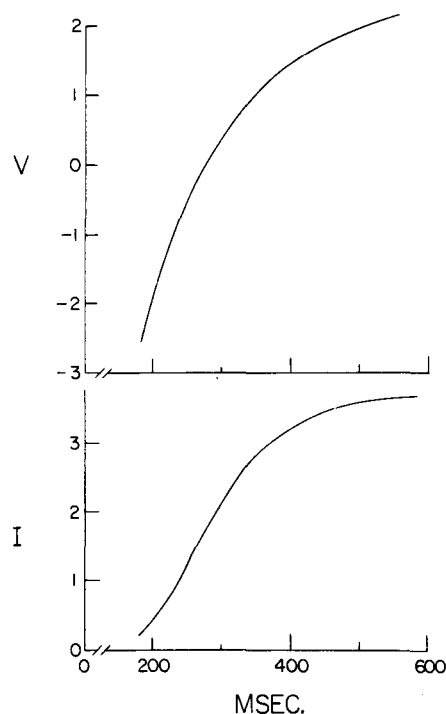

Figure 3. Component functions of the theory describing the DRT data. The upper function is the sensory strength (V) function from SRT, transformed to the DRT scale. The lower function indicates the growth of associative inhibition (I) as derived from the DRT, false-alarm data.

are presented in Figure 4 with plotted points indicating the empirical distributions. Calculated at 10-msec intervals, for correct responses the proportion of variance accounted for by the calculated values was .9997. The proportion for false alarms was .9406. This lower figure reflects primarily the lack of variance in the flat distribution. The average error of prediction was just .0026. (Aside from r^2 for linear regression, when proportion of variance accounted for is reported in this paper, it is simply the proportion accounted for by the relation $Y_{obtained} = Y_{calculated}$.) This model for the inhibitory strategy in DRT, which was derived by Grice et al. (1976) from individual subject data, also is seen to apply to the present group data.

Analysis of the CRT Deadline Experiment

The first step of the CRT and analysis is to obtain functions describing the growth of E as it depends only upon stimulus input and criterion variability as explained in the introduction. Since the distributions for the two stimuli were essentially the same, the data were combined, resulting in two distributions, one for correct responses and one for errors. The cumulative probability ($\hat{p}_{cum}$) that each of the responses would be equal to or greater than criterion by t_i was then calculated by Equation 7 at 10-msec intervals. The theoretical analysis was limited to points under 600 msec. These estimated probabilities were then transformed to the scale of 'E by transforming them to normal deviates. The resulting functions are indicated by the data points of Figure 5, plotted here at 20-msec intervals only. The functions are very regular, consisting of an initial nega-
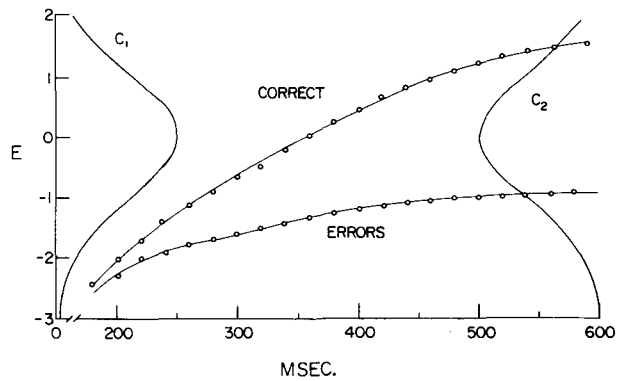
Figure 5. Functions indicating the growth of excitatory strength (E) for correct responses and errors in the deadline CRT experiment. Points are plotted at 20-msec intervals only. The two normal functions, C_1 and C_2, indicate the identical criterion distributions for the correct and incorrect responses. The smooth curves were calculated from functions derived by the variable criterion model. Theoretical cumulative probabilities that a response will have reached criterion by t_i, with the effect of response competition excluded, are given by areas of the criterion distributions below the functions at t_i.

tively accelerated phase, followed by a relatively linear section, and then resuming negative acceleration. Also included in Figure 5 are two normal functions representing the criterion distribution of the two responses. These two distributions are the same, with equal means and variances, but this would not be true in an experiment where response bias existed. It may also be noted that if there were stimulus bias or if the stimuli differed in intensity, there would be a separate pair of E functions for each stimulus. Here, the probability that either of the responses will be equal to or greater than the criterion by t_i is given by the proportion of the area of the normal function below either of the functions at t_i. But, of course, a trial is terminated when either response reaches criterion and occurs. Errors occur on these trials when the incorrect response criterion is sufficiently lower than that of the correct response so that the error function reaches criterion first. For example, if on a given trial the error criterion was -1.6 and the correct response criterion was at the mean, 0, the trial would end with an error of about 300 msec latency. The correct response would not have reached criterion until about 360 msec.

The next goal of the analysis is an attempt to provide a description of these functions in terms of the growth of component processes which determine them. The first step was to determine if the early portion of either the correct response function or the error function is determined purely by the growth of sensory strength. If this were true for errors, a form of fast-guess interpretation would be supported in which early errors were dependent only on detection and correct responses also depended
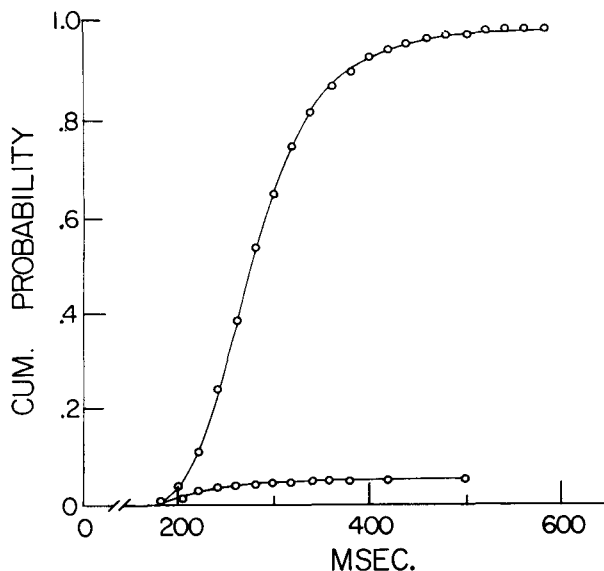

Figure 4. Calculated and obtained cumulative distributions for correct responses and false-alarms in the DRT experiment. Data points are at 20-msec intervals.

upon a positive associative component. On the other hand, if it were true for correct responses, the inhibitory interpretation would be supported. The hypothesis is evaluated by examining RECs relating the growth of E to the growth of V in order to determine if an early linear relation exists. The hypothesis was supported for correct responses, but not for errors. Figure 6 shows the REC relating the correct response E values to calculated values of the V function from SRT at 10-msec intervals up to 600 msec. There is, in fact, a highly linear relation up to 280 msec. Fitted to this portion of the REC only, the least squares regression line is

$$E_{CRT} = .839V - 1.478. \qquad (24)$$

The value of r^2 for the fit in this region is .9990. In terms of the model of Figure 1, this implies that, until 280 msec, the growth of E was determined entirely by sensory strength (V), but with a substantially higher mean criterion level and somewhat increased variability. The slope and intercept of Equation 24 give σ and $\overline{C}$ of the SRT experiment on the CRT scale. When Equation 16, the V function from SRT, is transformed to the deadline CRT scale by Equation 24, it describes the deadline E function up to 280 msec and the entire transformed function is used in the subsequent portions of the analysis.

Beyond 280 msec, scale values for the CRT condition begin to go above the extrapolation of the linear relation to Figure 6 in an orderly and curvilinear fashion. When plotted as a function of time, the residuals form a smooth function which begins gradually, grows with positive acceleration, and then becomes inflected to negative acceleration. This change from linearity is an extremely regular process, and the inference is quite compelling to us that it describes a second process which develops after the initial dependence of E purely on sensory strength. As in the case of a similar picture with DRT (Grice et al., 1976), we assume this second process to be associative strength relating the stimulus and the correct response. In both the models for DRT and conditioning (Grice, 1977), the combination of the sensory and associative components has been the simple additive one:

$$E_t = (V_t + A_t - \overline{C})/\sigma, \qquad (25)$$

where A_t is the associative strength at each t_i and E_t is the expected value of excitatory strength. Since $\sigma = 1$ and $\overline{C} = 0$, associative strength is estimated by:

$$A_t = E_t - V_t, \qquad (26)$$

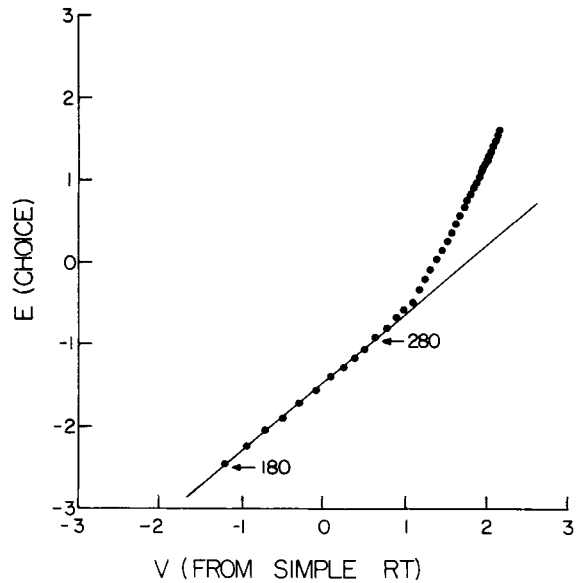where E_t are now the normal deviates of the prob-

Figure 6. REC relating correct response E values of the CRT deadline experiment to corresponding calculated values of the sensory strength (V) function from the SRT experiment. Points are at 10-msec intervals up to 600 msec. The regression line is fitted to points up to 280 msec.

abilities estimated by Equation 7, and V_t are calculated from Equation 16 transformed to the CRT scale. These estimates, which are also the residuals of Figure 6, were obtained at 10-msec intervals from 280 to 600 msec. The result was a smooth inflected function which was fitted with the following Gompertz growth function:

$$A = 1.400(1.0009 \times 10^{-52})^{.9879^t}, \qquad (27)$$

where t is in milliseconds. Indicating the regularity of the function obtained by Equation 26, the proportion of the variance of the estimates of A accounted for by Equation 27 is .998. This function, which shows the growth of associative strength with respect to time, is presented graphically in the second panel of Figure 7, along with the other component functions of the model. E as a function of time is then the sum of V and A according to Equation 25. This is the smooth function for correct responses presented in Figure 5.

We turn now to the analysis of the function describing the growth of E for errors. Since the growth of E for correct response is dependent purely upon sensory strength (V) until 280 msec, and since the correct and error functions clearly diverge during this time, the presence of associative inhibition may be inferred just as in DRT. However, in addition, there is the possibility of generalization of associative strength beginning at 280 msec when this process (A) begins to influence correct responses. The result would be a new increase in error rate at this time.

There is, in fact, a small change in curvature of the error function at this point, suggesting the influence of such generalization. Thus, the expression for errors which is analogous to Equation 25 for correct responses is

$$E_{et} = (V_t - I_t + A_{gt} - \overline{C})/\sigma, \qquad (28)$$

where E_{et} is the expectation of excitatory strength for errors at each t_i, I_t is the amount of associative inhibition, and A_{gt} is the amount of generalized associative strength to the incorrect response. The obvious source of the function for I_t is the one obtained from the DRT experiment. To be applicable, it must first be converted to the CRT scale. Since both CRT and DRT have been related to V from SRT, a separate REC relating the two is unnecessary. From Equations 17 and 24, the relationship may be shown to be

$$CRT = .582\,DRT - .965. \qquad (29)$$

The proportion of CRT variance up to 280 msec accounted for by this relation to DRT is .9970. CRT involved both a higher mean criterion level and greater criterion variability than DRT. Equation 23, describing the growth of I for DRT, was then transformed to the scale of CRT. Since $\overline{C}$ does not appear in this expression, only the slope constant of Equation 29, estimating the σ ratio, is used for this transformation to change the size of the unit. For CRT, estimates of I were based on Equation 28 with $\overline{C}$ = 0, σ = 1, and A_g = 0 before 280 msec. The estimates, then, were obtained by the relation.

$$\hat{I}_t = V_t - E_{et}, \qquad (30)$$

where V_t are computed values and E_{et} are the normal deviates of the probability estimates for errors obtained by Equation 7. These estimates were obtained at 10-msec intervals for the intervals below 280 msec. When these estimates were plotted against the transformed I function from DRT, the relation was clearly linear, but with a slope less than one. The fitted line passed through the origin with a slope of .751. What this means is that the function from DRT does accurately describe the development of associative inhibition in the CRT experiment but with a weighting factor less than one, i.e., .751. This is not a surprising result in view of previous findings of individual differences in this process (Grice et al., 1976). Thus, the I function for CRT was obtained by multiplying the transformed function from DRT by the weighting factor. The resulting function is shown in the bottom panel of Figure 7. It describes the increase, with the passage of time since stimulus onset, of the ability to inhibit errors below the level based upon detection.
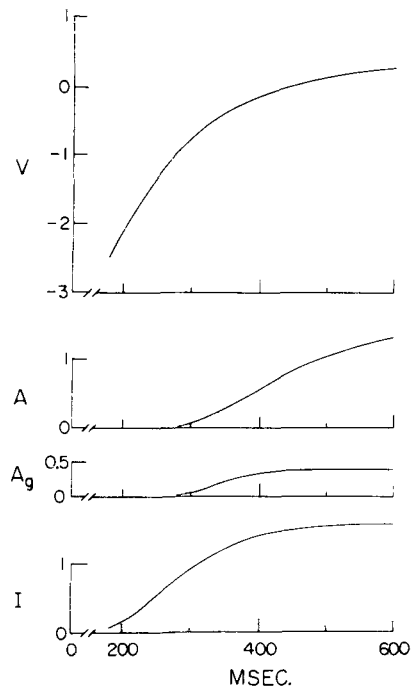

Figure 7. Graphs of the component functions of the theory describing the CRT deadline experiment. The upper function, describing the growth of sensory strength (V), was derived from the SRT experiment. The second function describes the growth of associative strength (A) for the correct response. The third function describes the growth of generalized associative strength (A_g) to the incorrect response. The bottom function, obtained from the DRT experiment, indicates the development of associative inhibition (I). All functions are on the scale of the CRT deadline experiment.

Finally, an estimate must be obtained of the growth of the generalized associative strength of the incorrect response. From Equation 28,

$$\hat{A}_{gt} = E_{et} - V_t + I_t, \qquad (31)$$

where values of I_t are computed from the transformed function from DRT. Such estimates were obtained at 10-msec intervals from 280 to 600 msec. The result was an inflected function fitted by the following Gompertz function:

$$A_g = .390\,(1.3700 \times 10^{-1.117})^{.9765^t}, \qquad (32)$$

with t in milliseconds. The function appears in the third panel of Figure 7. It is seen to have a lower limit than that of associative strength of the correct response but approaches its limit more rapidly. As a result, generalized associative strength is a decreasing proportion of the correct strength following its origin at about 280 msec. This is one of the factors in the dynamics of the speed-accuracy trade-

off. More generally, the tradeoff is described by the diverging E functions for correct responses and errors. This increasing separation is produced both by the growth of associative inhibition and by the later increasing difference between correct and generalized associative strength. Between experimental conditions, the tradeoff is also controlled by the values of $\overline{C}$ and σ.

From the three theoretical functions, excitatory strength for errors was computed at 10-msec intervals according to Equation 28. This is the smooth curve for errors shown in Figure 5. It is obvious that the theoretical function fits very well the scale values derived from the distributions. However, we shall evaluate the goodness of fit of the model with respect to the distributions themselves. From the two E functions, the implied CRT cumulative distributions for both correct responses and errors were calculated at 10-msec intervals by means of Equation 8. The resulting calculated distributions were presented in the upper panel of Figure 8 with data points plotted at 20-msec intervals only. Calculated at 10-msec intervals, the proportions of variance of the data points accounted by the calculated values is .9998 for correct responses and .9984 for errors. Of course, cumulative distributions are necessarily smooth and to some degree the goodness of the fit derives from the cumulative property. It is also desirable to observe the probability distributions directly. Calculated and obtained values of the distributions in noncumulative form are presented in the upper panel in Figure 9 for intervals of 30 msec. In this case, the proportions of variance accounted are .9940 for correct responses and .9396 for errors. While two of the component functions were obtained from other experiments, it must be recognized that this fit of the data also depends upon a total of nine parameters estimated from the data themselves. In doing this, however, we have produced the framework of a theory which describes the data with great accuracy in terms of psychologically interpretable underlying processes. The theory will gain strength when it is applied to data from which it was not derived, and which involve an additional experimental manipulation. In that case, the component functions are already defined, and only the experimentally manipulated parameters are free to vary.

Individual Differences and Evaluation of the Mixed Model Solution

From the outset (Grice, 1968), it has been suggested that differences in the criterion parameters constitute the major source of individual difference in the evocation of simple responses. For example, Grice (1972a) showed that wide individual differences in both the level and form of conditioning curves could be accounted for in this way. It is also
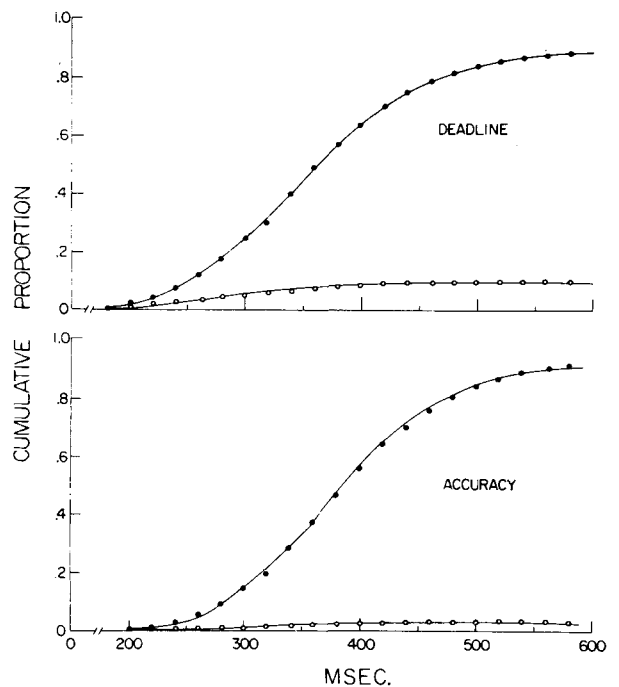

Figure 8. Calculated and obtained cumulative probability distributions for the CRT deadline and accuracy experiments. In each case, the lower, flat distributions are for errors. Data points are at 20-msec intervals.

desirable to approach this problem with respect to the present analysis. If it should turn out that the form of the theoretical functions obtained from the mixed model derive primarily from the pattern of individual differences, they would be of less interest than if they also describe laws of individual behavior. The first step of the individual difference analysis is to relate the temporal growth of E for each subject to the group function for correct response as shown in Figure 5. Thus, we obtained the cumulative distributions of correct responses and errors for each of the 24 subjects and, by means of Equation 7, computed the theoretical cumulative probabilities that the response would equal or exceed the criterion by each t_i. These computations were made at 20-msec intervals up to 600 msec, and were transformed to normal deviates giving the E functions for each subject. The 24 functions for correct responses were then related to the group function by means of RECs in the following manner. First, we identified the intervals of t in which each subject was nearest the following percentiles in the cumulative function: .05, .10, .25, .50, .75, .90, and .95. The RECs consisted of conventional linear regression between the normal deviates at each of these intervals and the group scale value at the same interval, with the subject values as the predicted variable. These relationships were highly linear. The values of r^2
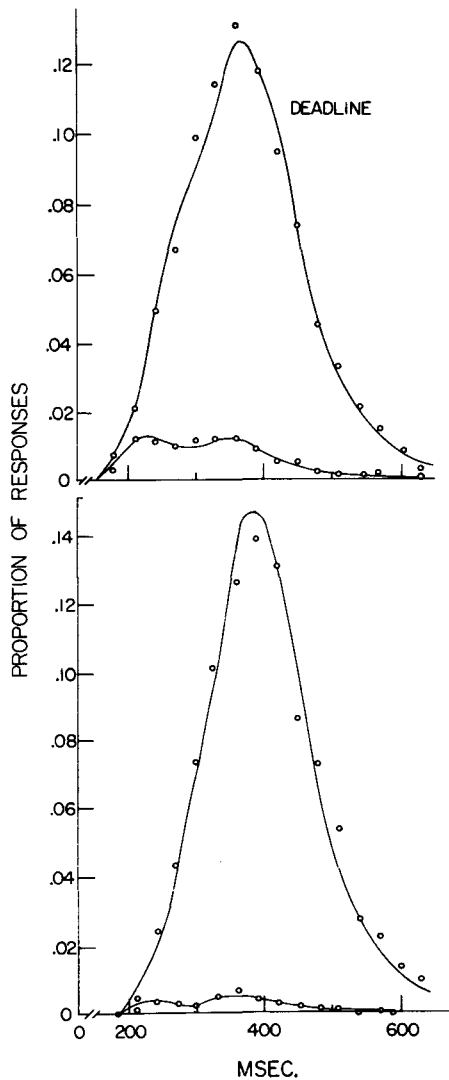
Figure 9. Calculated and obtained probability distributions for the two CRT experiments. The lower, flat distributions are for errors. Both calculated and obtained densities are for 30-msec intervals.

ranged from a low of .972 to a high of .999. The mean value of r^2 was .991, and the median was .994. Thus, on the average, the group function fails to account for less than 1% of the variance in the growth of E for individual subjects when individual differences in $\overline{C}$ and σ are estimated by the regression solution.

The slopes and intercepts of the 24 regression equations give estimates of the group σ and $\overline{C}$ on the scales of the individual subjects. However, the slopes and intercepts of the inverse of these equations estimate σ and $\overline{C}$ for each subject on the common group scale. With these estimates available for all subjects, the variance and the third and fourth

moments of the mixed distribution may be calculated by Equations 10, 11, and 13, and measures of skewness and kurtosis may be obtained under the assumption of criterion normality within subjects only. The measure of skewness, β_1, is .0004, indicating negligible asymmetry. The measure of kurtosis, β_2, is 3.51. This indicates moderate leptokurtosis, as expected. However, these two measures indicate that the mixed distribution is essentially identical to that of t with 16 degrees of freedom. Such a distribution would clearly not introduce detectable distortion in a scaling solution based on the normal model.

While the strong linear relations between the group and individual E functions clearly indicate the applicability of the model to individuals, it is desirable to carry the analysis further in order to evaluate the fit of the individual CRT distributions and to investigate the factors underlying individual differences in error rate. The individual values of $\overline{C}$ and σ have been estimated from the correct response data, but it is also necessary to estimate individual differences in the inhibitory process. This requires a two-stage process. The first stage was to convert the four functions of the model to the scales of each of the subjects by means of the 24 regression equations, and to compute the correct and error E functions for each subject. These error functions have equal weighting of I for all subjects. At this stage, it is possible to compute the proportion of variance in error rate accounted for by the criterion parameters estimated from the correct response functions only. Error rates varied from .008 to .242 with a mean of .098 and SD of .07. Calculated error rates for each subject were obtained by the application of Equation 8 at 20-msec intervals. These accounted for .484 of the individual difference variance. This is a substantial proportion, but leaves considerable variance unaccounted for.

Because of the individual differences in error rates, estimation of the inhibition weighting factor below 280 msec was not possible for all subjects. Therefore, the weights were estimated at 360 msec, a time generally not in either tail of the distributions and by which all subjects had made at least one error. The procedure was to obtain the difference between the obtained value of E_e and that calculated from the equal weighting solution, and to assume that these discrepancies were produced by differential weighting of the I function. The expression used was

$$W_I = \frac{I - (E_{eo} - E_{ec})}{I}, \qquad (33)$$

where I is the value for the subject computed from the equal weight solution, E_{eo} is the obtained value of E_e, E_{ec} is the calculated value, and W_I is the

weight. The weights obtained in this way varied from .528 to 1.805, with mean of 1.006 and SD of .289. The I function of each subject was then multiplied by the weight, and the error functions were recalculated by Equation 28. Cumulative distributions for correct responses and errors were then obtained at 20-msec intervals by Equation 8. The calculated distributions accounted for high proportions of the variance of the individual cumulative correct response distributions. These proportions varied from a low of .973 to a high of .999. The mean proportion was .993 and the median was .995. In noncumulative form, the distributions for unpracticed, individual subjects are, of course, quite irregular. However, calculated and obtained proportions for the distributions in this form were obtained at 60-msec intervals. In this form, the proportions of variance accounted for varied from .618 to .989. The mean proportion was .860, and the median was .908. The proportion of variance in error rate accounted for by the weighted solution was .923.

Alternate Solution with the Individual Difference Model

There appears to be ample reason for confidence in the solution based on the mixed model both because of its applicability to individual subject data and because of the evidence concerning the form of the mixed distribution. Previously, however, we have used another method for combining the data of a smaller number of subjects, which does not make use of the mixed model (Grice, 1972b). We also apply this procedure here for comparison purposes. This involves a complete between-subjects scaling solution under the assumption that they share a common function for the growth of E, but that they differ in their criterion parameters. These parameters are estimated by RECs relating all subjects to each other, i.e., the number of RECs is one less than the number of subjects. There are various ways of going about this, but the most simple is to do all of the scaling with respect to a single reference subject. Here, we selected as the reference subject one whose correct response E function covered, insofar as possible, the entire range of the group. We then identified for this subject the class intervals most closely approximating the same seven percentile points used in the previous analysis. RECs were then computed relating the scale values obtained by Equation 7 at these intervals to those of each of the other subjects at the same intervals. Because of lack of overlap, the RECs for two subjects were based on only six points, and that for one subject on only five. Computations were by unweighted, least squares, mutual linear regression (Grice et al., 1974). In regression of this kind, for high correlations, r^2 closely approximates the proportion of total variance of

both subjects accounted for by the mutual regression line (Grice et al., 1976). The 23 values of r^2 varied from .969 to .999, with a mean of .990 and a mean of .991. This model is the same as that of Figure 1, with 24 criterion distributions representing subjects rather than experimental conditions. This degree of linearity indicates that the model closely applies. The slopes and intercepts of the 23 regression equations give the values for $\overline{C}$ and σ for each subject on the scale of the reference subject, the values for the reference subject being zero and one, respectively. However, for convenience in comparing the result with that of the mixed model, the values were transformed to a scale with the mean of the subject $\overline{C}$s at zero and with the mean of the variances as the unit.

The goal of the individual difference analysis is to estimate the points on $E = f(t)$ as the means of the individual values when all of the individual functions are on a common scale. From Equation 3:

$$E_s = \sigma_k E_{kt_i} + \overline{C}_k, \qquad (34)$$

where E_s is on the common scale, σ_k is the σ for subject k, E_{kt_i} is subject k's scale value at t_i, $\overline{C}_k$ is the mean of the subject k's criterion distribution. Thus, any value of E is transformed to the common scale by Equation 34. For each subject, all of the values of E obtained from Equation 7 were transformed to the common scale at 20-msec intervals. At this point, we have 24 estimates of the theoretical function, $E = f(t)$, all on the common scale. The joint estimate is the mean of the 24 values at each t_i. Actually, at short and long latencies, the number is less than 24 because of the lack of full overlap of the subject distributions. Mean estimates of $E = f(t)$ were obtained in this way for both correct responses and errors. To the extent that the functions from the mixed and individual difference models are the same, their values should be equal at all values of t_i except for a linear transformation relating the two scales. The origin of the individual difference scale is at the mean of the subject means, which is also the mean of a mixed distribution and the origin of that scale. Hence, the transformation will have zero intercept. The slope of the transformation is the ratio of the size of the units. The unit of the individual difference scale is based on the mean of the subject variances, and the unit of the mixed model scale is the σ of the mixed distribution. Calculated from Equation 10, this is 1.041 in relation to the unit of the individual differences scale. Thus, if values of E from the individual differences solution are plotted against those from the mixed model at corresponding times, the relation should be linear with a slope of 1.041 and an intercept of zero.

In Figure 10, the mean values of E for correct

responses from the individual differences solution are plotted against corresponding calculated values from the mixed model solution. Points are at 20-msec intervals. The regression line is not a fitted one, but is the one with the rational slope of 1.041. The proportion of variance of the mean E values accounted for by this relation is .9977. The two functions for errors were also linearly related by the same relationship, the proportion of variance accounted for being .9927. It is of interest to note that the proportion of variance of the original data points accounted for by the correct response function presented in Figure 5 was .9998, and the proportion accounted for by the error function was .9966. Thus, in applying the functions to data points derived by the individual difference model, from which the functions were not derived, the loss in terms of proportions of variance accounted for is just .0021 for correct responses and .0039 for errors. It is our conclusion that, except for the small differences in unit size, the mixed and individual difference models produce essentially identical results. The difference in unit size is accurately described by Equation 10, which relates the variance of the mixed distribution to the variance between and within subjects. ·

Application of the Theory to the Accuracy Experiment

The theory of CRT which has been derived obviously gives an accurate description of the data as well as providing an analysis of the underlying processes. However, confidence in the theory would be strengthened if it also applies to other experiments from which it was not derived. This is especially so if a second experiment involves an additional experimental manipulation with predictable consequences. Such a manipulation is the use of accuracy instructions rather than a deadline procedure. Within this theory, accuracy may be increased in two ways—by increasing the level of the criterion and by decreasing its variability. Both of these effects will reduce the intersection of the criterion distribution with the error function and consequently reduce the error rate. Elevation of the criterion distribution will also increase the latency of correct responses since its intersection with the correct response function is later. Thus, the between-experiments speed-accuracy tradeoff is a part of the theory. The theory will be supported if it will describe the data of the accuracy experiment with the predicted changes in the criterion parameters. Whether the weighting of the associative inhibition function should be affected is uncertain, but it is recognized that this factor is at least subject to sampling error due to individual differences.

Analysis of the accuracy experiment began with the estimation from the cumulative distributions of functions for the growth of E for correct responses
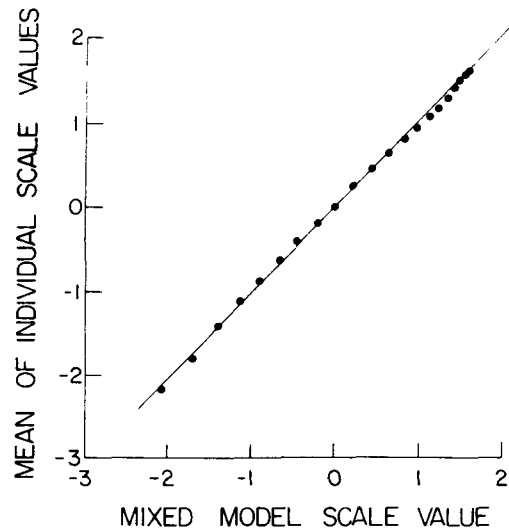

Figure 10. Mean correct response scale values obtained from the individual differences solution, applied to the CRT deadline experiment, plotted against computed scale values from the theory derived from the mixed model. Points are at 10-msec intervals. The slope of the unfitted line is the ratio of the standard deviation of the mixed, criterion distribution to the size of the unit of the individual differences scale.

and errors by means of Equation 7. For correct responses, these values were then plotted at 10-msec intervals against calculated values of the theoretical function from the deadline experiment. This relation was linear and is described by the least squares regression line

$$E_{Acc} = 1.139\,E_{DL} - .323. \tag{35}$$

The value of r^2 is .9988. This means that the proportion of variance accounted for by the theoretical function is just .001 less than it was for the deadline data from which it was obtained. This is convincing evidence that, for correct responses, the same underlying process determines the growth of E in both experiments. The slope and intercept of Equation 35 give σ and $\overline{C}$ of the deadline experiment on the accuracy scale. On the deadline scale, the parameters for the accuracy experiment are $\sigma = .878$ and $\overline{C} = +.289$. Thus, as predicted, we observe both a decrease in criterion variability and an increase in the criterion mean.

At this point, the four component function from the deadline experiment were transformed to the scale of the accuracy experiment by Equation 35 and the functions of E obtained by Equations 25 and 28. These scale values were then converted to probabilities, and calculated cumulative probability distributions were obtained by Equation 8. The calculated error rate was .047, while the obtained rate was .035. While this is not a large discrepancy, it

indicates that it would be worthwhile to estimate a new weighting factor for the inhibition function. Due to the absence of reliable scale values for errors under 280 msec, the estimate was made in the same way as for individual subjects in the previous analysis. Its value was 1.077, differing only slightly from one, as would be expected from the small discrepancy in predicted error rate. The associative inhibition function was weighted by this factor, and the E function for errors was obtained from Equation 28.

The predicted cumulative distributions for correct responses and errors were then calculated by Equation 8 using the error E function containing the weighted inhibition function. These calculated distributions are presented in the lower panel of Figure 8 with data points plotted at 20-msec intervals. Calculated at 10-msec intervals, the proportion of variance accounted for by the calculated correct distribution was .9994. For the error distribution, the proportion was .9917. Thus, the quality of the fit is of the same order as it was for the deadline experiment. The distributions in noncumulative form are also presented in the lower panel of Figure 9. Here, the proportions of variance accounted for are .9835 for correct and .7707 for errors. Of course, there is little variance to account for in the flat, slightly bimodal error distribution. Actually, the fit is surprisingly good. It may be noted that, in applying the theory to the accuracy experiment, only three parameters were estimated from the data—the two criterion parameters estimated from the correct data only and W_I estimated from the error data. All parameters of the theoretical functions defined for the deadline experiment are invariant in this application.

The final step in the analysis of the accuracy experiment is an individual difference analysis to evaluate the applicability of the theory to the individual subjects and provide evidence concerning the form of mixed distribution. RECs were computed, in the same manner as previously, relating the correct response E functions to computed values of the group function. Again, the relations were highly linear. The 24 values of r^2 varied from a low of .9798 to a high for two subjects with .9996. The mean was .9933, and the median was .9958. Thus, the theory accounts for the individual subject data to an even greater degree than it did in the case of the deadline experiment.

Estimates of $\overline{C}$ and σ were also obtained for each subject from the 24 regression equations, and from these the variance and third and fourth moments of the mixed distribution were calculated. Under the assumption of normality within subjects, these resulted in values of .0025 for β_1 and 3.094 for β_2. These values essentially describe the t distribution

with 64 degrees of freedom, which is not perceptibly different from normal.

DISCUSSION

One unusual feature of the present theory is the apparent identification of associative inhibition as an important process in CRT. While the possible role of inhibition has been previously suggested by Kornblum (1965), this is, to the best of our knowledge, the first explicit identification of the process in CRT theory. Certainly, it is possible that eventually some other explanation of the short latency separation of the correct and error distributions will be found preferable. At the present time, however, we find the evidence convincing. The fact that it is the early portion of the correct distribution rather than the error distribution which resembles SRT indicates that it is these correct responses which depend upon detection and that errors are being inhibited. This finding is also contrary to the expectation, based on fast guess theory, that errors should be SRT responses. Evidence of this kind has been consistent now, not only in the present experiments, but in previous results with DRT (Grice et a., 1976). Of course, in DRT, it is logically possible for correct performance to be entirely based on this principle, but this is not true in CRT, since each trial must end with a response. One kind of evidence suggesting the reasonability of our assumption that inhibition may be effective at very short latencies is that obtained by Lappin and Eriksen (1966). They found that, when a delayed second signal countermanded the go signal, RT need be extended only by a few milliseconds beyond the delay in order to obtain reliable inhibition.

An account of the speed-accuracy tradeoff is an integral part of the theory. The dynamics of the tradeoff are described by the divergence of the functions for excitatory strength for correct responses and errors as presented in Figure 5. Of course, the tradeoff is also affected by both parameters of the criterion distributions. If one desired a single tradeoff function, the difference between the two functions could be plotted as a function of time. Such a function would not, however, provide an account of the dynamics which produced it, and it now appears that functions of this kind are not of as much significance as commonly believed. The theory does provide an accurate description of the basic facts of the tradeoff as they appear in the present data and as they have been commonly reported in the literature. For example, the faster a response, the greater is the probability of an error. For the deadline experiment and the functions presented in Figure 5, at 200 msec the calculated conditional probability of an error, given that a response has occurred, is .39. This prob-

ability is reduced to .17 at 300 msec and .10 at 500 msec. With the parameters of the accuracy experiment, the corresponding conditional probabilities are .33, .08, and .04. The theory also predicts the common finding that the average latency of errors is faster than that for correct responses. The theory accurately predicts the median latencies which illustrate this point. In theoretical terms, this effect is produced primarily by the continued late growth of associative strength to the positive stimulus.

Another feature of this approach is the inclusion of individual differences within the theory. We suggest that it would be desirable if this were more common in theories originating within experimental psychology. When the three individual difference variables are taken into account, the theory describes the individual subject data with considerable accuracy, as well as the group data. One possible kind of future research is to investigate whether these variables represent consistent traits measured by conventional ability and personality tests. Here, we have performed one simple individual differences analysis with respect to error rate in the deadline experiment. The multiple correlation for predicting error rate from the three individual difference variables is .968. The multiple regression equation in standard score form is

$$\hat{z}_{\text{errors}} = -.097z_{\overline{C}} + .732z_{\sigma} - .481z_{W_1}. \quad (36)$$

The most interesting feature of this result is that the σ parameter, which was estimated from the correct response functions only, has the greatest weight in predicting error rate. The low weight of the criterion mean $(\overline{C})$ is attributable to the fact that there was relatively small variance in this variable.

Certainly, it is desirable to provide both experimental and conceptual comparisons of this theory with other theoretical accounts of CRT. It is our intention, however, to defer experimental and extensive conceptual comparison until further development of the theory, and perhaps others as well, provide more points of meaningful contact. Only one contrasting approach will be discussed here. This is the recent example of a random walk model presented by Link (1975). It appears to be a promising approach for analyzing relations between experimentally manipulated variables, mean correct and incorrect latencies to each stimulus, and the proportions of correct responses and errors to each stimulus within a set of trials. In this model, the subject compares by subtraction, a transduced representation of the stimulus with a psychological referent. Both are values on a psychological dimension, and the referent is bounded by represen-

tations of the two stimuli. The transduced stimulus is a random variable, and the referent may be as well. The basic random variable is the difference between them. Comparison continues sequentially over time and the difference is accumulated algebraically. The result is a random walk which terminates with a response when the difference reaches one of the two fixed criteria for the two responses. We have previously expressed and still affirm our skepticism concerning the reasonability of fixed criteria. However, one of the features of our model is achieved by indicating that response bias is achieved when the start of the random walk is not at the midpoint between the two criteria. Our intent here is not to criticize the· random walk model, but merely to explain our own preference on the basis of greater scientific optimism. The random walk model treats the events which take place between stimulus onset and the response as random processes. Variable criterion theory places variability in trial-to-trial variation of the criterion, which is treated as constant for the duration of the trial. The intratrial events are regarded as orderly sensory and associative processes which develop with time in a lawful manner. Our goal is to identify them and describe their dynamics. If such processes do, in fact, exist, they are more likely to be discovered if one looks for them than if one begins with the assumption that the intratrial events are largely random. Variable criterion theory provides a framework as well as methods for the analysis of a variety of experimental variables. We hope that such analysis will lead to a more detailed understanding of this area than has so far been achieved.

One way of expressing the results of our analysis is that we have obtained transformations of the time dimension which make the distributions Gaussian on the transformed scale. The rationale for seeking such transformations is our assumption of a normal distribution of the criterion. This is a rational assumption based upon our conception of criterion processes. Under this assumption, the RT distributions would be normal if the growth of excitatory strength were linear. However, one would hardly expect this to be the case. More reasonably, one would expect the temporal growth of such processes to be according to a principle of decreasing gain, either with or without an early recruitment phase with positive acceleration. This is the class of functions we have found. It is true that in the analysis of the deadline experiment the nature of the functions was not defined prior to the analysis. This is a part of our strategy of theory development which recognizes that theoretical processes cannot be quantitatively described prior to their discovery. The transformation which will normalize any given dis-

tribution is unique within the limits of curve fitting accuracy. On the other hand, if the transformation is to consist of component functions, a particular combination is not unique. However, given our hypothesis that early responses may be dependent on the growth of detection information only, the solution becomes unique if the implied early linearity with V is present in the data. If it were not present, the hypothesis would be rejected. Had the linearity been with errors, different theoretical conclusions would have followed. It should be emphasized that the linear relation is specific to the time dimension and the growth of V and not to the portions of the distributions involved. In the case of SRT, the linearity is in the range of cumulative proportions from .12 to .74. In terms of proportion of correct responses for the deadline experiment, the range was from less than .01 to .19. For the accuracy experiment in the time of the linear relation with V, the proportions were from less than .01 to .09. The application of the inhibition function to the error function was based on the inference that since early correct responses were based on detection, the incorrect response must be subject to early inhibitory control. The source of the function was from the DRT experiment where the evidence indicated that it was in its pure form throughout its range. If the two stimuli were sufficiently similar for the generalization of associative strength (A), it would be expected that E for error would begin to rise above the value of V-I at the time of origin of A for correct responses. This expectation was precisely confirmed.

It is conventional, and sometimes proper, to express a degree of skepticism about approaches which rely heavily upon curve-fitting. For example, with data as regular as that of these experiments, if one "peels off" one smooth function, there is bound to be another smooth function left. What this means is that goodness of fit alone is not a sufficient basis for the evaluation of such analysis. The skepticism is justified only to the extent that the analysis is arbitrary and unguided by adequate rational and theoretical considerations. When the analysis is guided by an adequate model, the approach may be powerful. In the present analysis, each step has been determined by the logical properties of the variable criterion model or by the nature of accumulated evidence. The result is a system with substantial internal consistency, which interrelates several kinds of RT experiments, and within which the component functions are psychologically interpretable. The precise confirmation of the theory in a second experi-

ment with experimental manipulation of the criterion parameters provides a strong basis for increasing confidence.

REFERENCES

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

Grice, G. R. Stimulus intensity and response evocation. Psychological Review, 1968, 75, 359-373.

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, 1972. Pp. 1-65. (a)

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

Grice, G. R. Information-processing dynamics of human eyelid conditioning. Journal of Experimental Psychology: General, 1977, 106, 71-93.

Grice, G. R., Hunt, R. R., Kushner, B. A., & Morrow, C. Stimulus intensity, catch trial effects, and the speed accuracy tradeoff in reaction time: A variable criterion theory interpretation. Memory & Cognition, 1974, 2, 758-770.

Grice, G. R., Hunt, R. R., Kushner, B. A., & Nullmeyer, R. Associative processes and strategies in disjunctive reaction time. Memory & Cognition, 1976, 4, 433-445.

Kornblum, S. Response competition and/or inhibition in two-choice reaction time. Psychonomic Science, 1965, 2, 55-56.

Lappin, J. S., & Eriksen, C. W. Use of a delayed signal to stop a visual reaction-time response. Journal of Experimental Psychology, 1966, 72, 805-811.

Lewis, D. Quantitative methods in psychology. New York: McGraw-Hill, 1960.

Link, S. W. The relative judgment theory of two choice response time. Journal of Mathematical Psychology, 1975, 12, 114-135.

Luce, R. D., & Green, D. M. A neural timing theory for response times and the psychophysics of intensity. Psychological Review, 1972, 79, 14-57.

McGill, W. J. Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), Handbook of mathematical psychology (Vol. I). New York: Wiley, 1963. Pp. 309-360.

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

NOTE

1. The Gompertz function has been used for the description of growth or developmental processes. It is of the form:

$$Y = vg^{h^T},$$

where Y is the amount of growth at time T, v is the asymptote or amount of growth at maturity, g is the intercept at T = 0 or the amount of initial growth, and h determines the rate of growth. A discussion of the properties of the function is given by Lewis (1960, pp. 80-88 and 182-185).