UC Irvine

UC Irvine Previously Published Works

Title

Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time

Permalink

<https://escholarship.org/uc/item/1h60z7fv>

Journal

Journal of Mathematical Psychology, 8(2)

ISSN

0022-2496

Author

Yellott, John I

Publication Date

1971-05-01

DOI

10.1016/0022-2496(71)90011-3

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Correction for Fast Guessing and the Speed-Accuracy Tradeoff in Choice Reaction Time¹

JOHN I. YELLOTT, JR.

University of Minnesota, Minneapolis, Minnesota 55455

In choice reaction time tasks, response latency varies as the subject changes his bias for speed vs accuracy; this is the speed-accuracy tradeoff. Ollman's Fast Guess model provides a mechanism for this tradeoff by allowing the subject to vary his probability of making a guess response rather than a stimulus controlled response (SCR). It is shown that the mean latency of SCR's (μ_s) in two-choice experiments can be estimated from a single session, regardless of how the subject adjusts his guessing probability. Three experiments are reported in which μ_s apparently remained virtually constant despite tradeoffs in which accuracy varied from chance to near-perfect. From the standpoint of the Fast Guess model, this result is interpreted to mean that the tradeoff here was produced almost entirely by mixing different proportions of fast guesses and constant (mean) latency SCR's. The final sections of the paper discuss the question of what other models might be compatible with μ_s invariance.

1. INTRODUCTION

It is well known that response latency in choice reaction time (RT) experiments depends strongly on the subject's willingness to sacrifice speed for accuracy. This paper deals with the problem of finding parameters of choice RT performance that will be invariant—for a given subject in a given task—regardless of changes in that subject's bias for speed vs accuracy. To illustrate the problem in a concrete fashion, consider the following results obtained from a well-practiced subject in two conditions of a choice reaction time experiment.² The experimental situation was identical in both conditions, except that in one case the subject was motivated (by payoffs) to respond very rapidly,

¹ I am indebted to Robert Ollman, especially for pointing out to me that Eq. 6, which is valid for the special case $\pi_1 = \pi_2 = .5$, holds more generally for any presentation probabilities in the form of Eq. 3. I thank also Jean-Claude Falmagne, David LaBerge, Stephen Link, and Saul Sternberg for helpful comments, Peter Van Gelder for technical assistance, and Joan Tonn and Betty Hamre for running the experiments. Support for this research has been provided by the University of Minnesota Graduate School, Space Science Center, and Center for Research in Human Learning, and by the U.S. Public Health Service (MH 16270).

² The data are from Experiment 2 of the present study—Subject 2002, Schedules 3 and 9 (cf. Table 3).

and in the other case to respond very accurately. In the "speed" condition the proportion of correct responses was .63 (chance was .50) and the mean latency (of correct responses) was 219 msec. In the "accuracy" condition the proportion correct was .96, and the mean latency was 324 msec. Since the psychophysical aspects of the task were constant across conditions, it seems reasonable to suppose that the speed with which the subject could extract discriminative information from the choice stimuli was the same in both cases, even though performance was so dramatically different. The problem is how to estimate parameters that reflect this invariance.

In spirit, at least, the problem here is quite analogous to that of finding a measure of stimulus detectability which will be invariant under changes in the observer's response criterion—the problem for which Signal Detection Theory provides such an elegant solution (Green and Swets, 1966). However, the solution to be considered here for the choice RT problem is quite unlike those suggested by Signal Detection Theory. It is, in fact, more closely analogous to the "correction for guessing" in high threshold detection theory. This solution is based on Ollman's (1966) Fast Guess model for choice reaction time, the simplest case of which is described in Sec. 2 below. According to this model the subject in a choice RT task makes two kinds of responses: fast "guess" responses that convey no information about stimulus identity, and slower responses which represent the outcome of a (possibly imperfect) recognition process. The latter have been called "stimulus controlled responses" (SCR's) (Yellott, 1967). On any given trial the subject has some probability of making either a guess or a SCR. The relative frequency of guesses and SCR's, and also the latency distribution of the guess responses, are controlled by the subject and can be expected to vary with instructions and payoffs. These parameters in themselves are of little interest, except insofar as they contribute to the variability of observable performance. The parameters of the SCR's, however, reflect limitations imposed by underlying sensory-motor processes, and consequently are relevant to the psychophysics of reaction time. To increase his speed the subject can make more guesses and/or increase the speed of his guesses—the latter because guess responses can be initiated during the foreperiod. To increase his accuracy, the subject can make more SCR's. But regardless of how the subject chooses to operate, it is possible to estimate the mean latency of the SCR's made over any set of trials (Yellott, 1967). Section 2 shows how this estimate is obtained.

The behavior of this estimate of mean SCR latency under changes in speed-accuracy bias is the main experimental concern of the present paper. The Fast Guess model itself does not *predict* that mean SCR latency will be invariant under changes in speed-accuracy bias, because it does not *assume* that varying his proportion of guess responses is the only way in which a subject can produce a speed-accuracy tradeoff. In fact, as will be apparent in Sec. 2, the Fast Guess model really does little more than formalize the notion that some (unspecified) proportion of the subject's responses are guesses; nothing is assumed about possible effects of speed-accuracy bias on the latency of

nonguess responses (i.e., SCR's). However, the Fast Guess model does allow us to determine experimentally whether changes in SCR latency play any substantial role in the speed-accuracy tradeoff. By examining estimates of mean SCR latency obtained from conditions corresponding to different accuracy levels—conditions involving, for example, different payoff contingencies favoring speed or accuracy—we can determine the extent to which the speed-accuracy tradeoff involves changes in the latency of nonguess responses. If estimated mean SCR latency remains invariant despite changes in speed-accuracy bias, then we may conclude that the speed-accuracy tradeoff is due entirely to changes in the frequency (and/or the latency) of guess responses. In this case, estimated mean SCR latency—in effect, the mean latency “corrected for fast guessing”—would provide a solution to the problem described in the opening paragraph of the paper.

Preliminary tests with the Fast Guess model have suggested that under certain fairly standard conditions, mean SCR latency does in fact remain approximately constant regardless of fluctuations in speed-accuracy bias (Ollman, 1966; Yellott, 1967). If it could be reliably demonstrated, this result would be of interest for two reasons. First, it would provide a surprisingly simple solution to the problem of finding a measure of RT performance that is unaffected by changes in speed-accuracy preferences. Secondly, such a result is not at all what one would expect on the basis of the widely held notion that the temporal process of stimulus identification involves an evolution through several levels of accuracy, any one of which can be the level represented in performance. Regarded simply as a device for estimating SCR latency, the Fast Guess model is not incompatible with models (e.g., Stone, 1960; LaBerge, 1963; Edwards, 1965; Laming, 1968) which suppose that the speed-accuracy tradeoff is produced by the subject's sampling more or less information about the stimulus before initiating a response. According to these models, stimulus identification consists of accumulating statistical evidence over time, and there can be many levels of accuracy between chance and maximum accuracy, each level corresponding to a larger average sample size, and hence to a longer mean latency. This amounts to assuming that there can be several levels of SCR accuracy and several corresponding SCR latency distributions. But if several such levels exist and play a role in the speed-accuracy tradeoff, then we should find that mean SCR latency estimated according to the Fast Guess model would vary with performance accuracy. The preliminary results just mentioned suggest instead that this does not happen (at least for certain kinds of highly discriminable choice stimuli), that there is only one level of SCR accuracy (the maximum), and that the speed-accuracy tradeoff is produced entirely by mixing SCR's of this sort and guesses in various proportions.

Sections 3, 4, and 5 below report three experiments designed to test these notions over a broad range of incentive conditions. Specifically, all three were designed to test the prediction that mean SCR latencies estimated according to the Fast Guess model would be constant (for a given subject and a given set of choice stimuli) regardless of

the speed-accuracy balance. Mathematical background for these experiments is provided in the next Section, which describes the simplest special case of the Fast Guess model proposed originally by Ollman (1966). This special case, which I will call the Simple Fast Guess (SFG) model, assumes that SCR's made to either choice stimulus have the same probability of being correct, and the same latency distribution whether correct or incorrect. The general case of Ollman's model involves more realistic assumptions; that model (designated FG1) is considered in Sec. 8, along with another general Fast Guess model (FG2) which also contains model SFG as a special case. For the purpose of motivating the experiments, however, the SFG model is adequate, since the formula it provides for estimating mean SCR latency has essentially the same interpretation in models FG1 and FG2.

It should be clear from what follows that the SFG model avoids making any assumptions about trial to trial changes in the subject's bias parameters. Consequently it is not a complete model for RT performance—it does not, for example, generate predictions about variances. But for the same reason, the predictions that it does generate about means and proportions will be valid no matter how the bias parameters behave from trial to trial. This is important because the behavior of these parameters appears to be very complicated and idiosyncratic (Laming, 1969).

2. THE SIMPLE FAST GUESS MODEL

Consider a choice RT experiment in which two stimuli, s_1 and s_2 , are presented randomly over a sequence of trials, with $P(s_i)$ equal to a constant π_i on every trial, independent of past events. The subject is instructed to make response r_i ($i = 1, 2$) when s_i is presented. The SFG model assumes that on any trial the subject makes either a stimulus controlled response (SCR), with probability q , or a guess response, with probability $1 - q$ (see Fig. 1). If the subject guesses, he makes response r_i with probability b_i , regardless of which stimulus is presented. The latency of each guess response is a random variable with cumulative distribution function Φ_g and mean μ_g . If the subject makes a SCR, the response is correct with probability a ($a > .5$), and its latency is a random variable with distribution function Φ_s and mean μ_s . The parameter

FIG. 1. Branching process of the Simple Fast Guess model for a single trial. The choice stimulus here is s_i ; r_i is the correct response, r_j the incorrect response

a , and the distributions Φ_s and Φ_g (and hence their means μ_s and μ_g) are interpreted as constants (for a given sequence of trials under constant experimental conditions), but to allow for the possibility of trial to trial sequential effects, q , b_1 , and b_2 are regarded as random variables ($b_1 + b_2 = 1$). Figure 1 shows the branching process of the model.

To relate the model to observables, consider first a single trial. Let p_{ij} denote $P(r_j | s_i)$, and M_{ij} the expected latency given s_i and r_j . Then the following relationships hold regardless of the distribution of q , b_1 , and b_2 , and provide a basis for estimating a , μ_g , and μ_s :

$$p_{11}M_{11} + p_{22}M_{22} = c_1(p_{11} + p_{22} - 1) + \mu_g. \quad (1)$$

$$c_1 = \frac{2a\mu_s - \mu_g}{2a - 1}. \quad (1')$$

$$p_{12}M_{12} + p_{21}M_{21} = c_2(p_{11} + p_{22} - 1) + \mu_g. \quad (2)$$

$$c_2 = \frac{2(1-a)\mu_s - \mu_g}{2a - 1}. \quad (2')$$

$$\frac{(p_{11}M_{11} + p_{22}M_{22}) - (p_{12}M_{12} + p_{21}M_{21})}{2} = \mu_s(p_{11} + p_{22} - 1). \quad (3)$$

Each equation expresses a linear relationship between response accuracy (i.e., $p_{11} + p_{22} - 1$, which ranges from zero at chance accuracy, when $p_{11} = p_{21} = 1 - p_{22}$, to 1.0 when $p_{11} = p_{22} = 1$) and a weighted average of mean latencies.

In the special case $\pi_1 = \pi_2 = .5$, Eqs. 1, 2, and 3 can be written in a simpler form. Let p_c denote the probability of a correct response, $p_e = 1 - p_c$, M_c the expected latency of a correct response, and M_e the expected latency of an incorrect response. Then when $\pi_1 = \pi_2 = .5$, $p_{11} + p_{22} - 1 = p_c - p_e$, $p_{11}M_{11} + p_{22}M_{22} = 2p_cM_c$, and $p_{21}M_{21} + p_{12}M_{12} = 2p_eM_e$. Substituting into (1), (2), and (3) yields

$$2p_cM_c = c_1(p_c - p_e) + \mu_g, \quad (4)$$

$$2p_eM_e = c_2(p_c - p_e) + \mu_g, \quad (5)$$

$$p_cM_c - p_eM_e = \mu_s(p_c - p_e). \quad (6)$$

Derivation of Eqs. 1, 2, and 3. For convenience, we assume that q , b_1 , and b_2 are discrete random variables, and that Φ_s and Φ_g are differentiable. To derive Eq. 1,

consider first the quantity $p_{11} + p_{22} - 1$. Let G denote a random variable which equals 1 if a guess occurs, zero if a SCR occurs. Then

$$\begin{aligned}
 p_{ii} &= \pi_i^{-1} P(r_i, s_i) \\
 &= \pi_i^{-1} \sum_q \sum_{b_i} \sum_G [P(r_i | s_i, G, q, b_i) P(s_i | G, q, b_i) P(G | q, b_i) P(q, b_i)] \\
 &= \sum_q \sum_{b_i} [qa + (1 - q)b_i] P(q, b_i) \\
 &= aE(q) + E[(1 - q)b_i],
 \end{aligned} \tag{7}$$

where the sums are taken over all possible values of q , b_i , and G , and we have used the fact that $P(s_i | G, q, b_i) = \pi_i$. Since $b_1 + b_2 = 1$, it follows from (7) that

$$p_{11} + p_{22} - 1 = (2a - 1)E(q). \tag{8}$$

Next, consider the quantity $p_{ii}M_{ii}$. Let L denote response latency, and $F_{ij}(t) = P(L \leq t | s_i, r_j)$. Then

$$\begin{aligned}
 p_{ii}F_{ii}(t) &= (\pi_i)^{-1} P(L \leq t, s_i, r_i) \\
 &= (\pi_i)^{-1} \sum_q \sum_{b_i} \sum_G [P(L \leq t | r_i, s_i, G, q, b_i) \\
 &\quad \times P(r_i | s_i, G, q, b_i) \pi_i P(G | q, b_i) P(q, b_i)] \\
 &= \sum_q \sum_{b_i} [aq\Phi_s(t) + b_i(1 - q)\Phi_g(t)] P(q, b_i) \\
 &= a\Phi_s(t)E(q) + \Phi_g(t)E[(1 - q)b_i].
 \end{aligned} \tag{9}$$

Hence

$$p_{11}F_{11}(t) + p_{22}F_{22}(t) = [2a\Phi_s(t) - \Phi_g(t)]E(q) + \Phi_g(t). \tag{10}$$

Differentiating (10) with respect to t , multiplying by t , and integrating from 0 to ∞ , we obtain

$$p_{11}M_{11} + p_{22}M_{22} = (2a\mu_s - \mu_g)E(q) + \mu_g, \tag{11}$$

and using (8) to substitute for $E(q)$, we have

$$p_{11}M_{11} + p_{22}M_{22} = \left(\frac{2a\mu_s - \mu_g}{2a - 1} \right) (p_{11} + p_{22} - 1) + \mu_g, \tag{12}$$

which is Eq. 1 with c_1 written out. Equation 2 is derived in the same fashion, and Eq. 3 follows immediately on subtracting (2) from (1), since $c_1 - c_2 = 2\mu_s$.

It should be noted that in the process of showing that (1), (2), and (3) hold for the means of the various distributions, we have incidentally shown that these equations also hold for the distributions themselves, i.e., (1), (2), and (3) hold if $F_{ij}(t)$ is substituted for M_{ij} , $\Phi_s(t)$ for μ_s , and $\Phi_g(t)$ for μ_g . [All that is required is to substitute for $E(q)$ at the level of Eq. 10 rather than Eq. 11.] Although we make no use of this fact in the present paper, it clearly provides a basis for estimating the underlying SCR latency distribution Φ_s .

Without carrying out all the details, it is also worth mentioning that Eq. 3 (and hence also Eq. 6) holds even if μ_g is regarded as a random variable. That is, we might imagine a family of guessing latency distributions $\{\Phi_{g,k}\}_{k=1}^{\infty}$, from which the subject could "select" a distribution on any given trial depending on events on earlier trials—for example, the subject might select a faster distribution if earlier guesses had been too slow to meet some deadline. In this case $\Phi_g(t)$ is a random variable (for any fixed value of t), and a simple extension of the derivation of Eq. 10 leads to

$$p_{11}F_{11}(t) + p_{22}F_{22}(t) = 2a\Phi_s(t)E(q) + E[(1 - q)\Phi_g(t)],$$

and also

$$p_{12}F_{12}(t) + p_{21}F_{21}(t) = 2(1 - a)\Phi_s(t)E(q) + E[(1 - q)\Phi_g(t)].$$

These equations do not imply (1) and (2) unless we make the additional assumption that q and $\Phi_g(t)$ are uncorrelated. However, if we subtract the second from the first, then differentiate, multiply by t , and integrate, the result is Eq. 3, without any additional assumptions.

Application of the Model to a Sequence of Trials. Suppose a , μ_s , and μ_g are constant over a sequence of trials, but $E(q)$ and $E((1 - q)b_i)$ may vary from trial to trial. In this case p_{ij} and M_{ij} may vary over trials, and it is necessary to write $p_{ij,n}$ and $M_{ij,n}$ to refer specifically to $P(r_j | s_i)$ and $E(L | s_i, r_j)$ on trial n . But since (1), (2), and (3) hold for any distribution of q , b_1 , and b_2 , they hold for every trial, i.e., with $p_{ij,n}$ and $M_{ij,n}$ substituted for p_{ij} and M_{ij} . To translate (1), (2), and (3) into equations involving observable proportions and averages, let $S_{i,n}$ ($i = 1, 2; n = 1, 2, 3, \dots$) denote indicator random variables which index the occurrence of s_1 and s_2 : $S_{i,n} = 1$ if s_i occurs on trial n , zero otherwise; let $R_{i,n}$ ($i = 1, 2; n = 1, 2, 3, \dots$) denote indicator random variables indexing the occurrence of r_1 and r_2 : $R_{i,n} = 1$ if r_i occurs on trial n , zero otherwise; and let L_n denote response latency on trial n . Then for a sequence of N trials in which $P(s_i) = \pi_i$, we define

$$\hat{p}_{ij} = \frac{\sum_{n=1}^N S_{i,n} R_{j,n}}{\pi_i N} \quad (i = 1, 2; j = 1, 2), \quad (13)$$

$$\hat{M}_{ij} = \frac{\sum_{n=1}^N S_{i,n} R_{j,n} L_n}{\sum_{n=1}^N S_{i,n} R_{j,n}} \quad (i = 1, 2; j = 1, 2). \quad (14)$$

Since $E(S_{i,n}R_{j,n}) = \pi_i p_{ij,n}$ and $E(S_{i,n}R_{j,n}L_n) = \pi_i p_{ij,n} \bar{M}_{ij,n}$, we need only attach the trial subscript n to p_{ij} and $\bar{M}_{ij}$ in (1), (2), and (3), and sum both sides of each equation from $n = 1$ to N to establish that

$$E[\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}] = c_1 E[\hat{p}_{11} + \hat{p}_{22} - 1] + \mu_g, \quad (15)$$

$$E[\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21}] = c_2 E[\hat{p}_{11} + \hat{p}_{22} - 1] + \mu_g, \quad (16)$$

$$\frac{1}{2}E[(\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}) - (\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21})] = \mu_s E(\hat{p}_{11} + \hat{p}_{22} - 1), \quad (17)$$

provided a , μ_s , and μ_g are constant throughout the sequence.

In the case of a sequence with $\pi_1 = \pi_2 = .5$, we can define $\hat{p}_c$, $\hat{p}_e$, $\bar{M}_c$, and $\bar{M}_e$ in the natural way, i.e., $\hat{p}_c$ as the proportion of trials with correct responses, $\hat{p}_e = 1 - \hat{p}_c$, $\bar{M}_c$ as the average latency of the correct responses, and $\bar{M}_e$ as the average latency of incorrect responses. With these definitions $\hat{p}_{11} + \hat{p}_{22} - 1 = \hat{p}_c - \hat{p}_e$, $2\hat{p}_c\bar{M}_c = \hat{p}_{11}\bar{M}_{11} + \hat{p}_{22}\bar{M}_{22}$, and $2\hat{p}_e\bar{M}_e = \hat{p}_{12}\bar{M}_{12} + \hat{p}_{21}\bar{M}_{21}$, and consequently (15), (16), and (17) take the form

$$E(2\hat{p}_c\bar{M}_c) = c_1 E(\hat{p}_c - \hat{p}_e) + \mu_g, \quad (18)$$

$$E(2\hat{p}_e\bar{M}_e) = c_2 E(\hat{p}_c - \hat{p}_e) + \mu_g, \quad (19)$$

$$E(\hat{p}_c\bar{M}_c - \hat{p}_e\bar{M}_e) = \mu_s E(\hat{p}_c - \hat{p}_e). \quad (20)$$

Equations 15 through 20 hold regardless of trial by trial sequential dependencies at the level of the SCR probability q and the response biases b_1 , b_2 . Consequently estimates of a , μ_s , and μ_g based on these equations are compatible with any learning model that might describe trial by trial changes in q and b_i . Equations (17) and (20) hold, in addition, even if the latencies of the guesses are subject to sequential effects, and consequently can be used to estimate μ_s even when there is reason to believe that the guessing latency is varying systematically over trials.

Experimental Applications. On the basis of Eqs. 15 and 16 (or 18 and 19), one way to test whether μ_s remains invariant under changes in speed-accuracy bias is to vary incentive conditions over sessions (to generate several different values of $\hat{p}_{11} + \hat{p}_{22} - 1$), plot $\hat{p}_{11}\bar{M}_{11} + \hat{p}_{22}\bar{M}_{22}$ and $\hat{p}_{12}\bar{M}_{12} + \hat{p}_{21}\bar{M}_{21}$ vs $\hat{p}_{11} + \hat{p}_{22} - 1$, and see whether the predicted linear relationships obtain (Ollman, 1966). If the parameters a , μ_s , and μ_g are constant over sessions, we should find the data points lying on two straight lines, with slopes c_1 and c_2 and common intercept μ_g . But this prediction depends on invariance of all three parameters; if the subject were to vary the speed of his guesses (i.e., vary μ_g) from one session to another, linearity could fail even though a and μ_s remained constant. Equation 17 (or 20 in the case $\pi_1 = \pi_2 = .5$), however, provides a test of μ_s invariance that does not require either the mean latency or the proportion of guesses to remain

constant. On the basis of this equation, we can plot $\frac{1}{2}[(\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}) - (\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21})]$ against $(\hat{p}_{11} + \hat{p}_{22} - 1)$, and look for the linear relationship

$$\frac{1}{2}[(\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}) - (\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21})] = \mu_s(\hat{p}_{11} + \hat{p}_{22} - 1), \quad (21)$$

which should be found if μ_s is invariant over sessions. Equation 17 also suggests the following estimate of μ_s based on the data from a single sequence of trials:

$$\hat{\mu}_s = \frac{(\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}) - (\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21})}{2(\hat{p}_{11} + \hat{p}_{22} - 1)}. \quad (22)$$

In the case $\pi_1 = \pi_2 = .5$, we plot $\hat{p}_c\hat{M}_c - \hat{p}_e\hat{M}_e$ vs $\hat{p}_c - \hat{p}_e$, and look for the linear relationship

$$\hat{p}_c\hat{M}_c - \hat{p}_e\hat{M}_e = \mu_s(\hat{p}_c - \hat{p}_e), \quad (23)$$

and in this case the single-sequence estimate of μ_s is

$$\hat{\mu}_s = \frac{\hat{p}_c\hat{M}_c - \hat{p}_e\hat{M}_e}{\hat{p}_c - \hat{p}_e}. \quad (24)$$

*Consistency of $\hat{\mu}_s$.*³ The distribution of the single-sequence estimate $\hat{\mu}_s$ (Eq. 22) depends on the unknown latency distributions Φ_s and Φ_g , and the unknown stochastic process $\{(q_n, b_{1,n}) \mid n = 1, 2, \dots, N\}$ corresponding to the trial by trial sequence of SCR and response bias probabilities. Since we have no basis (or motivation) for making specific assumptions about these unknowns, it is difficult to say anything specific about the small sample distribution of $\hat{\mu}_s$. In practice, then, the case for using $\hat{\mu}_s$ to estimate μ_s must rest on an appeal to the law of large numbers, and consequently it is important to determine the conditions under which such an appeal is justified, i.e., the conditions under which $\hat{\mu}_s$ converges in probability to μ_s as the sequence length $N \rightarrow \infty$. It turns out that sufficient conditions can be specified quite precisely, and that these conditions are weak enough to seem quite reasonable.

The general argument goes as follows. Let T_N and B_N denote, respectively, the numerator and denominator of the right side of (22) for a sequence of N trials, so that $\hat{\mu}_s = T_N/B_N$. If $\text{Var } T_N$ and $\text{Var } B_N$ converge to zero as $N \rightarrow \infty$, then T_N and B_N converge in probability to constants $\tau [= \lim_{N \rightarrow \infty} E(T_N)]$ and $\beta [= \lim_{N \rightarrow \infty} E(B_N)]$, and because of Eq. 17 we know that $\tau = \mu_s\beta$. Suppose $\beta > 0$. Then as $N \rightarrow \infty$, the ratio $T_N/B_N = \hat{\mu}_s$ converges in probability to $\tau/\beta = \mu_s$, i.e., $\hat{\mu}_s$ is a consistent estimate of μ_s . (Fisz, 1963, Theorem 6.14.1, p. 236). The consistency of $\hat{\mu}_s$ depends then on the limits of $E(B_N)$, $\text{Var } B_N$, and $\text{Var } T_N$. These quantities can be calculated

³ This section treats a special topic not referred to elsewhere in the paper. The reader can skip directly to Sec. 3 with no loss of continuity.

explicitly in terms of the SFG model using straightforward methods, and the results are as follows.

First, for B_N , we have

$$B_N = \hat{p}_{11} + \hat{p}_{22} - 1 = \frac{1}{\pi_1 \pi_2 N} \sum_{n=1}^N R_{1,n}(S_{1,n} - \pi_1), \quad (25)$$

where for convenience the 2 in the denominator of (22) is regarded as belonging to the numerator. It is easily shown that⁴

$$E(B_N) = (2a - 1) \frac{1}{N} \sum_{n=1}^N E(q_n), \quad (26)$$

and after a good deal more calculation it can be shown that

$$\begin{aligned} \text{Var } B_N &= [a + \pi(1 - 2a)] \frac{1}{\pi_1 \pi_2 N^2} \sum_{n=1}^N E(q_n) \\ &\quad + \frac{1}{\pi_1 \pi_2 N^2} \sum_{n=1}^N E[b_{1,n}(1 - q_n)] \\ &\quad - (2a - 1)^2 \frac{1}{N^2} \sum_{n=1}^N E^2(q_n) \\ &\quad + \frac{2}{(\pi_1 \pi_2 N)^2} \sum_{n=1}^{N-1} \sum_{m=n+1}^N \text{Cov}[R_{1,n}(S_{1,n} - \pi_1), R_{1,m}(S_{1,m} - \pi_1)], \quad (27) \end{aligned}$$

where (for $m > n$)

$$\begin{aligned} &\text{Cov}[R_{1,n}(S_{1,n} - \pi_1), R_{1,m}(S_{1,m} - \pi_1)] \\ &= \pi_1 \pi_2 (2a - 1) E\{R_{1,n}(S_{1,n} - \pi_1)[E(q_m | S_{1,n}, R_{1,n}) - E(q_m)]\}. \quad (28) \end{aligned}$$

Now consider Eq. 26. Since $a > \frac{1}{2}$ by assumption, $\lim_{N \rightarrow \infty} E(B_N)$ is positive iff $\lim_{N \rightarrow \infty} N^{-1} \sum_{n=1}^N E(q_n)$ is positive. This will be true, if, for example, $\lim_{n \rightarrow \infty} E(q_n)$ exists and is positive, e.g., if the q_n process is asymptotically stationary and $\lim_{n \rightarrow \infty} E(q_n) > 0$. Next consider Eq. 27. The first three terms on the right side clearly converge to zero as $N \rightarrow \infty$, so the limit of $\text{Var } B_N$ depends on the covariance term. Using (28), this term can be written as

$$\frac{2(2a - 1)}{\pi_1 \pi_2 N^2} \sum_{n=1}^{N-1} E \left\{ R_{1,n}(S_{1,n} - \pi_1) \left[\sum_{m=n+1}^N [E(q_m | S_{1,n}, R_{1,n}) - E(q_m)] \right] \right\}. \quad (29)$$

⁴ q_n here denotes the value of q on trial n .

The critical quantity here is $\sum_{m=n+1}^N [E(q_m | S_{1,n}, R_{1,n}) - E(q_m)]$, which can be thought of as representing the expected cumulative effect of the trial n stimulus and response on the subsequent SCR probabilities $q_{n+1}, q_{n+2}, \dots, q_N$. Let $K(N, n, S_{1,n}, R_{1,n})$ denote this quantity. If the effects of trial n wear off sufficiently fast that $K(N, n, S_{1,n}, R_{1,n})$ remains uniformly bounded (for all n) as $N \rightarrow \infty$, then (29), and hence $\text{Var } B_N$, converge to zero as $N \rightarrow \infty$. These conditions are satisfied if, for example, (1) the effects of any trial n last at most some fixed number of trials M , i.e., if $E(q_m | S_{1,n}, R_{1,n}) = E(q_m)$ for $m - n > M$; or more generally, if (2) the effects of any trial n wear off in a geometrically bounded fashion, i.e., for some constant C $|E(q_m | S_{1,n}, R_{1,n}) - E(q_m)| \leq C\theta^{m-n}$, where $0 \leq \theta < 1$. The second condition is commonly satisfied by stochastic learning models which have been applied to psychophysical experiments, and it seems reasonable to expect that it would be satisfied by similar models for the q_n process in RT experiments.

For T_N we have

$$\begin{aligned} T_N &= \frac{1}{2}[\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22} - \hat{p}_{12}\hat{M}_{12} - \hat{p}_{21}\hat{M}_{21}] \\ &= \frac{\sum_{n=1}^N L_n(2R_{1,n} - 1)(S_{1,n} - \pi_1)}{2\pi_1\pi_2 N}, \end{aligned} \quad (30)$$

$$E(T_N) = \mu_s(2a - 1) \frac{1}{N} \sum_{n=1}^N E(q_n), \quad (31)$$

$$\begin{aligned} \text{Var } T_N &= \frac{\mu_{2,s}}{2\pi_1\pi_2 N^2} \sum_{n=1}^N E(q_n) + \frac{\mu_{2,g}}{4\pi_1\pi_2 N^2} \sum_{n=1}^N (1 - E(q_n)) \\ &\quad - \frac{\mu_s^2(2a - 1)^2}{2N^2} \sum_{n=1}^N E^2(q_n) + \frac{1}{2(\pi_1\pi_2 N)^2} \\ &\quad \times \sum_{n=1}^{N-1} \sum_{m=n+1}^N \text{Cov}[L_n(2R_{1,n} - 1)(S_{1,n} - \pi_1), L_m(2R_{1,m} - 1)(S_{1,m} - \pi_1)] \end{aligned} \quad (32)$$

where $\mu_{2,s}$ and $\mu_{2,g}$ are the second raw moments of Φ_s and Φ_g . Assuming these quantities to be finite, the first three terms of (32) converge to zero as $N \rightarrow \infty$, and here again the limit of the variance depends on the covariance term. This term can be expressed in the form

$$\frac{(2a - 1)\mu_s}{\pi_1\pi_2 N} \sum_{n=1}^{N-1} E \left\{ L_n(2R_{1,n} - 1)(S_{1,n} - \pi_1) \left[\sum_{m=n+1}^N [E(q_m | S_{1,n}, R_{1,n}, L_n) - E(q_m)] \right] \right\}. \quad (33)$$

In (33), just as in (29), convergence depends on how quickly the influence of the past

(in this case the stimulus, response, and latency on trial n) wears away, and the remarks made in connection with (29) apply here also.

Roughly speaking then, sufficient conditions for $\hat{\mu}_s$ to be a consistent estimate of μ_s are (i) the distributions Φ_s and Φ_g must have finite second moments (we might as well say finite means and variances), (ii) expected long run performance must be better than chance, i.e., $\lim_{N \rightarrow \infty} E(\hat{p}_{11} + \hat{p}_{22} - 1)$ must be positive, and (iii) the effects of the trial n stimulus, response, and latency on subsequent q_m values must not be too prolonged, i.e., the convergence of $E(q_m | S_{1,n}, R_{1,n}, L_n)$ to $E(q_m)$ must not be too slow—geometric convergence is satisfactory. None of these conditions seems especially unrealistic.

3. EXPERIMENT 1⁵

Method. Three highly practiced subjects were run in a choice reaction time experiment with two stimuli and two responses. Each subject participated in 20 sessions consisting of 480 trials each. The choice stimulus on each trial was a red or green illumination of the screen (4×3 cm) of an IEE series 10 readout located at a viewing distance of 40 cm. Stimulus rise time was approximately 25 msec. In every session the red and green stimuli occurred randomly over trials with equal probability. Responses were key depressions (80 g) using the left and right forefingers, the left hand key being associated with one color, the right hand key with the other. For a given subject the stimulus-response assignment remained the same throughout the experiment. In all but two sessions (see below) a fixed "deadline" prevailed; responses faster than this deadline were rewarded by computing the percentage of all such responses within each session and awarding 0.5 ¢ for each percentage point. In addition to the amount earned in this fashion, subjects were paid 50 ¢/session.

Each trial began with an 800-msec display of an "A" on the readout screen. During the first 50 msec of this display an 800-Hz tone burst also occurred if and only if the response on the previous trial had been slower than the prevailing deadline. The A was followed by a 400-msec display of the letter "B," and this in turn was followed by the red or green choice stimulus. Total time for a 480-trial session was approximately 12 min.

Ten different motivational conditions were studied, with two consecutive sessions being devoted to each. In eight of these conditions, motivation was manipulated by varying the deadline; in the other two by changing the instructions. In the eight "Deadline" conditions the following deadlines were used: 150, 200, 250, 300, 350, 400, 500, and 800 msec. In all of these conditions, responses faster than the deadline were rewarded as described above, and no explicit penalties were imposed for errors. In the instructions, however, subjects were warned that if their error rate consistently exceeded 5%, they might be dropped from the experiment. (In fact, this threat was never carried out even though the 5% limit was frequently exceeded.) The two remaining conditions, a "Speed" condition and an "Accuracy" condition, were run last. In the Speed condition the deadline was set at 150 msec, and the subjects were instructed to respond as fast as they could without regard for errors. In the Accuracy condition the subjects were instructed to try to be correct on every trial without regard for speed, and no deadline was imposed.

⁵ This experiment was briefly reported in a previous paper (Yellott, 1967).

The experiment was controlled by a CDC 160 computer, which also calculated the percentage of fast responses and errors immediately after each session so that this information could be given to the subjects.

Results. Since the presentation probabilities were $\pi_1 = \pi_2 = .5$ in every session, Eqs. 18, 19, and 20 are applicable, and the $\hat{p}_c$, $\hat{p}_e$, $\hat{M}_c$, $\hat{M}_e$ notation can be used in describing results. For purposes of analysis, data from the two sessions corresponding to each condition have been combined. Table 1 shows the observed values of $\hat{p}_c$ (proportion correct), $\hat{M}_c$ (mean latency of all correct responses), $\hat{M}_e$ (mean latency of all incorrect responses), and $\hat{\mu}_s$ for each subject in each condition. The $\hat{\mu}_s$ value in each case is the value obtained by substituting the observed values of $\hat{p}_c$, $\hat{p}_e$, $\hat{M}_c$, and $\hat{M}_e$ into Eq. 24. Averages of the individual subject $\hat{\mu}_s$ values are also shown for each condition.

TABLE 1
Results of Experiment 1

Condition	Subject 1				Subject 2				Subject 3				Average
	$\hat{p}_c$	$\hat{M}_c$	$\hat{M}_e$	$\hat{\mu}_s$	$\hat{p}_c$	$\hat{M}_c$	$\hat{M}_e$	$\hat{\mu}_s$	$\hat{p}_c$	$\hat{M}_c$	$\hat{M}_e$	$\hat{\mu}_s$	
150	.85	296	222	312	.76	265	195	297	.85	270	174	291	300
200	.83	292	230	307	.86	272	201	286	.74	246	185	280	291
250	.81	267	230	278	.91	242	210	245	.80	267	219	284	269
300	.85	268	241	274	.91	275	236	280	.87	281	252	287	280
350	.91	306	296	308	.97	286	262	286	.96	299	276	300	298
400	.91	284	247	288	.93	266	233	268	.95	302	276	304	287
500	.96	303	266	304	.97	287	257	288	.96	304	285	305	296
800	.94	303	287	305	.99	283	251	284	.95	286	276	286	292
Accuracy	1.00	529	—	529	.99	303	252	304	.99	354	365	354	396
Speed	.51	161	156	283	.63	209	162	276	.55	170	154	242	267

An idea of the overall constancy of μ_s across conditions can be obtained from Fig. 2, which shows a plot of $(\hat{p}_c \hat{M}_c - \hat{p}_e \hat{M}_e)$ vs $(\hat{p}_c - \hat{p}_e)$ for each subject, and also for all subjects (i.e., all 30 data points) combined. According to Eq. 20, if μ_s is constant over conditions, all of these points should tend to fall on a straight line through the origin with a slope equal to the common value of μ_s . Straight lines were fit to the eight points corresponding to Deadline conditions in each individual subject graph, and to the 24 Deadline points in the All Subjects graph. (These lines were fit by least squares, subject to the constraint of passing through the origin.) The μ_s values indicated in Fig. 2 are the slopes of the fitted lines. All the data points fall quite close to the fitted

lines, except for the points generated by Subjects 1 and 3 in the Accuracy condition. Since the $\hat{p}_c$ values achieved by these subjects in the Accuracy condition were only slightly higher (no more than four percentage points) than those achieved under the longer deadlines, these points could be interpreted to mean that subjects responded to the Accuracy instructions by "wasting time," i.e., using much more time than was

FIG. 2. Plots of $\hat{p}_c M_c - \hat{p}_c \hat{M}_c$ vs $\hat{p}_c - \hat{p}_c$ for Experiment 1. Numbered points correspond to Deadline conditions; "Speed" and "Accuracy" points correspond to the two Instruction conditions. μ_s in each graph is the slope of the fitted straight line.

actually required for maximum accuracy. This would not have been unreasonable, since the instructions in this condition stressed accuracy without regard for speed, and no feedback on response speed was provided.

Leaving aside the Accuracy condition, Fig. 2 shows little evidence of any systematic variation in μ_s as a function of changes in performance accuracy. In particular, there is

no sign of any substantial increase in μ_s as p_c increases; the relationship between $\hat{p}_c \hat{M}_c - \hat{p}_e \hat{M}_e$ and $\hat{p}_c - \hat{p}_e$ appears to remain essentially linear throughout the range of $\hat{p}_c - \hat{p}_e$, exactly as predicted by Eq. 23. On this basis, it appears that μ_s remained approximately constant across all conditions (except the Accuracy condition), despite sizable changes in performance accuracy.

The μ_s estimates in Table 1 tend to support the conclusion that mean SCR latency does not undergo any dramatic systematic change as accuracy changes from one condition to another. By and large, differences between $\hat{\mu}_s$ values associated with the same levels of $\hat{p}_c$ (e.g., Subject 1, Deadline conditions 350 and 400) seem to be about as large as the differences between $\hat{\mu}_s$ values associated with different levels of $\hat{p}_c$, and exceptionally large and small values of $\hat{\mu}_s$ can be found at both high and low values of $\hat{p}_c$ (e.g., Subject 2, Deadline condition 250; Subject 1, condition 150). There is, however, some statistical evidence in Table 1 which suggests that despite the approximate invariance shown by Fig. 2, there may nevertheless be some tendency for μ_s to increase with p_c . This evidence comes from the rank order correlations between $\hat{\mu}_s$ and $\hat{p}_c$: Excluding the Accuracy condition, these correlations were .40, .08, and .88 for Subjects 1, 2, and 3, respectively. Though only the last of these is significant at the .05 level, the fact that all three are positive suggests the possibility of a small but systematic increase in mean SCR latency associated with increasing accuracy. Presumably this increase, if it exists, is too slight to be detectable in the graphs of Fig. 2.

Figure 3 shows a plot of $2\hat{p}_c \hat{M}_c$ and $2\hat{p}_e \hat{M}_e$ vs $\hat{p}_c - \hat{p}_e$ for each subject and for all subjects combined. (Only the eight Deadline conditions are included in Fig. 3.) According to (18) and (19), both relationships should be linear if the parameters a , μ_s , and μ_g are all constant across conditions. In each graph the lower (solid) line was fit to the $2\hat{p}_e \hat{M}_e$ data points by least squares. The slope and intercept of this fitted line were then used (in combination with the μ_s estimates of Fig. 2) to estimate a and μ_g ; and these parameter estimates were used to generate a predicted line for the $2\hat{p}_c \hat{M}_c$ points. These predicted lines are the upper (broken) lines in each graph. The estimate of μ_g shown in each graph is the intercept of the fitted $2\hat{p}_e \hat{M}_e$ line. The a estimates were obtained by solving for a in Eq. 2' using the μ_s values from Fig. 2, the μ_g estimates just described, and setting c_2 equal to the slope of the fitted $2\hat{p}_e \hat{M}_e$ line. To generate the predicted line for $2\hat{p}_c \hat{M}_c$, these parameter values were substituted into Eq. 1' to obtain a value of c_1 ; the predicted line is $\hat{c}_1(\hat{p}_c - \hat{p}_e) + \hat{\mu}_g$. The data in the "All Subjects" graph were treated in the same fashion, except that here a single line was fit to all 24 $2\hat{p}_e \hat{M}_e$ points, and parameter estimates based on this line were used to generate a single predicted line for all 24 $2\hat{p}_c \hat{M}_c$ points.

By and large, the data points in Fig. 3 fall quite near the straight lines predicted by Eqs. 18 and 19, indicating that μ_g , as well as a and μ_s , remained quite constant across conditions. Apparently, in this experiment, our anxiety about the possibility of changes in the guessing latency was unnecessary. The only serious deviations from

the predicted lines are the two circled points in the graph for Subject 2, and Table 1 suggests that these two discrepancies can be attributed to exceptionally low values of $\hat{\mu}_s$, rather than to changes in μ_g . The parameter estimates shown in Fig. 3 also seem quite sensible: $\hat{\mu}_g$ is about the right value for simple reaction time, as would be expected if the guess responses were simply nondiscriminative reactions to the onset of the choice stimulus. The fact that $\hat{a}$ is close to 1.0 in every case reflects the fact that the

FIG. 3. Plots of $2\hat{p}_c \hat{M}_c$ and $2\hat{p}_e \hat{M}_e$ vs $\hat{p}_c - \hat{p}_e$ for the Deadline conditions of Experiment 1. μ_g , μ_s , and a are the parameter estimates used to generate the pair of straight lines in each graph.

choice stimuli here were perfectly discriminable. Finally, in Fig. 3, as in Fig. 2, the "All Subjects" graph shows that individual differences in this experiment—at the level of the underlying latency distributions—were very small indeed.

Overall then, the results of Experiment 1 support the notion that nonguess responses

(i.e., SCR's) have approximately the same mean latency at all levels of performance accuracy. Because of the uniformly positive rank order correlations between $\hat{p}_c$ and $\hat{\mu}_s$ we cannot rule out the possibility of some slight systematic increase in μ_s as accuracy increases, but Fig. 2 shows clearly that the only substantial change in μ_s occurred in the Accuracy condition, where the subjects were under no obligation to respond as quickly as possible. Across the other nine conditions, μ_s apparently changed very little. However, the experiment failed to produce any substantial amount of data bearing on performance at low levels of accuracy; except for the Speed condition, the lowest observed value of p_c was .74. This is understandable in view of the rather severe instructions used in the Deadline conditions. Experiment 2 was designed to produce a wider range of variation in performance by manipulating the payoff contingencies.

4. EXPERIMENT 2

Method. The reaction time task here was identical to that of Experiment 1 except that each session now consisted of 500 trials, and a different pair of response keys was used. Four new subjects were hired and run for 31 sessions each. Subjects were paid 75 ¢ for each session, plus an amount based on performance. Within each session a fixed deadline was in effect, and the subject was paid 1 ¢ for each percent of responses faster than this deadline. From this amount, however, the experimenter subtracted an amount equal to the percentage of incorrect responses multiplied by a factor that varied from session to session as shown in Table 2. Before each session

TABLE 2
Design of Experiment 2

Sessions	Schedule	Deadline (msec)	Loss (¢) per % errors
1-2	Pretrain	300	-2
3-5	1	300	-2
6-8	2	200	-2
9-11	3	200	$-\frac{1}{2}$
12-14	4	150	$-\frac{1}{2}$
15-17	5	300	$-\frac{1}{2}$
18-20	6	150	-2
21-23	7	300	-4
24-26	8	500	-4
27-29	9	800	-4
30-32	10	150	-0
33-35	11 ^a	150	$-\frac{1}{2}$
36-37	12 ^b	300	-2

^a Schedules 11 and 4 were identical.

^b Schedules 12 and 1 were identical.

the subject was told what deadline was in effect, and how much he would lose for each percent of incorrect responses. The first two sessions were devoted to pretraining. Then over the next 35 sessions the deadline-payoff conditions were changed every fourth session to yield 12 "schedules," as shown in Table 2. Note that Schedules 11 and 12 involved exactly the same conditions as Schedules 4 and 1, respectively. These duplications were included to provide information on the reliability of effects produced by a given set of conditions.

Results. Table 3 shows the results obtained from each subject on each schedule, and the average value of $\hat{\mu}_s$ for every schedule in which estimates could be computed for all four subjects. Each individual subject entry is based on all the responses made by that subject on a given schedule.

The experiment succeeded in generating a wide range of accuracy levels; $\hat{p}_c$ varied from the chance level of .50 to .99. It is evident from Table 3 that the estimate μ_s provided by (24) cannot be expected to be very reliable when $\hat{p}_c$ is close to .5 (cf., for example, the $\hat{\mu}_s$ values for subject 2003 under Schedules 4 and 6). When $\hat{p}_c\hat{M}_c - \hat{p}_e\hat{M}_e$ is plotted against $\hat{p}_c - \hat{p}_e$, however, the effects of this unreliability become less noticeable, as shown in Fig. 4. The straight lines in Fig. 4 were fit to the data points by least squares (with the line constrained to pass through the origin, as in Fig. 2), and the indicated μ_s values are the slopes of the fitted lines.

For the most part, the correspondence between data points and fitted straight lines in Fig. 4 seems about as close as in Fig. 2, indicating that despite very large changes in accuracy, μ_s remained approximately constant across almost all the conditions of Experiment 2. The only notable exception to this conclusion occurred under Schedule 9, which placed the heaviest emphasis on accuracy and involved only a nominal deadline (800 msec); all four graphs show some increase in μ_s under this schedule. But it does not appear that this tendency for μ_s to increase with accuracy was consistently maintained across the entire range of $\hat{p}_c$. The rank order correlations between $\hat{p}_c$ and $\hat{\mu}_s$ (across all 12 schedules) were .01, -.41, .33, and -.05 for Subjects 2001, 2002, 2003, and 2004, respectively; and none of these is significant at the .05 level. It should also be noted that the increase in accuracy from Schedule 8 to 9 is uniformly very small, less than .02 in every case. The overall picture, then, seems very much like that found in Experiment 1: μ_s remained approximately constant across all levels of accuracy, except when time pressure was effectively removed and the subject was motivated to be extremely accurate; in that case μ_s increased noticeably, while accuracy improved very little.

Figure 5 shows plots of $2\hat{p}_c\hat{M}_c$ and $2\hat{p}_e\hat{M}_e$ vs $\hat{p}_c - \hat{p}_e$ for each subject. Recall that the predicted relationship between these quantities depends on μ_g as well as μ_s (Eqs. 18 and 19); if μ_g changes (in particular if the subject begins initiating guesses further back in the foreperiod), we cannot expect the sort of linearity found in Experiment 1 (i.e., in Fig. 3) even if μ_s is constant. Figure 5 shows that linearity did in fact fail in Experiment 2 when the deadline was very low: The open points corresponding to schedules with a deadline of 150 msec fall in every case below the fitted and predicted

TABLE 3
Results of Experiment 2

Schedule no.	Subject 2001				Subject 2002				Subject 2003				Subject 2004				Average $\hat{\mu}_s$
	$\hat{p}_c$	$\hat{M}_c$	$\hat{M}_s$	$\hat{\mu}_s$	$\hat{p}_c$	$\hat{M}_c$	$\hat{M}_s$	$\hat{\mu}_s$	$\hat{p}_c$	$\hat{M}_c$	$\hat{M}_s$	$\hat{\mu}_s$	$\hat{p}_c$	$\hat{M}_c$	$\hat{M}_s$	$\hat{\mu}_s$	
1	.90	311	249	318	.86	290	233	300	.87	302	261	309	.80	288	237	305	308
2	.79	297	216	324	.78	287	185	329	.83	288	236	302	.72	261	192	306	315
3	.64	243	191	313	.63	219	143	329	.62	235	192	302	.69	250	48	330	319
4	.51	117	102	416	.59	164	75	356	.52	147	124	425	.58	154	102	296	373
5	.86	299	225	314	.71	247	191	285	.64	249	199	309	.66	254	211	301	302
6	.65	242	155	342	.57	150	71	383	.56	157	118	207	.55	151	91	413	336
7	.86	311	238	324	.85	275	207	290	.86	299	244	309	.69	254	203	296	305
8	.97	321	273	323	.94	312	275	314	.98	333	306	333	.93	333	323	334	326
9	.98	340	295	341	.96	324	312	324	.99	360	302	361	.95	361	342	363	347
10	.50	114	96	— ^a	.58	148	81	330	.49	77	73	— ^a	.54	90	49	324	—
11	.59	182	126	310	.61	144	69	275	.50	80	79	— ^a	.52	87	54	488	—
12	.82	299	206	325	.73	260	214	289	.68	246	203	284	.66	225	160	288	297

^a When $\hat{p}_c \leq p_c$, $\hat{\mu}_s$ cannot be estimated by Eq. 24. Presumably in this case all the responses are guesses.

lines, which are based on data from the schedules with higher deadlines. These fitted and predicted lines were generated in the same fashion as those of Fig. 3 using only the solid data points, i.e., the data points from schedules with deadlines over 150 msec. The estimates of μ_s , μ_g , and a used to generate these lines are shown in each graph. The μ_s values are those of Fig. 4, while the estimates of a and μ_g are based on the

FIG. 4. Plots of $\hat{p}_c \hat{M}_c - \hat{p}_e \hat{M}_e$ vs $\hat{p}_c - \hat{p}_e$ for Experiment 2. Points correspond to the schedules identified by number in Table 2. μ_s in each graph is the slope of the fitted straight line.

lines fitted to the solid $2p_e M_e$ points, and consequently reflect only the schedules with deadlines greater than 150 msec. Since $\hat{\mu}_g$ is greater than 150 msec in every case, it is not surprising that the subjects apparently adopted a faster guessing latency when the deadline was lowered to 150 msec. A comparison of Figs. 4 and 5, however, shows that this change in the guessing latency had no apparent effect on μ_s : In Fig. 4, the goodness of fit is quite consistent throughout the entire range of $\hat{p}_c - \hat{p}_e$, except in Schedule 9.

FIG. 5. Plots of $2\hat{p}_c\hat{M}_c$ and $2\hat{p}_e\hat{M}_e$ vs $\hat{p}_c - \hat{p}_e$ for Experiment 2. μ_s , μ_g , and a are the parameter estimates used to generate the pair of straight lines in each graph. The μ_g estimate is based only on schedules with Deadlines > 150 msec. Poor fit to open points indicates a change in guessing latency produced by lowering the deadline to 150 msec.

5. EXPERIMENT 3

In Experiments 1 and 2 the stimulus presentation probabilities were $\pi_1 = \pi_2 = .5$ in every session. Experiment 3 was designed to see whether unequal presentation probabilities would have any effect on the invariance of μ_g . In addition, since Experiments 1 and 2 both used the same visual choice stimuli, auditory stimuli were used here.

Method. This experiment followed the same general format as the other two, the only major differences being that auditory choice stimuli were used, and that the presentation

probabilities were not 50–50 in every condition. The stimuli were a pair of tones, one at 1500 Hz (s_1), the other at 1000 Hz (s_2), presented binaurally over earphones at approximately 70 dB (re 0.0002 dynes/cm²). Stimulus onset was controlled by an electronic switch; the total rise time was approximately 25 msec. There were 500 trials/session. Each trial began with a 400 msec display of the digit "1" on the screen of the IEE readout. This was followed by 400-msec display of the digit "2," and this display in turn was followed by the presentation of either s_1 or s_2 . As soon as the subject responded, visual information feedback as to the speed and accuracy of his response was provided by a pair of messages which appeared simultaneously on the readout screen. The speed message was either "speed: O.K.," or "speed: slow" depending on whether or not the preceding response had been faster than the deadline which prevailed during that session. The accuracy message was either "choice: O.K." or "choice: wrong" as appropriate. These feedback messages were displayed for 800 msec, and the digit 1 then reappeared to start the next trial. The responses were key depressions as in Experiments 1 and 2. Each session required about 20 min.

TABLE 4
Design of Experiment 3

Condition	Deadline (msec)	$P(s_1)$
1	300	.5
2	300	.7
3	250	.5
4	250	.3
5	400	.5
6	400	.7
7	300	.1
8	250	.1
9	400	.1

Four naive subjects were hired at the rate of 75 ¢/session plus an amount depending on performance. In each session, this second amount was equal to 1 ¢ for each percent of responses faster than the prevailing deadline, minus 2 ¢ for each percent of incorrect responses. Nine sessions of pretraining were given; in these sessions the deadline was 325 msec, and $P(s_1) = P(s_2) = .5$. After completing the pretraining sessions, subjects were run for two consecutive sessions on each of the nine different conditions shown in Table 4. The nine conditions were run in numerical order for all subjects. For personal reasons, Subject 302 had to leave the experiment after finishing Condition 7. The other three subjects completed all nine conditions.

Results. Data from both sessions of each condition have been combined to yield the results in Table 5. The table shows $\hat{p}_{ij}$, $\hat{M}_{ij}$, and $\hat{\mu}_s$ for each subject in each condition. For Conditions 1–7, in which all four subjects participated, the average values of $\hat{\mu}_s$ were 268, 264, 254, 256, 258, 266, and 268 msec, respectively, which suggests that μ_s did not change to any substantial extent over these conditions.

TABLE 5
Results of Experiment 3

Cond.	Subject 301							Subject 302 ^a						
	$\hat{P}_{11}$	$\hat{P}_{22}$	$\hat{M}_{11}$	$\hat{M}_{12}$	$\hat{M}_{21}$	$\hat{M}_{22}$	$\hat{\mu}_s$	$\hat{P}_{11}$	$\hat{P}_{22}$	$\hat{M}_{11}$	$\hat{M}_{12}$	$\hat{M}_{21}$	$\hat{M}_{22}$	$\hat{\mu}_s$
1	.86	.84	286	232	232	279	.293	.92	.91	265	243	249	269	.269
2	.92	.52	220	229	181	311	.294	.97	.86	253	265	229	275	.266
3	.77	.71	247	186	202	244	.273	.87	.86	250	219	218	259	.261
4	.47	.92	299	153	210	203	.294	.57	.87	254	179	218	225	.265
5	.70	.85	258	195	216	245	.272	.94	.96	275	259	222	281	.280
6	.93	.79	249	229	210	284	.274	.98	.91	271	283	252	296	.285
7	.18	.98	350	110	292	132	.298	.69	.98	287	186	277	232	.270
8	0	1.00	—	88	—	91	—							
9	.08	1.00	398	166	—	174	.332							

Cond.	Subject 303							Subject 304						
	$\hat{P}_{11}$	$\hat{P}_{22}$	$\hat{M}_{11}$	$\hat{M}_{12}$	$\hat{M}_{21}$	$\hat{M}_{22}$	$\hat{\mu}_s$	$\hat{P}_{11}$	$\hat{P}_{22}$	$\hat{M}_{11}$	$\hat{M}_{12}$	$\hat{M}_{21}$	$\hat{M}_{22}$	$\hat{\mu}_s$
1	.94	.90	259	245	258	260	.260	.83	.89	248	199	210	238	.250
2	.98	.85	247	263	243	268	.258	.89	.59	200	188	158	229	.238
3	.86	.79	239	232	229	230	.237	.70	.63	205	166	157	201	.246
4	.63	.93	245	195	233	218	.240	.39	.91	214	147	168	171	.224
5	.92	.88	254	236	250	241	.248	.75	.82	222	181	184	217	.233
6	.99	.91	258	272	256	271	.265	.93	.68	218	209	185	244	.242
7	.60	.99	259	185	268	210	.242	.51	.98	249	163	205	188	.231
8	.29	.99	247	163	200	174	.225	.22	.99	253	158	201	158	.207
9	.54	.99	254	178	231	196	.232	.34	.99	291	172	225	179	.243

^a Subject 302 completed only the first seven conditions.

This impression of approximate μ_s invariance is supported by Fig. 6, which shows, for each subject, the relationship between $\hat{p}_{11} + \hat{p}_{22} - 1$ and $(\frac{1}{2})(\sum \hat{p}_{ii}\hat{M}_{ii} - \sum \hat{p}_{ij}\hat{M}_{ij})$. If μ_s is constant across conditions, this relationship should be linear with slope μ_s and intercept zero (Eq. 17). The straight lines in Fig. 6 were fit by least squares, subject to the constraint of passing through the origin; μ_s in each graph is the slope of the fitted line. If anything, the data points in Fig. 6 lie even closer to the predicted straight lines than those in Figs. 2 and 4. Apparently, despite the unequal presentation probabilities which varied from condition to condition, μ_s remained invariant to a good approximation across all levels of performance accuracy. But at the same time, just as in Experiment 1, this overall conclusion has to be qualified on the basis of the rank order correlation between $\hat{p}_c$ and $\hat{\mu}_s$. These correlations were $-.79$, $.67$, $.93$, and $.60$ for Subjects 301, 302, 303, and 304, respectively. All of these values are significant at the .05 level, suggesting that μ_s tended to increase with accuracy in the case of Subjects

FIG. 6. Plots of $(\frac{1}{2})(\sum \hat{p}_{ii}\hat{M}_{ii} - \sum \hat{p}_{ij}\hat{M}_{ij})$ vs $(\hat{p}_{11} + \hat{p}_{22} - 1)$ for Experiment 3. Points correspond to the Conditions identified by number in Table 4. μ_s is the slope of the fitted line in each graph.

302, 303, and 304, and to decrease as accuracy increased in the case of Subject 301. Figure 6 makes it clear, however, that the changes involved here are relatively small, and this conclusion is reinforced by Table 5. Over the first seven conditions, for example, the differences between the largest and smallest values of $\hat{\mu}_s$ are 26, 24, 22, and 26 msec, respectively, for Subjects 301, 302, 303, and 304.

Figure 7 shows $\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}$ and $\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21}$ plotted against $\hat{p}_{11} + \hat{p}_{22} - 1$ for each subject. Both relationships should be linear if a , μ_s , and μ_g all remained constant across conditions (Eqs. 15 and 16). Straight lines were fit to the $\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21}$ points by least squares; and the slope and intercept of these lines were used to estimate a and μ_g in the manner described in connection with Fig. 3. These parameter estimates are shown in each graph. Predicted lines for the $\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}$ points were generated by substituting these estimates into Eq. 1' to obtain a value of c_1 ; the predicted lines ($\hat{c}_1(\hat{p}_{11} + \hat{p}_{22} - 1) + \hat{\mu}_g$) are the broken lines in each graph. Except for the points corresponding to Conditions 7 and 8 in the graph for Subject 301, the data points are uniformly quite close to the fitted and predicted straight lines, indicating that μ_g normally remained roughly constant across conditions.

FIG. 7. Plots of $\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}$ (x); and $\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21}$ (o) vs $\hat{p}_{11} + \hat{p}_{22} - 1$ for Experiment 3. μ_s , μ_g , and a are the parameter estimates used to generate the pair of straight lines in each graph.

6. SUMMARY AND DISCUSSION OF EXPERIMENTS 1, 2, AND 3

SCR latency. The purpose of all three experiments was to determine the extent to which the speed-accuracy tradeoff involves changes in the latency of nonguess (i.e., stimulus controlled) responses. In all three, a tradeoff was produced by varying deadlines and payoffs over conditions to generate accuracy levels ranging from near

chance to near-perfect. And in all three experiments we were primarily concerned with the linear relationship represented by Eq. 21 [or (23) when $P(s_1) = P(s_2) = .5$]. According to the SFG model, this relationship should hold for a given set of conditions if the mean latency of stimulus controlled responses is invariant across those conditions, i.e., if the speed-accuracy tradeoff is produced simply by mixing different proportions of fast guesses and SCR's, with the latter having a constant mean latency.

The general trend of the results, as shown in Figs. 2, 4, and 6, indicates that mean SCR latency did remain invariant, to a good approximation, across almost all the conditions of these three experiments. The only notable exceptions were the Accuracy condition of Experiment 1, and to a lesser degree, Schedule 9 of Experiment 2. In both of these cases there were obvious and consistent departures from linearity indicating an increase in mean SCR latency. These two exceptional cases involved similar and rather extreme incentive conditions: A heavy emphasis on accuracy, and either no time pressure at all (in the Accuracy Condition), or nominal time pressure in the form of an 800 msec deadline (in Schedule 9). And in both cases the improvement in accuracy which accompanied the apparent increase in mean SCR latency was quite small: In the Accuracy Condition the largest increase in percent correct was from .96 to 1.0 (Subject 1, Table 1), and in Schedule 9, from .93 to .95 (Subject 2004, Table 3).

Across the remaining 29 conditions of the three experiments Figs. 2, 4, and 6 show little sign of any substantial systematic departure from linearity, despite changes in accuracy ($\hat{p}_c$) covering the range .51 to .99. On this basis, it seems fair to conclude that the speed-accuracy tradeoff over these conditions involved no substantial systematic change in SCR latency. Evidently the tradeoff over these conditions was produced almost entirely by varying the mixture of fast guesses and SCR's, and changes in mean SCR latency played at most a negligible role. In other words, the mean latency "corrected for fast guessing" apparently remained approximately invariant over these conditions, despite a speed-accuracy tradeoff in which accuracy ranged from near chance to near-perfect.

But this is not to say that there was no systematic change at all in SCR latency over these 29 conditions. Four of the 11 subjects showed a significant ($p < .05$) positive rank order correlation between accuracy ($\hat{p}_c - \hat{p}_e$ or $\hat{p}_{11} + \hat{p}_{22} - 1$) and estimated mean SCR latency ($\hat{\mu}_s$), and one subject showed a significant negative correlation. (None of these significant correlations involved data from the two conditions for which Figs. 2 and 4 show substantial departures from linearity.) Although Figs. 2, 4, and 6 make it clear that none of these correlations reflect substantial changes, they nevertheless indicate systematic second-order tendencies which must be taken into account. At the same time, however, one should not lose sight of the equally important fact that 6 of the 11 subjects showed no significant rank-order correlation between $\hat{p}_c$ and $\hat{\mu}_s$.

Guess response latency. A secondary concern in all three experiments was whether

the mean latency of guess responses (μ_g) would remain constant across conditions. If this parameter is affected by the speed-accuracy tradeoff, the linear relationships represented by Eqs. 15 and 16 will not hold across conditions, and therefore cannot be used to estimate the parameter a , even though μ_s may be invariant. In Experiments 1 and 3, μ_g evidently remained approximately constant across all conditions, as shown by the high degree of linearity in Figs. 3 and 7. In Experiment 2, μ_g apparently remained constant (at around 180 msec) over most of the conditions, but decreased appreciably when the deadline was lowered to 150 msec.

SCR accuracy. Estimates of the parameter a were obtained for all subjects in each experiment. All of these estimates were at least .90, reflecting the very easy discrimination required in all three experiments. The average values of $\hat{a}$ were .973 and .955 in Experiments 1 and 2, both of which involved a red-green discrimination, and .920 in Experiment 3, which used tones at 1000 and 1500 Hz.

Discussion. Our problem at the beginning of the study was to find a measure of choice RT performance which would be invariant under the speed-accuracy tradeoff. Clearly the simplest solution that could be hoped for is the one provided by the Simple Fast Guess model under the assumption that mean SCR latency is constant for a given subject working on a given discrimination. To the extent that this interpretation of the tradeoff is adequate, "correction for fast guessing" (i.e., estimation of μ_s) provides an estimate of exactly the sort of invariant parameter we hoped to find.

The results of Experiments 1, 2, and 3 suggest that this simple solution can be regarded as a good first approximation for experiments involving easy discriminations, at least for tradeoffs in which accuracy ranges from chance up to about 95% correct. Over this accuracy range the data indicate that any changes in mean SCR latency are small enough to be negligible for many practical purposes—comparing latencies from different discrimination tasks, for example. For purposes of this sort one might well use the overall μ_s estimates in Figs. 2, 4, and 6 to represent each subject's "true mean latency" at all levels of accuracy between chance and 95% correct.

For accuracy levels between 95 and 100% correct, the data are not so clearcut, and it is more difficult to be comfortable with the correction-for-guessing solution. Although one subject (Subject 1 in Experiment 1) was able to achieve an accuracy of 99% correct without any appreciable increase in $\hat{\mu}_s$, other subjects working on the same discrimination showed consistent, and sometimes very sizable, increases when the incentive conditions were modified to encourage very high accuracy. One problem in interpreting this tradeoff near perfect accuracy is the fact that accuracy has a ceiling, but latency does not. Nothing in the choice RT design actually prevents a subject from using more time than necessary to achieve a given level of accuracy, and when deadline pressure is removed to encourage very high levels of accuracy, this sort of inefficiency does not entail any penalty. Indeed, from the subject's point of view, waiting longer than necessary may seem the safest course, even though the extra delay adds nothing

to accuracy. The only empirical test of whether a certain latency is actually the minimum for a given level of accuracy is to ask whether that same accuracy was also achieved at a lower latency. By this test, the All Subjects graph in Fig. 2 (which shows a nearly vertical slope in the neighborhood of $\hat{p}_c - \hat{p}_a = .99$) suggests that much of the dramatic μ_s increase found in the Accuracy Condition of Experiment 1 was due simply to wasted time.

Even if the results of the Accuracy Condition are largely discounted, however, the modest but consistent increase in μ_s generated by Schedule 9 of Experiment 2, and the rank-order correlation data from Experiments 1 and 3, suggest that "wasted time" may not be the entire story. An alternative (or supplementary) hypothesis is that increases in SCR latency near perfect accuracy reflect a true speed-accuracy tradeoff, i.e., a tradeoff between the accuracy and latency of nonguess responses which is enforced by characteristics of the underlying discrimination process. Both LaBerge's (1962) Recruitment theory and the Sequential Probability Ratio Test (SPRT) theory proposed by Stone (1960), Edwards (1965) and Laming (1968) can provide models for such a tradeoff which are compatible with the results of the present study. Both theories assume that the choice stimulus on any trial generates a temporal sequence of independent, identically distributed, random variables $X_1, X_2, \dots$ which the subject examines successively until he has accumulated sufficient evidence to decide in favor of s_1 or s_2 . In Recruitment theory the $\{X_i\}$ correspond to "stimulus elements" which may be of type 1 (associated with stimulus s_1) or type 2 (associated with s_2). The difficulty of the discrimination depends on the probabilities $p_1(i)$ and $p_2(i)$, where $p_j(i)$ denotes the conditional probability that an arbitrary element is of type i ($i = 1, 2$), given that the stimulus is s_j . Unless $p_1(1) = p_2(2) = 1.0$, a single element does not uniquely determine the stimulus which generated it. The subject is assumed to set criteria R_1 and R_2 , and to examine the incoming sequence of elements until he obtains R_1 s_1 elements or R_2 s_2 elements, whichever comes first. Assuming for simplicity that $p_1(1) = p_2(2)$, the subject can increase his accuracy from $p_i(i)$ (when $R_1 = R_2 = 1$) up to 1.0 (the limit as R_1 and $R_2 \rightarrow \infty$) by increasing the R_i , but only at the cost of increasing the number of elements which must be collected before reaching a decision.

Now suppose that for the easy discriminations required in the present experiments, $p_1(1)$ and $p_2(2)$ were both around .95. In this case the subject could achieve a 5% error rate by sampling only a single element, and with a two-element criterion (i.e., $R_1 = R_2 = 2$) his error rate would be less than 1%. Under these conditions the only way to produce error rates higher than 5% would be to mix in some proportion of fast guess trials, i.e., trials on which no elements are sampled. If we assume that stimulus elements do not begin to arrive until about 100 msec after the subject becomes aware of stimulus onset, and that the time for each additional stimulus element is on the order of 20–30 msec, then it is possible to give a fairly complete interpretation of all our results: Guesses (normally responses to nondiscriminative information provided by the stimulus onset) will usually average 100 msec faster than SCR's; mean SCR

latency will be approximately constant as long as the subject typically samples only one element on his SCR trials; there may be some tendency for SCR latency to increase a bit with accuracy as the subject mixes in more two-element criterion trials; and finally, when time pressure is relaxed and high levels of accuracy are required, the subject will shift from a one-element criterion to some higher criterion—perhaps $R_1 = R_2 = 2$, which would produce only a relatively small change in SCR latency, or perhaps a much higher criterion, which would produce no real improvement in accuracy (already at .99+ for $R_1 = R_2 = 2$), but which would lead to a dramatic increase in latency (“wasted time”).

A similar interpretation can be based on the Stone-Edwards-Laming model. In that model the $\{X_i\}$ are arbitrary random variables with (for example) probability density functions $p_1(x)$ or $p_2(x)$ depending on whether s_1 or s_2 is presented. After the arrival of each successive random variable, the subject computes the current likelihood ratio $L(X_1, X_2, \dots, X_n) = (p_1(X_1)p_1(X_2) \cdots p_1(X_n))/(p_2(X_1)p_2(X_2) \cdots p_2(X_n))$. As soon as likelihood ratio exceeds an upper criterion C_1 , or falls below a lower criterion C_2 the subject stops sampling and responds r_1 or r_2 as appropriate. If the distributions $p_1(\)$ and $p_2(\)$ are concentrated on disjoint sets, of course, the subject need only sample X_1 to be perfectly accurate. In other cases the subject can increase his accuracy by increasing C_1 and C_2^{-1} , but only at the cost of increasing the average sample size required to reach criterion. Minimum time and accuracy are achieved when $C_1 = C_2 = 1$, since in this case likelihood ratio always reaches criterion after one observation. If we suppose that in our experiments $P[L(X_1) > 1 | s_1]$ and $P[L(X_1) < 1 | s_2]$ were both around .95, then the subject could only achieve accuracy levels below 95% correct by fast guessing (i.e., responding before sampling any of the $\{X_i\}$) on some proportion of the trials. Our results then could be interpreted by assuming $p_i(\)$ distributions of this sort, and supposing that under most conditions, subjects either guessed or made minimum latency SCR's based on a single observation.

These interpretations assume that the amount of tradeoff between SCR accuracy and latency in our experiments was limited by the fact that even a single information “quantum” (i.e., a single one of the $\{X_i\}$) allowed almost perfect stimulus identification. In this case, the subject's options are essentially limited to waiting for the first one or two quanta, and achieving almost perfect accuracy, or responding before any quanta have arrived and achieving chance accuracy. One obvious implication is that a true speed-accuracy tradeoff (involving SCR's alone) might be found with more difficult discrimination tasks. A recent study by Swensson (1968) provides evidence that this is so. This study was designed to find experimental conditions which would produce a true speed-accuracy tradeoff of the sort envisioned by the Stone-Edwards-Laming model. Tradeoffs were induced by using a linear cost for time function (so many points lost for each additional msec of RT) combined with various payoffs for correct and incorrect responses, rather than by using deadlines as in the present study. In his initial experiments, which involved easy discriminations, Swensson found that

subjects made only two kinds of responses: fast guesses, and highly accurate stimulus controlled responses which had constant latencies and accuracies. (Unlike subjects in the present study, however, Swensson's subjects did not alternate between these two kinds of responses in a random fashion from trial to trial. Instead they apparently used either one strategy or the other exclusively for long blocks of trials, which made it possible for Swensson to separate fast guesses and SCR's by direct observation, rather than using the statistical methods of the Fast Guess model.⁶) In later experiments, Swensson was able to produce a true tradeoff involving intermediate error rates (between zero and chance) by using very difficult shape discrimination problems (in effect, identifying the longer edge of rectangles which had length to width ratios ranging from 35 : 34 to 85 : 84) and allowing the subjects a "free delay" (typically 250 msec) after the stimulus onset before starting the cost-for-time clock. Under these conditions, tradeoffs were found which involved only long latency (i.e., presumably stimulus controlled) responses. For the easiest discrimination (a length to width ratio of 35 : 34) the amount of latency change involved in this tradeoff was apparently quite small (about 30 msec as accuracy varied from a lower limit of 80% correct up to 98% correct), but for the hardest discrimination (a length to width ratio of 85 : 84) mean latency varied over a range of approximately 250 msec as accuracy increased from 65 to 97% correct. (Data from Subject EK, Experiment IV, Fig. 9 in Swensson, 1968.)

Swensson's results indicate that it is possible to obtain a substantial tradeoff involving only SCR's if the discrimination task is made sufficiently difficult and the incentive conditions sufficiently elaborate. Whether his free delay and linear cost for time procedures would yield any sizable SCR tradeoff with easy discriminations of the sort used in the present study is an open question; such an experiment has not been done. But in view of the results obtained with the easiest of Swensson's hard discrimination problems (all of which were much more difficult than those of the present study) it seems unlikely that they would. What seems more likely is that highly discriminable choice stimuli allow the subject very little leeway for trading SCR speed against accuracy, regardless of how we arrange the payoff contingencies. In other words, it seems likely that for discriminations of this sort, SCR accuracy is bounded between near-perfect and perfect, and SCR mean latency is approximately constant—unless the subject wastes time trying to achieve further reductions in an already negligible error rate. If this interpretation is generally correct, then the correction for guessing supplied by the SFG model provides a simple, first-approximation solution to the problem described at the beginning of the paper.

⁶ According to the SFG model a subject should always set q equal to zero or 1 to maximize expected payoffs—under either a deadline or linear cost for time procedure. Swensson's subjects apparently operated in something close to this optimal fashion, while subjects in the present study did not. Probably this difference has to do with the more precise trial by trial information feedback provided in Swensson's experiment, which encouraged his subjects to learn the optimum strategy.

7. WHAT OTHER MODELS ARE COMPATIBLE WITH A LINEAR TRADEOFF FUNCTION?

The principal result of Experiments 1, 2, and 3 was the nearly perfect linear relationship between $(\hat{p}_{11} + \hat{p}_{22} - 1)$ and $(\frac{1}{2})[(\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}) - (\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21})]$ shown in Figs. 2, 4, and 6. The SFG model, of course, implies that (the expectations of) these quantities will be linearly related iff mean SCR latency is invariant (Eq. 17), and up to this point our interpretation of results has been carried out entirely in terms of that model. That is, we have acted as though a strictly linear relationship, of the sort implied by Eq. 17, could have only one interpretation: that the SFG model is correct, and mean SCR latency is invariant. But clearly there is a possibility that the same result might also be predicted by other models in which it could have a very different interpretation. So the question arises, what other models, besides SFG, are compatible with a strictly linear relationship between the expectations of $(\hat{p}_{11} + \hat{p}_{22} - 1)$ and $(\frac{1}{2})[(\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}) - (\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21})]$?

I have not found any general answer to this question, and to prove even the partial results I have been able to obtain would require an unjustifiably longwinded development. Nevertheless, the question deserves some discussion; it does not seem intractable, and someone else might well have better success with it. So this section presents a brief heuristic treatment of the general problem, long on explication and short on results. The final two sections of the paper then describe several specific models which are compatible with the linear relationship in question and also seem to have some independent importance.

To explicate the general problem we first need some shorthand notation. Let

$$\hat{A} = \hat{p}_{11} + \hat{p}_{22} - 1, \quad (34)$$

$$\hat{T} = \frac{1}{2}[(\hat{p}_{11}\hat{M}_{11} + \hat{p}_{22}\hat{M}_{22}) - (\hat{p}_{12}\hat{M}_{12} + \hat{p}_{21}\hat{M}_{21})]. \quad (35)$$

Then the broad question is, what can the relationship between $\hat{T}$ and $\hat{A}$ tell us about the mechanism responsible for the speed-accuracy tradeoff? Specifically, if this relationship appears to satisfy a linear equation of the form $\hat{T} = \mu_s \hat{A}$, what models can and cannot be ruled out? The experimental context here is a two-choice RT experiment in which we run the same subject for many session on the same discrimination problem, varying incentive conditions from session to session to encourage a speed-accuracy tradeoff. Each session generates a pair of values $(\hat{A}, \hat{T})$, and over sessions these pairs trace out a function of the form $\hat{T} = f(\hat{A})$ —the sort of function pictured in Figs. 2, 4, and 6. I will call this the *tradeoff function*, and speak of a “linear tradeoff function” when $\hat{T} = \mu_s \hat{A}$.⁷ (μ_s here is simply a label for the slope of the tradeoff

⁷ The general linear form $\hat{T} = \mu_s \hat{A} + C$ does not need to be considered; it could arise only because of sampling error, or as an artifact of unusual experimental procedures. For if $\hat{A} = 0$, then $\hat{p}_{11} = \hat{p}_{12}$ and $\hat{p}_{22} = \hat{p}_{21}$, which imply that the subject is simply guessing. But in this case we expect also $\hat{M}_{11} = \hat{M}_{21}$ and $\hat{M}_{22} = \hat{M}_{12}$, since a subject who is guessing cannot adjust

function; its substantive interpretation will depend on what model is being considered).

Now the theoretical implications of an empirical tradeoff function of this sort are determined by comparing it with the theoretical relationship between $E(\hat{A})$ and $E(\hat{T})$ implied by various models. And in asking whether a model is compatible with a linear tradeoff function, we are really asking whether the pairs $(E(\hat{A}), E(\hat{T}))$ generated by that model can satisfy a linear relationship. More precisely: Models for the speed-accuracy tradeoff generally specify mechanisms involving two kinds of parameters: *task parameters*, which are assumed to remain constant for a given subject throughout a given experiment, and *bias parameters*, which the subject can adjust as often as he likes to take account of changing incentive conditions. (In the Recruitment model described in Sect. 6, for example, the criteria R_1 and R_2 would be regarded as bias parameters, while $p_1(1)$, $p_2(2)$, and various time constants, would be task parameters.) The speed-accuracy tradeoff then is attributed to changes in the bias parameters, and the theoretical tradeoff function for any given model is simply the collection of all $(E(\hat{A}), E(\hat{T}))$ pairs which that model allows the subject to generate using constant task parameters—different pairs being produced by assigning different values to the bias parameters, or more generally, by assuming different stochastic processes for the trial to trial values of these parameters. Our concern is whether these pairs satisfy a linear relationship, and are thus consistent with an empirical linear relationship between $\hat{A}$ and $\hat{T}$.

Three possibilities arise. First, a model may imply that *all* pairs $(E(\hat{A}), E(\hat{T}))$ which can be generated by a given subject in a given experiment must satisfy a common linear equation,

$$E(\hat{T}) = \mu_s E(\hat{A}), \quad (36)$$

for some constant μ_s . Models of this sort can be said to *require* a linear tradeoff function. It was shown in Sec. 2 that the SFG model requires a linear tradeoff function if its SCR parameters a and μ_s are assumed to be constants [Eq. 36 is just (17) rewritten in the notation of (34) and (35)]. Section 8 below describes two generalizations of that model which have the same property if their SCR parameters are taken to be constant. In these two models, however, interpretation of the slope constant μ_s is more complicated than in the SFG model. These two generalizations of SFG are the only models I know of which require a linear tradeoff function. But there is no reason to believe that others cannot be devised.

A second possibility is that a model might allow the subject to generate either a

his latency according to which stimulus is presented. Consequently if $\hat{A} = 0$, we expect $\hat{T} = 0$ also, except for sampling error. The only exception I can think of would be if the subject were responding indiscriminately to stimulus onsets, and the two stimuli have different onset times. This possibility is one we would probably want to avoid experimentally, and can afford to ignore theoretically.

linear or a nonlinear tradeoff function, dependency on how he adjusts his bias parameters. Such a model would not require a linear tradeoff function, and consequently the statistic $\hat{T}/\hat{A}$ (i.e., μ_s in Eq. 22) could not be regarded as an estimate of an invariant parameter. But such a model would be *compatible* with a linear tradeoff function. Section 9 describes a model of this sort which is particularly interesting because, like the SFG model, it provides a mechanism for the speed-accuracy tradeoff in which mean stimulus controlled response latency remains invariant.

Finally, a model may be *incompatible* with a linear tradeoff function in the sense that it does not allow the subject to generate a set of $(E(\hat{A}), E(\hat{T}))$ pairs lying on a common straight line. Here it is necessary to be precise about the range of $E(\hat{A})$ values required. We will say that a model is incompatible with a linear tradeoff function over the interval (a, b) if that model does not allow the subject to generate a colinear set of $(E(\hat{A}), E(\hat{T}))$ pairs in which $E(\hat{A})$ assumes every value in (a, b) . It can be shown that LaBerge's Recruitment model and the Stone-Edwards-Laming SPRT model are both incompatible with a linear tradeoff function over $(0, 1)$, and that this is true even if the conventional assumptions of these models are augmented to allow the subject to guess as often as he likes (as was done in Sec. 6), and the criterion parameters are allowed to vary from trial to trial according to any arbitrary stochastic process—so long as that process is first-order stationary.

This last result illustrates in a specific way the sort of general theorems one would like to obtain. Rather than demonstrate that particular models are incompatible with a linear tradeoff function, one would like to be able to show this for general families of models which share some common assumption about the mechanism responsible for the speed-accuracy tradeoff. In particular, it would be useful to be able to prove that any model which attributes the tradeoff to changes in stimulus controlled response latency must be incompatible with a linear tradeoff function over $(0, 1)$. I do not know of any counter-example disproving this claim, but neither do I know how to prove it, or how to weaken the claim in such a way that it becomes provable and still remains useful. So for the moment, at least, the general question of what models are compatible with a linear tradeoff function remains open.

But as I indicated at the start of this section, the question does not seem intractable; it is possible to obtain some general results by narrowing down the class of models one is prepared to consider. To suggest the sort of strategy that might be followed in proving strong general theorems, it seems useful to conclude this section by proving a weak general result—the only one I've been able to obtain. The idea here is to show that if we restrict our attention to what I will call *Pure Tradeoff models*, then any model which attributes the speed-accuracy tradeoff to changes in mean SCR latency must be incompatible with a linear tradeoff function over the (closed in this case) interval $[0, 1]$. A Pure Tradeoff model is defined as follows. Suppose the subject has available $N + 1$ "states" $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_N$, where a state σ_i corresponds to a pair of parameters (a_i, λ_i) . a_i denotes the probability of a correct response, and λ_i the mean

latency, which apply when the subject is in state σ_i .⁸ We assume that σ_0 is a guessing state ($a_0 = .5$), and that the remaining states (the SCR states) are ordered in such a way that increasing accuracy is associated with increasing mean latency: $.5 = a_0 < a_1 < a_2 \cdots < a_N = 1.0$, and $\lambda_0 < \lambda_1 < \lambda_2 \cdots < \lambda_N$. And we assume that the subject can change states from one trial to the next, but in such a way that within any session the trial by trial sequence of states is a first-order stationary stochastic process. Thus for any session there will be some set of state probabilities $\{c_0, c_1, \dots, c_N\}$, where c_i is the marginal probability of state σ_i for that session.

It follows from these assumptions that for any session we have

$$E(\hat{A}) = \sum_{i=1}^N c_i(2a_i - 1), \quad (37)$$

$$E(\hat{T}) = \sum_{i=1}^N \lambda_i c_i(2a_i - 1). \quad (38)$$

(Note that the guessing latency λ_0 plays no role here.) Now how could such a model allow the subject to produce a linear tradeoff function, with slope μ_s , in which $E(\hat{A})$ assumes every value in $[0, 1]$? One way, clearly, is to set $\lambda_N = \mu_s$, and then assume that the subject never uses any state except σ_N and σ_0 . This, of course, is the SFG model. Is there any other way—could the subject ever use any of the intermediate states (that is, could he ever use a mean SCR latency other than λ_N) and still produce a linear tradeoff function?

The answer is no. For if $E(\hat{T}) = \mu_s E(\hat{A})$, then it follows from (37) and (38) that

$$\sum_{i=1}^N c_i(2a_i - 1)(\lambda_i - \mu_s) = 0, \quad (39)$$

and consequently, to produce a linear tradeoff function the subject must confine his choice of state probabilities to those which do not violate (39). But since the value $E(\hat{A}) = 1.0$ can only be achieved by setting $c_N = 1.0$, we can see from (39) that λ_N must equal μ_s , and therefore $\lambda_1, \lambda_2, \dots, \lambda_{N-1}$ are all strictly less than μ_s . Consequently, since $2a_i - 1$ is positive for all $i \geq 1$, the left side of (39) will have to be negative if

⁸ That is, when the subject is in state σ_i , $p_{11} = p_{22} = a_i$, and $M_{11} = M_{22} = M_{21} = M_{12} = \lambda_i$. It is these restrictions that prevent the Pure Tradeoff models from supplying a general representation for models of the speed-accuracy tradeoff. For example, the Recruitment model is not a Pure Tradeoff model, because in the natural "states" of that model errors have a longer expected latency than correct responses (i.e., $M_{21} > M_{11}$ and $M_{12} > M_{22}$ for any fixed pair of criterion values $R_1 > 1, R_2 > 1$). On the other hand, the assumption that the number of states is finite, with $a_N = 1.0$, is not a serious restriction, since the result proved above remains valid if there are a countable number of states and $\lim_{n \rightarrow \infty} a_n = 1.0$.

the subject ever assigns nonzero probabilities to any of the intermediate SCR states $\sigma_1, \sigma_2, \dots, \sigma_{N-1}$. In other words, the only Pure Tradeoff model compatible with a linear tradeoff function over $[0, 1]$ is the Simple Fast Guess model.

The weakness of this result, of course, lies primarily in the fact that the class of Pure Tradeoff models is too narrow. The open problem that remains is to extend the same result to more general families of models.

8. TWO GENERALIZATIONS OF THE SIMPLE FAST GUESS MODEL

The General Case of Ollman's Fast Guess Model (FG1). The SFG model assumes that only two kinds of responses are made in any session: SCR's with mean latency μ_s and probability a of being correct, and guesses with mean latency μ_g . No provision is made for latency differences between various kinds of SCR's (i.e., various $s_i - r_j$ combinations), or for differences between the latency of r_1 and r_2 guess responses. The general case of the Fast Guess model proposed by Ollman (1966) takes such possibilities into account. Figure 8 shows the single-trial branching process of this model, which is referred to below as FG1. Model FG1 differs from the SFG model described in

FIG. 8. Branching process of model FG1 for a single trial.

Sec. 2 only in that (i) when the subject makes a SCR, $P(r_j | s_i) = a_{ij}$, and $E(L | s_i, r_j) = \mu_{ij}$; (ii) when the subject guesses, $E(L | s_i, r_j) = \mu_{gj}$. When $a_{11} = a_{22} = a$, $\mu_{11} = \mu_{22} = \mu_{12} = \mu_{21} = \mu_s$, and $\mu_{g1} = \mu_{g2} = \mu_g$, model FG1 reduces to the SFG model.

Using the same sort of derivation as that outlined in Sec. 2, it is easily shown that

model FG1 implies Eq. 3, and thus Eq. 17 [which is the same as (36)] with μ_s equal to

$$\frac{\mu_{11} + \mu_{22}}{2} + \frac{a_{21}(\mu_{11} - \mu_{21}) + a_{12}(\mu_{22} - \mu_{12})}{2(a_{11} + a_{22} - 1)}. \quad (40)$$

Consequently FG1 requires a linear tradeoff function with a slope μ_s given by (40), provided the SCR parameters a_{ij} and μ_{ij} are invariant.

In general, if a_{ij} and μ_{ij} are unrestricted, it is hard to think of any concise label for the quantity given by (40). But if the second term in (40) is neglected, what remains is simply the average of μ_{11} and μ_{22} , i.e., the average mean latency of correct SCR's. The second term in (40) vanishes if $\mu_{1i} = \mu_{2i}$, i.e., if correct and incorrect r_i SCR's have the same mean latency (Stone has shown that this is always true in the SPRT model), or if $a_{21} = a_{12} = 0$. More generally, if a_{21} and a_{12} are quite small, say .05 or less, the second term in (40) will also be relatively small: on the order of 5 msec even if $(\mu_{11} - \mu_{21})$ and $(\mu_{22} - \mu_{12})$ are both as large as 100 msec. Consequently in situations where it is known that the subject can achieve small error rates under proper motivation, the distortion involved in identifying the slope of the tradeoff function as "the average mean latency of correct SCR's" will be negligible—at least as far as model FG1 is concerned.

A Fast Guess Model with Stimulus Biases (FG2). The fact that FG1 predicts a linear tradeoff function is not unwelcome. On the contrary, it would have been disconcerting to find that a linear tradeoff function could only be predicted by theories which fit the rigid specifications of the SFG model. If linear tradeoff functions are a common empirical phenomena, then they must be compatible with other results commonly found in RT experimnts, and consequently they must be predictable by models which account for a fairly broad range of phenomena. Model FG1 is a start in that direction, since it can predict intrinsic response asymmetries not allowed in the SFG model. But FG1 is severely limited in its ability to account for sequential effects: changes in performance to particular stimuli as a function of the local history of the stimulus presentation sequence. It is well established that subjects show complex and fairly dramatic changes in their latency for particular stimuli depending on events over the immediately preceding sequence of trials—sometimes positive recency effects (e.g., Falmagne, 1965), and sometimes negative recency (e.g., Laming, 1969). Apparently these sequential effects cannot simply be attributed to changes in response bias alone; they must also involve changes in stimulus bias, i.e., changes in the subject's readiness to process a particular stimulus—or at any rate, an entire stimulus–response pair (LaBerge, Van Gelder, and Yellott, 1970). Model FG1 provides no mechanism to account for sequential effects of this sort, since the only trial by trial bias changes possible in that model are changes in the guess response biases b_1 and b_2 .

We now describe a generalization of the SFG model, called FG2, which provides a mechanism to account for trial by trial sequential effects, and at the same time

predicts a linear tradeoff function. The essential idea of FG2 is that the subject "decides" at the beginning of each trial whether to be set for stimulus s_1 or s_2 , and whether to make a guess or an SCR. Figure 9 shows the branching process for a single trial.

Suppose that at the start of any trial the subject is set for s_1 (state $\mathcal{S}_1$) with probability b_1 , or for s_2 (state $\mathcal{S}_2$) with probability b_2 ($b_1 + b_2 = 1$), and, in addition, set to make either an SCR (with probability q) or a guess (with probability $1 - q$). If the subject makes an SCR and the stimulus presented agrees with his stimulus set (i.e.,

FIG. 9. Branching process of model FG2 for a single trial.

s_i for state $\mathcal{S}_i$), his response is correct with probability a . The expected latency of SCR's in this case is μ_{sc} . If the subject is set for s_i , and s_i is presented, and he makes an incorrect SCR, his expected latency is μ_{se} . If the subject makes an SCR and the stimulus presented is the one he is not set for, his probability of responding correctly is a' , and his expected latency is μ'_{sc} for correct responses, μ'_{se} for incorrect responses. Finally, if the subject decides to guess, he makes the response corresponding to his stimulus set (i.e., r_i for state $\mathcal{S}_i$), and his expected latency is μ_{g1} or μ_{g2} depending on whether the response is r_1 or r_2 . It is natural to suppose that $\mu_{sc} \leq \mu'_{sc}$, and that $a \geq a'$, i.e., the subject responds more quickly and accurately to the stimulus he is set for. We assume that all the SCR parameters a , a' , μ_{sc} , μ_{se} , μ'_{sc} , and μ'_{se} are constants, but q , b_1 and b_2 can vary over trials. If $a = a'$, $\mu_{sc} = \mu_{se} = \mu'_{sc} = \mu'_{se}$, and $\mu_{g1} = \mu_{g2}$, model FG2 reduces to the SFG model.

Given these assumptions, it is easy to see that by allowing b_1 and b_2 to vary over trials according to an appropriate learning model [as, for example, Falmagne (1965)]

has done with a similar model] one could account (at least qualitatively) for such phenomena as decreases in the mean latency of responses to s_i following homogeneous runs of that stimulus. Moreover, FG2 requires a linear tradeoff function, since it can be shown to imply Eq. 3, and consequently Eq. 36, with μ_s equal to

$$\frac{\mu_{sc} + \mu'_{sc}}{2} + \frac{(1-a)(\mu_{sc} - \mu_{se}) + (1-a')(\mu'_{sc} - \mu'_{se})}{2(a + a' - 1)}. \quad (41)$$

Here, as with FG1, the slope of the tradeoff function does not correspond to any easily interpretable substantive quantity unless we place some restrictions on the parameters. But if the probability of a SCR being incorrect is small, i.e., .05 or less, then μ_s [i.e., (41) here] will be approximately equal to the average mean latency of correct SCR's—the first term of (41) in this case. Consequently we can say that for both FG1 and FG2 the slope of the tradeoff function has the same interpretation in cases where small error rates are known to be possible under proper motivation: In that case μ_s can be regarded as the average mean latency of correct stimulus controlled responses, plus or minus a few milliseconds. According to both models then, this is the quantity which is estimated by correction for fast guessing.

9. A DEADLINE MODEL

The SFG model proposes to account for the entire speed-accuracy tradeoff in terms of a single invariant "true mean latency," together with guessing probabilities and guess response latencies which the subject can adjust according to motivation. That model, and its generalizations FG1 and FG2, require a linear tradeoff function in which the constancy of the slope reflects the invariance of SCR mean latency. On this basis, one might suppose that any model which specifies an invariant mean SCR latency ought to also require a linear tradeoff function, since if SCR latency is fixed the subject can only produce a tradeoff by adjusting the parameters of guess responses. The following example, however, shows that this is not the case.⁹

Assumptions. Suppose that on every trial, information about the identity of the choice stimulus takes the form of a single quantum which arrives S msec after stimulus onset. S is a continuous random variable with cumulative distribution function $S(t)$ ($=P(S \leq t)$) and density function $s(t)$. If the subject waits until the arrival of the information quantum, and then responds, his response is correct with probability one. On each trial, however, the subject presets a deadline D ; if the information quantum has not arrived D msec after stimulus onset, the subject guesses (with some bias probabilities b_1 and b_2 for responses r_1 and r_2). D is a continuous random variable with

⁹ This model was suggested to me by R. C. Atkinson in the mathematically equivalent form of a race between guesses and SCR's: Guesses have latency D ; SCR's have latency S ; whichever occurs first on any trial determines the nature of the response on that trial. The same model has also been suggested by Ollman (personal communication).

distribution function $D(t)$ ($=P(\mathbf{D} \leq t)$) and density function $d(t)$. For a given subject, working on a given discrimination problem, the distribution $S(\cdot)$ is assumed to be fixed. The deadline distribution $D(\cdot)$, however, is adjustable from one session to another.

[According to these assumptions, response latency on any trial is either $\mathbf{S}$ (if $\mathbf{S} \leq \mathbf{D}$), or $\mathbf{D}$ (if $\mathbf{S} > \mathbf{D}$). For any serious application of the model, one probably would want to assume some motor response time random variable $\mathbf{R}$, in which case latency would be $\mathbf{S} + \mathbf{R}$ or $\mathbf{D} + \mathbf{R}$. But, for our purposes, there is no loss of generality in setting $\mathbf{R} = 0$. One would also probably want to assume that the deadline is measured not from stimulus onset—which the subject can hardly sense instantaneously—but from some later point, i.e., the instant at which he becomes aware of stimulus onset. But this notion is already subsumed in the general definition of $\mathbf{D}$, since we do not assume anything about its distribution.]

Specialization to the SFG Model. The Deadline model, like the SFG model, assumes that SCR mean latency is an invariant—in the sense that if the subject makes nothing but SCR's, his mean latency is a constant $E(\mathbf{S})$. But here the subject does not predetermine on every trial whether he will guess or make an SCR. Rather, he sets a deadline which may or may not be long enough to allow him to make an SCR. Of course, if $S(\cdot)$ is concentrated on some interval I_s , the subject can guarantee that his response will be correct by setting $\mathbf{D}$ above I_s , and by the same token he can guarantee that his response will be a guess by setting $\mathbf{D}$ below I_s . So to obtain the SFG model as a special case, imagine that $P(\mathbf{D} \in I_s) = 0$, $P(\mathbf{D} < I_s) = 1 - q$, and $P(\mathbf{D} > I_s) = q$; any $D(\cdot)$ distribution which satisfies these conditions yields a model identical to the SFG model, with $a = 1$. Consequently under these conditions, the Deadline model predicts a linear tradeoff function with slope $\mu_s = E(\mathbf{S})$. But in general it does not, as we now show.

Nonlinear Tradeoff Functions for the General Case. To show that the Deadline model does not require a linear tradeoff function, it is sufficient to show that the ratio $E(\hat{T})/E(\hat{A})$ is not invariant under arbitrary transformations of the deadline distribution $D(\cdot)$. In that case, (36) need not hold if the subject varies $D(\cdot)$ from session to session; hence the tradeoff function need not be linear. We consider the special case $\pi_1 = \pi_2 = .5$. The problem is to calculate $(p_c M_c - p_e M_e)$ [i.e., $E(\hat{T})$ in this case] and $p_c - p_e(E(\hat{A}))$, and show that their ratio depends on $D(\cdot)$. Assuming that $S(\cdot)$ and $D(\cdot)$ are fixed, on any trial we have

$$p_c = \int_0^{\infty} P(\text{correct} \mid \mathbf{D} = x) dD(x) \quad (42)$$

$$= \int_0^{\infty} \left[S(x) + \frac{1}{2}(1 - S(x)) \right] dD(x), \quad (43)$$

$$p_e = \int_0^{\infty} \frac{1}{2}(1 - S(x)) dD(x). \quad (44)$$

Hence

$$p_c - p_e = \int_0^x S(x) dD(x). \quad (45)$$

Or, using integration by parts, and changing the variable from x to t :

$$p_c - p_e = \int_0^\infty s(t)[1 - D(t)] dt. \quad (46)$$

Now let $F_c(t) = P(L \leq t \mid \text{correct})$ and $F_e(t) = P(L \leq t \mid \text{error})$ denote the cumulative distribution functions of correct and incorrect response latency. Since $p_c F_c(t) = P(\text{Correct} \& L \leq t)$, we have

$$p_c F_c(t) = \int_0^\infty P(\text{correct} \& L \leq t \mid \mathbf{D} = x) dD(x), \quad (47)$$

where

$$P(\text{correct} \& L \leq t \mid \mathbf{D} = x) = \begin{cases} S(t) & \text{if } t < x \\ S(x) + \frac{1}{2}(1 - S(x)) & \text{if } t \geq x. \end{cases}$$

Consequently

$$p_c F_c(t) = \int_t^\infty S(t) dD(x) + \int_0^t \left[S(x) + \frac{1}{2}(1 - S(x)) \right] dD(x). \quad (48)$$

And in the same fashion

$$p_e F_e(t) = \int_0^t \frac{1}{2}(1 - S(x)) dD(x). \quad (49)$$

Subtracting (49) from (48), differentiating with respect to t , then multiplying by t and integrating, we obtain

$$p_c M_c - p_e M_e = \int_0^\infty ts(t)[1 - D(t)] dt. \quad (50)$$

Consequently for the Deadline model:

$$\frac{p_c M_c - p_e M_e}{p_c - p_e} = \frac{\int_0^\infty ts(t)[1 - D(t)] dt}{\int_0^\infty s(t)[1 - D(t)] dt}. \quad (51)$$

For a fixed $s(\cdot)$, the right side of (51) will not in general be invariant under arbitrary transformations of the function $D(\cdot)$. But note that in the specialization to the SFG model described above, $[1 - D(t)] = q$ throughout the interval on which $s(t)$ is positive: In that case, (51) equals the invariant quantity $E(\mathbf{S})$, i.e., mean SCR latency.

The fact that the Deadline model does not require a linear tradeoff function does not,

of course, make it any the less interesting. On the contrary, this model seems well worth further investigation, since it provides mechanisms for explaining both linear and nonlinear tradeoff functions in terms of the relationship between the distributions $S(\cdot)$ and $D(\cdot)$. A question of particular interest is whether one can separate these distributions and obtain an estimate of $S(\cdot)$ that does not depend on any specific assumptions about the form of $D(\cdot)$. I do not know whether this is possible.

REFERENCES

- EDWARDS, W. Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*, 1965, 2, 312-329.
- FALMAGNE, J. C. Stochastic models for choice reaction time with applications to experimental results. *Journal of Mathematical Psychology*, 1965, 2, 77-124.
- FISZ, M. *Probability theory and mathematical statistics*. New York: Wiley, 1963.
- GREEN, D. M., AND SWETS, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- LABERGE, D. L. A recruitment theory of simple behavior. *Psychometrika*, 1962, 27, 375-396.
- LABERGE, D. L., VAN GELDER, P., AND YELLOTT, J. A cueing technique in choice reaction time. *Perception and Psychophysics*, 1970, 7, 57-62.
- LAMING, D. R. J. *Information theory of choice reaction times*. London: Academic Press, 1968.
- LAMING, D. R. J. Subjective probability in choice-reaction experiments. *Journal of Mathematical Psychology*, 1969, 6, 81-120.
- OLLMAN, R. Fast guesses in choice reaction time. *Psychonomic Science*, 1966, 6, 155-156.
- STONE, M. Models for choice reaction time. *Psychometrika*, 1960, 25, 251-260.
- SWENSSON, R. G. The elusive tradeoff: Speed versus accuracy in choice reaction tasks with continuous cost for time. Technical Report No. 13, Human Performance Center, University of Michigan, 1968.
- YELLOTT, J. Correction for guessing in choice reaction time. *Psychonomic Science*, 1967, 8, 321-322.