4 experiments in simple and choice reaction time (RT) are reported. Experiment I examines the effect of monetary payoffs on the accuracy and variability of time estimation. Experiment II examines the effect of moving the position of a narrow payoff band along the time axis on the variability of observed RTs. This appears to alter the proportion of bona fide reactions (of low variability) and of more variable time estimates of the foreperiod duration. Experiment III is designed to assess the factors responsible for the increased mean and variability of choice compared with simple RT distributions. It it concluded that information processing rather than motor factors is primarily responsible for the difference between simple and choice RT. Experiment IV studies the relation between RT of correct and error responses as a function of variations in stimulus probability in a 2-choice RT paradigm. Finally, several theoretical distributions are evaluated by the empirical distributions obtained in Experiments II, III, and IV; none seems wholly satisfactory, but those with rounded modes and an exponential tail (e.g., the gamma) are clearly not adequate.

A common problem in simple RT experimentation is that of anticipations, or voluntary time estimates of the foreperiod. These may occur prior to the reaction signal, in which case they are normally unrecorded, or after the reaction signal, in which case they are recorded as particularly fast RTs. Investigators use various techniques to circumvent the problem of anticipations.

The authors gratefully acknowledge the following sources of support for this work: Theoretical and experimental support was provided by National Science Foundation Grant GB 1462 to the University of Pennsylvania and by Office of Naval Research, Contract No. NONR 477(34) to the University of Washington. The extensive computer time was provided by National Institutes of Health Grant FR-15 to the Johnson Research Foundation, School of Medicine, University of Pennsylvania. A short report of this work was read to the Fifth Scientific Meeting of the Psychonomic Society, Niagara Falls, Ontario, Canada, October 1964.
unctions. These include elimination of responses prior to some time, usually approximately 100 msec., the so-called "irreducible minimum" (Woodworth & Schlosberg, 1954); employing variable rather than constant foreperiods; and introducing catch trials on some proportion of the trials (Drazin, 1961). Because it is questionable whether these techniques completely eliminate anticipations, the first two experiments were run. The first was designed to estimate the variability and accuracy of time estimation when payoffs with feedback are employed and the second to devise a payoff scheme which would simultaneously reward the fastest possible true reactions to the signal and penalize anticipations and slow responses.

GENERAL METHOD

Subjects.—Ten university undergraduates, three of whom were female, (CP, EM, and SR), served as Ss. The Ss BW and GC participated in Exp. I; Ss AB, ES, KM, and CP in Exp. II; AH, MO, and EM in Exp. III; and AH and SR in Exp. IV. Note that AH participated in both Exp. III and IV. All Ss had normal hearing and were right-handed, except for ES who nonetheless used his right hand to respond. They were paid as a result of their performance in the experiments, and in some cases they received additional payments for their time. Data for each S are analyzed separately.

Apparatus.—The Ss sat in an Industrial Acoustics Company acoustic chamber which attenuated outside noise by 30-65 db. over the audible-frequency range from 100 hz. The warning and reaction signals were pure tones generated by an oscillator and presented binaurally through Permofox PDR-8 earphones mounted in semiplastic cushions. Except in Exp. I and for S AB in Exp. II, the tones were turned on and off by a sine-zero switch when the sine wave crossed zero amplitude with positive slope. Tone duration was controlled by counting cycles of a 1,000-hz. tone generated by a second oscillator. In Exp. I, III, and IV, and for AB in Exp. II, the tone duration was 143 msec.; it was 124 msec. for remaining Ss in Exp. II. The warning signal (Sw) preceded either reaction signal (Sb, Sb) by 2,000 msec. (±5 msec.). The interval between the two signals was controlled by a synchronous motor-driven cam switch. The mechanical nature of this system prevented us from controlling the foreperiod interval more exactly. Tone frequency was controlled by a Krohn-Hite programmable oscillator. In Exp. I and II, Sw and Sb were 1,000-hz. tones; in Exp. III, Sb and 1,100 hz., Sb was 1,000 hz., and Sb was 1,200 hz.; in Exp. IV, Sb was 1,000 hz., Sb was 900 hz., and Sb was 1,100 hz.

The intensity of all tones was .030 v., or about 30 db. above .001 v. However, because the tones were not filtered, there was an unmeasured acoustical click at the beginning and end of the tones. All Ss in Exp. I and II, and SR in Exp. IV used a pair of hand-held joined microswitches actuated by typewriter-type keys. All Ss in Exp. III and AH in Exp. IV used two 6 in. × 1 in. Plexiglas keys which were set flush in a response board and were separated by a short partition. These keys were adjusted to require a small amount of pressure and a very short travel. The typewriter keys required somewhat more travel.

Times were measured to the nearest msec, by a Hewlett-Packard electronic counter and were recorded on an associated printer. In Exp. I, the counter was actuated by Sw. This could have been done in the later experiments and would have enabled us to record the times of responses prior to Sb, but it had the disadvantage of degrading the accuracy of measured RTs from Sb to ±5 msec. because of the inaccuracy of the cam controlling the onset of Sw. To increase the accuracy of measured RTs in Exp. II through IV, the timer was actuated by a 30-v. pulse that accompanied the onset of the reaction tone; this reduced the measurement error to ±1 msec. In these experiments, the occurrence, but not the time, of responses prior to Sb was recorded.

The presentation of Sb and Sa in the choice experiment was controlled by a punched tape read by a Western Union Teletype tape reader. Tapes of 500 random choices with P = .1, .3, .5, .7, and .9 were prepared subject to the condition that the proportion of Sb signals within each block of 100 trials exactly equalled P.

A display panel in front of S provided information feedback after every trial. The form of this information is described for each experiment. The amount of money won or lost on a trial was presented by
ACCURACY AND VARIABILITY OF REACTION TIMES

means of projection units in the panel. In Exp. IV, a bulb above the correct key was illuminated as soon as S responded to the reaction signal.

Procedure.—Explicit payoffs were presumed to fulfill part of the function of the usual instructions; therefore, the instructions consisted primarily of a detailed explanation of the payoffs to be used. The S was given as much information as possible about the experimental condition in the form of graphs, payoff matrices, information about the presentation probabilities of the signals, etc. In both the simple and choice situations, bands of latencies were differentially rewarded. If S responded either before or after the band he either lost money or failed to win anything. In the simple reaction conditions, payoffs were for latency alone, and in the two-choice conditions, for both latency and accuracy. The S was encouraged to earn as much money as he could.

After each trial, S was informed of the amount of money won or lost by that response. In Exp. III this information was sufficient for S to know whether he was too fast, too slow, or his relative position within the band. For Exp. II and IV, in which the payoff band was symmetric, the position on the display panel of the amount of money lost informed S whether he was too fast or too slow relative to the payoff band. In Exp. III, where payoffs for speed and accuracy were separate, the payoff for latency was displayed first, followed by that for accuracy.

The S was allowed as much practice as he wished before each new experimental condition. Brief rest periods were given after each block of approximately 100 trials or at S's request. During each rest, S was told the total amount of money that he had won or lost during the preceding block. Experimental sessions consisted of approximately 500 responses, excluding practice trials, and lasted between 1 and 11 hr.

EXPERIMENT I

Previous Es have investigated the relation between the variability and constant error of time estimates and the duration of the interval being estimated. The standard deviation of time estimates was found by Woodrow (1930, 1933) to be approximately 10% of the mean duration for intervals between 1 and 2 sec., in the absence of feedback and payoffs. Klemmer (1957) analyzed the joint effect of foreperiod (FP) variability and time estimation variability within the framework of information theory. He found mean RT to be a linear function of stimulus uncertainty, defined as the sum of the variance of S's time estimates and the variance of the FP distribution.

Procedure.—The two Ss, each of whom had previously made about 20,000 RTs, were each instructed to try to estimate times by responding when they felt that a prescribed time period had elapsed from the onset of a 1,000-hz. tone. The actual elapsed times were shown to S in milliseconds on the digital readout of the Hewlett-Packard timer. The S was paid off for accuracy as follows: .94 for responses with ±20 msec. of the goal, .84 for responses in the adjoining 20-msec. bands, and so on, with no payoffs for re-

| Table 1 |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| **Target Interval in msec.** | **BW** |     |     | **GC** |     |     |     |     |
|                  | μ   | σ   | σ/μ |      | μ   | σ   | σ/μ |     |
| 600              | 601.7 | 75.2 | 0.125 | 601.4 | 49.5 | 0.083 |
| 1000             | 991.4 | 119.3 | 0.120 | 1060.5 | 97.3 | 0.092 |
| 1900             | 1914.1 | 154.8 | 0.081 | 1894.8 | 209.7 | 0.111 |
| 3500             | 3504.3 | 401.8 | 0.115 | 3464.9 | 335.4 | 0.097 |
| 5000             | 4853.7 | 1536.8 | 0.316 | 4923.5 | 275.3 | 0.056 |
| 1900             | 1937.5 | 160.0 | 0.083 | 1898.0 | 160.6 | 0.085 |

* The first response was excessively long and has been omitted.
sponses more than 180 msec. from the goal. Each run consisted of 500 estimates of a given time; these times were 1.9, .6, 1.9, 3.5, 5.0, and 1.9 sec., in that order.

Results and discussion.—The distributions of responses were sufficiently symmetric so that, for our purposes, it is adequate to report only the means and standard deviations, which are shown in Table 1. Except, perhaps, for the longest time, for which BW seems excessively variable and GC excessively consistent, these data are stable and systematic. The stability of the repeated value, 1.9 sec., seems most satisfactory. In addition, plots of the trial-by-trial estimates do not exhibit any apparent time trends within a run. The distribution for the shortest time, .6 sec., is highly peaked at the mode, similar to the RT distributions shown later; however, unlike them, it is symmetric about the mode.

The primary conclusion is that, within the range from .6 to 5.0 sec., Ss can time estimate quite accurately, but that the standard deviation of their responses is of the order of 10% of the mean, even when payoffs are employed. This figure is comparable to that obtained by Woodrow (1930, 1933) for intervals of comparable duration obtained without payoffs or feedback. As we shall see, this relative variability is considerably in excess of that for RT when time is measured from the warning signal; therefore, however Ss may use time estimation in our later experiments when we ask them to react, it cannot be that they simply estimate the length of the foreperiod on every trial.

Experiment II

As noted previously, RT investigators have used various techniques, among them variable FPs and catch trials, to experimentally eliminate time estimates of the FP duration. We consider here an alternative scheme—that of rewarding Ss for responding within a latency band centered at $T$, where $t$ (the foreperiod duration) is fixed and $T$ is varied over a range from times less than the (apparent) true reaction time to times considerably greater. The S's task under the band payoff then becomes one of reacting consistently within the band, for he is only rewarded for those responses occurring within the latency band and punished for responses which are too fast or too slow for it.

If we assume that two response mechanisms are available, namely, S can either time estimate from $S_0$ or $S_1$ or react to $S_1$ and if further we assume as the data of Exp. I indicate that the variability of the time estimates increases with the time to be estimated, then we can predict qualitively how the variability of the RT distribution must change with $T$. For explicitness, although it is not essential to the argument, suppose that, as is approximately true (see Table 1), the standard deviation of time estimates is a fixed proportion $k$ of the mean estimate; according to our data $k$ is approximately .1. When $T$ is very short—so short that if S reacts to $S_1$ he will necessarily be slower than the payoff band—then he has no reasonable option but to time estimate from $S_0$, and so the standard deviation of the RT distribution should be approximately $k(t + T)$. As $T$ is increased sufficiently, reactions to $S_1$ become feasible. Since the variability of the reactions is less than that of the estimates the observed variability should decrease, reaching a minimum when all responses are reactions. With further increases in $T$, the reactions will fall short of the payoff band, and S is again forced to time estimate, but now it is to his advantage to estimate from $S_1$ rather than $S_0$ (because the variability
ACCURACY AND VARIABILITY OF REACTION TIMES

is less and so a positive payoff is more likely). The standard deviation of the RT distribution should be approximately $kT$. According to this argument, then, a plot of the variability of the RT distribution vs. $T$ should be U-shaped with (for all $T < t$) a minimum located approximately at the "true" reaction time.

Procedure.—Three Ss were run under each of six different payoff bands: (Band 1) 85-104 msec, (Band 2) 105-124 msec,. . . (Band 6) 185-204 msec. A fourth S, AB, was run only on Bands 2-5. The S won 2¢ for responding within the band and lost 1¢ for responding outside it. For AB, the bands were presented in random order; for the other three Ss, the order of presentation of bands was constrained as follows: no band was followed by an adjacent band; difficult bands (1, 2, and 6) alternated with easy ones; and the first band in the series was always easy. These three Ss had three practice sessions, consisting of 500 trials each, during which the 6 bands were used. Their order of use was different from that of the experimental sessions, but subject to the same restrictions. Each daily experimental session consisted of 500 trials under a single band.

Analyses are based upon distributions obtained from individual Ss under a single band. Each is based upon 500 trials, except for AB under Band 2, where only 190 observations are included in the analysis because the remaining RTs were mistimed.

Results and Discussion.—The proportion of responses prior to $S_1$ was highest for Band 1 (8.8%, 4.0%, and 3.0% for Ss KM, CP, and ES, respectively), next highest for Band 2 (.5%, 2.8%, .0%, and 1.2% for AB, KM, CP, and ES, respectively), and was negligible for Bands 3-6.

Figure 1 presents the means and the proportion of reactions occurring within the payoff band for each S as a function of band position. The latter seems to be a better measure of variability in this situation than the empirical standard deviation because, as we show later, the RT distributions may have high tails-falling off as a negative power of $t$ rather than as an exponential—in which case the empirical variance is a rather variable statistic.

It is clear from Fig. 1 that the mean RT follows the band position closely.
Except for the two fastest bands, all means are well within the band range. In terms of our measure of variability—the proportion of responses within the band—the least variable band for CP is 2, for ES and KM is 3, and for AB is 4. With the empirical variance, the minimum band is shifted from 3 to 5 for KM and the others are unchanged. Note that the maximum seems fairly clearly defined and that, for the best bands, a surprisingly high proportion (50-65%) of the responses lie within its narrow, 20-msec. width. Moreover, pressing for short times seems to force S's increasingly to time estimate from $S_0$ as judged both by the increased frequency of unambiguous anticipations (responses before $S_0$) and by the marked increase in variability. This shift to much greater variability appears to occur in the neighborhood of 90-100 msec., which is near the value of the irreducible minimum defined by others. In summary, then, although the observed RT distribution is affected by outcomes, this may be due entirely to replacing reactions by time estimations without in any way affecting the magnitude of the "true" RT.

**Experiment III**

We consider next the source of the difference between simple and choice RT. Not only are choice RTs longer on the average than simple ones, they are also more variable (Woodworth & Schlosberg, 1954). It seems desirable progressively to complicate the simple reaction until we reach the choice reaction and to measure both the mean and variability of the RTs for each of the successive experimental steps to determine which difference or differences between the two are crucial.

**Procedure.**—The Ss were run in situations of increasing complexity, the simplest being a single reaction to a single tone and the most complex, a two-choice reaction to one of two tones. Two relatively wide, non-overlapping, adjacent payoff bands were used. A "fast" band differentially rewarded responses between 100 and 200 msec., and a "slow" band differentially rewarded responses between 200 and 300 msec. Responses prior to the band were fined 5¢; those in the first third won 3¢, in the second, 2¢, and in the last third, 1¢; those slower than the band received no payoff at all.

Distributions of RTs were collected under the following procedures:

1. **Simple—1.** One signal, one response, a single payoff band rewarding all responses.
2. **Simple—2.** Two signals, one response, a single band rewarding all responses regardless of which signal was presented.
3. **Recognition.** Two signals, one response, the fast payoff band rewarding responses to one signal and the slow band rewarding responses to the other.
4. **Choice.** Two signals, two responses, a single band rewarding all responses, regardless of accuracy. In addition, $S$ won 1¢ for each correct response and lost 3¢ for each incorrect response.

Recognition is so named because in order to win money, $S$ must first recognize whether the signal presented is the fast one (i.e., the

<table>
<thead>
<tr>
<th>Condition</th>
<th>Response Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td>Simple-1</td>
<td>(1) S₁→F</td>
</tr>
<tr>
<td></td>
<td>(3) S₂→S</td>
</tr>
<tr>
<td>Simple-2</td>
<td>(5) S₁→F</td>
</tr>
<tr>
<td>Recognition</td>
<td>(7) S₁→F</td>
</tr>
<tr>
<td>Choice</td>
<td>S₁→F... (9)...S₂→F</td>
</tr>
<tr>
<td></td>
<td>S₂→S... (10)...S₁→S</td>
</tr>
</tbody>
</table>

Note.—S₁ = 1,000-hz. tone, S₂ = 1,200-hz. tone, F = fast band payoff, S = slow band payoff, numbers in parentheses are names for conditions.
one to which the fast payoff applies) or the slow one, for if he reacts faster than 200 msec. to the slow one, he loses 5¢. The S indicates that he has recognized the signal by making differential temporal responses with the same key. We define recognition time as the time to respond to the fast signal.

With two signals, one low (1,000 hz.) and one high (1,200 hz.) frequency, two payoff bands, fast and slow, and two response keys, right and left, each procedure generates a variety of conditions. The 10 shown in Table 2 were run. Of these, 6 result in two distributions of RTs, broken down by the signal presented and/or the response made; thus, there are a total of 16 distributions which are available for comparison. Of these, only the 8 run under the fast payoff band are relevant to our original question about differences between simple and choice RT. Conditions using the slow payoff band were included for practice in the recognition condition, where both payoffs were used simultaneously, and to answer the general question of what, if any, change in the form of the distribution occurs under a simple shift of the payoff function along the time axis.

Each condition was run during each of eight experimental sessions, resulting in 16 distributions (two for each stimulus) of 240 RTs each. Certain restrictions were imposed upon the order of presentation of the 10 conditions within a session: any recognition or choice condition was preceded by both of the relevant simple reaction conditions, e.g., Cond. 7 was always preceded by both 1 and 3; and the recognition and choice conditions were always preceded by approximately equal numbers of fast and slow payoff conditions. Because both Simple-2 conditions were run with fast payoffs only, these always appeared at the end of the session. With these restrictions, 16 orderings of the conditions were possible, of which 8 were selected. These were presented in a random order to the 3 Ss.

Results and discussion.—A session-by-session analysis of median RTs for all conditions indicated that the data for AH and MO were stable over sessions, but the median RT for EM declined appreciably from Day 1 to Day 2 after which it was stable. Therefore, the results reported for EM are based upon only her last seven sessions, and so her RT distributions consist of 210, rather than 240, observations.

Figure 2 shows the mean RT for the four conditions run under the fast payoff band and the three conditions run under the slow payoff band. Results for the fast payoff are unequivocal: mean RTs for both recognition and choice conditions are much longer than those for either of the simple conditions; indeed, recognition and choice-mean times are not even within the fast payoff region. The position of the payoff band affected RT means for the simple conditions markedly, but it

![Fig. 2. Mean RTs of individual Ss for conditions under the fast and slow payoff bands. (Conditions are: S-1 = Simple-1; S-2 = Simple-2; R = Recognition; and C = Choice.)](image-url)
changed those for the recognition and choice conditions little, if at all. The choice condition results in somewhat longer mean RTs than recognition under both payoff conditions.

The standard deviations for the same conditions exhibit a similar pattern to that of the means (Fig. 3). Again, the two simple conditions are very similar under fast payoffs, both having relatively low variances, whereas recognition and choice are more variable and do not appear to differ much from one another. Location of payoff band has little effect on variances of recognition and choice, but it affects the variance of the simple distributions markedly. In fact, variances of the simple condition under the slow payoff are not much smaller than those of the recognition distribution for two Ss, and the variance for one S in the simple condition is actually larger than the corresponding choice variance under the slow payoff. Thus, it appears that lengthening simple RTs by payoffs adds almost as much variability as does increasing the complexity of the decision.

We conclude from the results of this experiment that the major part of the increase in time and variability between simple and choice RT is contributed by the process of recognition.

Experiment IV

We have already seen that increasing the number of responses from one to two, i.e., from simple to choice RT, increases both the mean and variability of the RT distribution. We next consider the effects of different presentation probabilities of the two signals on the mean and variability of both the correct and error distributions.

Several investigators have found that changes in presentation probabilities of fixed sets of signals had effects on the latencies of correct responses (Crossman, 1953; Fitts, Peterson, & Wolpe, 1963; Hyman, 1953). Specifically, the mean RT of the correct response to a signal decreases as the probability of that signal being presented is increased. Typically, the results have been interpreted within the theoretical framework of information theory. One characteristic of the method of data analysis usually employed has been to average across signal-response pairs; when this is done, a measure of redundancy or information transmitted by the selection of a signal is linearly related to the mean RT over all signals. However, the information measure does not accurately predict the mean RT of responses to specific signals (Hyman, 1953). Various models of choice RT predict that stimulus presentation probability and latencies of correct responses will be inversely related and sometimes they specify the form of the latency distribution (e.g., Audley, 1960; LaBerge, 1962; Luce, 1960; Stone, 1960). However, it is often difficult to determine what they predict about the latencies of error responses.

One problem in evaluating these models, or of developing new ones, is the lack of detailed empirical information about the relation between the latencies of erroneous and correct responses and about the forms of the resulting empirical distributions.

The purpose of the present experiment is to obtain data for which detailed relationships between correct and error latencies can be determined, and some attributes of the distributions observed.

Procedure.—Five presentation probabilities, .1, .3, .5, .7, and .9, were used with the two Ss. Only one probability was used during a single experimental session. The S was informed of its value at the beginning of the session. The same payoff function was used with each probability. The S won 1¢ for responding correctly from 150 to 200
msec., lost $1\$ for responding incorrectly in the same band, and received no payoff for all other responses. This relatively difficult payoff function was chosen to produce high error rates. When $S$ failed to respond within the payoff band, he received information about the direction of his error.

Results and discussion.—The data may be summarized in a $2 \times 2$ table in which each cell has two entries, the conditional response probability $\text{Pr}(R_i/S_i)$ (here denoted $p_{ij}$) and the RT distribution $R(y)(t)$. The relation between $p_{11}$ and $p_{21}$ that is generated by varying presentation probability ($P$), the isosensitivity plot, is shown in Fig. 4 for each $S$. These data are similar to those reported in the signal-detection literature where $\text{Pr}(\text{Yes}/\text{Signal})$ and $\text{Pr}(\text{Yes}/\text{Noise})$ covary with the presentation probability and payoffs. In contrast to the signal-detection results, however, the points generated by our $S$s do not appear to lie on a single isosensitivity curve. In particular, the $P = .5$ point for AH and the $P = 7$ point for SR lie closer to the chance diagonal than do the others. This anomaly presumably arises because the two reaction signals are, in fact, perfectly discriminable when $S$ takes enough time to respond, and the errors arise because $S$ decides to take less information than is necessary for more accurate responding. If this interpretation is valid, the discriminability of the signals is essentially under $S$'s control, rather than $E$'s, and our $S$s apparently varied the discriminability from probability to probability by observing for shorter or longer times, thus causing a shift from one isosensitivity curve to another.

Next, we examine mean RT as a function of several probabilities. A given RT distribution $R(y)$ has at least three probabilities associated with it: (a) the presentation probability, $\text{Pr}(S_i)$, (b) the unconditional response probability, $\text{Pr}(R_i)$, and (c) the conditional response probability, $p_{ij}$. In the ordinary choice RT experiment in which errors are negligible, only the correct latencies are shown, usually as a function of the presentation probability. Because in nearly error-free performance response and stimulus probability are nearly identical, a plot of RT vs. response probability would be about the same. In this experiment, presentation and response probabilities for the correct responses are almost identical, because our $S$s tended to match response probabilities with presentation probabilities. However, for incorrect responses, response and presentation probabilities are distinctly not the same. Figure 5 shows mean RTs (filled circles) and error RTs (open circles) as a function of each of the three probabilities. It is clear that the most nearly identical functions for both error and correct latencies occur when they are plotted against unconditional response probability.

![Fig. 4. Iso-sensitivity plots relating, $\text{Pr}(R_i/S_i)$ to $\text{Pr}(R_i/S_2)$ as a function of presentation probability, $P_i$ of $S_i$, for each $S$ separately. (The value of $P$ is shown in parentheses beside each point. The dotted line indicates chance performance.)](image-url)
However, even across a fixed response probability there seem to be nontrivial differences between the error and correct latencies. Figure 6 presents in greater detail the relation between the means and the standard deviations of correct and error distributions for the nonsymmetric probabilities $P = .1, .3, .7,$ and $ .9$ for AH. The data are plotted in terms of the more prob-
variable stimulus, $S_1$ or $S_2$, depending on $P$. For this $S$, the ordering of the means and standard deviations is consistent but different: if we define $i$ and $j$ so that $Pr(S_i) > Pr(S_j)$, then the order is $ii$, $ji$, $ij$, and $jj$ for the means and $ij$, $ii$, $ji$, and $jj$ for the standard deviations. These relations for the other $S$, although in the same general direction, are not as consistent and they are not shown.

Two aspects of these data are of relevance in evaluating or developing models of the choice process. The first is that latencies of error and correct responses are more closely related to responses than to stimuli. This seems to be true whether the errors are generated by time pressure, as in this RT experiment, or by low stimulus discriminability, as the experiments of Carterette et al. (1965) suggest. The second finding is that error latencies tend to lie intermediate to correct latencies.

**Form of the RT Distribution**

A problem which has remained unsolved in spite of its long history is the form or forms of the underlying RT distribution. Perhaps the earliest proposal was the normal distribution; however, RT distributions, especially simple ones, are notoriously skewed and are bounded from below. These facts led Woodworth and Schlosberg (1954) to propose the log normal as a substitute.

Conceptions of the reaction process as a series of intervening, nonobservable steps in a chain have been the basis of many recent theoretical proposals. In particular, if one assumes that the overall RT results from a finite sum of independent, identically distributed exponential random variables, the gamma distribution results. If these random variables are not identically distributed, but still exponential, the general Erlang or generalized gamma distribution results (McGill & Gibbon, 1965). Under fairly general conditions, the normal distribution results if a very large (infinite) number of independent, identically distributed components are summed (central-limit theorem).²

The gamma distribution has been suggested for RT by Bush and Mosteller (1955), Christie (1952), Luce (1960), and Restle (1961). A discrete analog of the gamma, the negative binomial, was derived by LaBerge (1962) in his theory of RT.

As a preliminary step, the empirical RT distributions collected in Exp. II, III, and IV were classified as to Pearson type by the procedures described in Appendix A. As noted there, none were classified as ordinary gamma and only 13 were classified as generalized gamma. A striking aspect of the results of the analysis is that the simple RT distributions in particular are far outside the limits for the gamma. In fact, by inspecting the simple distributions themselves, it becomes evident that any Pearson distribution is unlikely to account for them: they are both too peaked and have too high tails (which are usually found in RT distributions and have sometimes been attributed to experimental error) to be fit by distributions with rounded modes and exponential tails.

In lieu of any suitable theoretical alternative, we attempted to fit the double monomial (DM) to our RT distributions. The DM was derived by Luce and Galanter (1963, p. 288) as the form of a generalization function in magnitude estimation. It consists of two power functions that meet at the mode. As a result it has two qualitative features of our empirical distributions—peakedness at the mode and high tails (in the sense that a negative power approaches zero much more slowly than an exponential). The density func-

² It should be pointed out that if the commonly accepted assumptions of the central-limit theorem are altered, a variety of other latency distributions can arise (Gnedenko & Kolmogorov, 1954); however, the results are stated in characteristic function form and are not readily translatable back into distribution terms. This, however, is an important direction that must ultimately be clarified by those interested in the form of the RT distribution.
tion of the DM is:

\[ f(t) = \frac{(\delta + 1)(\epsilon - 1)}{\epsilon + \delta} \]

\[ \times \begin{cases} 0 & \text{if } t < 0 \\ \left( \frac{t}{t_0} \right)^{\delta} & \text{if } 0 \leq t \leq t_0 \\ \left( \frac{t}{t_0} \right)^{-\epsilon} & \text{if } t_0 \leq t, \end{cases} \]

where \( t_0 \) is the mode.

In addition, we have fit these data by the ordinary gamma distribution with its origin located at a value \( t_0 \), not necessarily 0, which we here call the displaced gamma. The density function of the displaced gamma is:

\[ f(t) = \frac{\lambda^\alpha (t - t_0)^{\alpha - 1} e^{-\lambda(t-t_0)}}{\Gamma(\alpha)}. \]

Procedures for estimating the three parameters for each distribution are given in Appendix B.

Table 3 presents the results of the comparison between likelihood values for the two theoretical distributions. The DM fits the simple RT distributions and the gamma the choice distributions more often by this test. We do not know, however, how good a test of goodness-of-fit the likelihood comparison is.

Even though Table 3 indicates that the gamma distribution does better than the table:

<table>
<thead>
<tr>
<th>Exp.</th>
<th>Type</th>
<th>Number of Cases Favoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Double Monomial</td>
</tr>
<tr>
<td>II</td>
<td>simple</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>simple</td>
<td>20</td>
</tr>
<tr>
<td>III</td>
<td>recognition</td>
<td>7</td>
</tr>
<tr>
<td>IV</td>
<td>choice</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>choice*</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>53</td>
</tr>
</tbody>
</table>

*Only distributions of more than 50 observations were analyzed; these include some distributions of incorrect responses.

Fig. 7. Two empirical RT distributions (solid lines) are plotted twice, each time with a best fitting theoretical density function (dotted lines). (The left pair are fitted by the DM, the right pair by the gamma. The theoretical mode for the DM and the theoretical starting point for the gamma are indicated by \( t_0 \) on the abscissas.)
DM on almost half of the empirical distributions, plots of the fits of the theoretical distributions to the empirical data points, such as those shown in Fig. 7, are disturbing. The distribution in the lower half of the figure is best fit (by the likelihood criterion) by the gamma and yet to the eye the DM seems to be the better representation. The difference in evaluation arises from the details of the tails, which are difficult to see by eye and which may be of secondary importance in many contexts.

Table 4 shows the results of $\chi^2$ tests on the DM. In the overwhelming majority of the narrow band, simple RT distributions from Exp. II, the DM is rejected at the .05 level, whereas the wide band simple distributions from Exp. III were fit more often than not.

Although the DM distribution is not completely satisfactory as judged by $\chi^2$, the source of its failure is interesting and makes clear that the classically proposed distributions, in particular the gamma, would fare even worse. The maximum likelihood estimates of the parameters take into account both the tails and the mode of the distributions, and to get the tails of the DM sufficiently high ($\delta$ and $\epsilon$ sufficiently small) the mode is not sufficiently peaked. In addition, the empirical distributions tend to be rather more jagged than would be predicted by a smooth, continuous density function and such large sample sizes. However, in light of the two processes in simple RT—reaction and time estimation—for which evidence was obtained by Exp. II, this fact is not surprising. Perhaps if we were able to obtain RT distributions uncontaminated by time estimates, the DM would fit better.

**Conclusion**

The most striking conclusion we draw from these studies is that, even though our data are highly systematic, we do not have an adequate theoretical understanding of them. It is clear that Ss respond sensibly and effectively to differential temporal payoffs and that some use is probably made of time estimation as well as of reactions, but we do not understand the detailed mechanisms very well. This may be a result of faulty experimental procedures; for example, some Ss probably used hand position to control their RTs in some payoff conditions. However, we cannot hide all of our theoretical difficulties behind experimental problems. The form of the individual empirical RT distribution is another indication that existing theories are probably inadequate. These distributions appear simultaneously to have high tails and a peaked mode. If this is a reliable finding, summary statistics based on the moments are much less satisfactory than those, such as median, interquartile range, and percentage of responses in an interval, that are less sensitive to the occasional very long RTs.

In addition to the regularity that RT exhibits as a function of several experimental manipulations—including the location of the payoff bands and the type of reaction required—there are also systematic trading relations between pairs of psychological measures, such as RT and response probability. Since comparable results have been found by other investigators and since similar trading rela-

---

**Table 4**

RESULTS OF $\chi^2$ TESTS BETWEEN DM AND EMPirical DISTRIBUTIONS

<table>
<thead>
<tr>
<th>Type</th>
<th>Number of Distributions Which are</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not Rejected</td>
</tr>
<tr>
<td>Exp. II simple</td>
<td>2</td>
</tr>
<tr>
<td>Exp. III simple</td>
<td>17</td>
</tr>
<tr>
<td>recognition</td>
<td>6</td>
</tr>
<tr>
<td>choice</td>
<td>6</td>
</tr>
<tr>
<td>Exp. IV choice*</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
</tr>
</tbody>
</table>

* Only distributions of more than 100 observations were analyzed; these include some distributions of incorrect responses.

* $p < .05$. 

---

* We did not calculate $\chi^2$ for the best fitting gamma distributions because, in the light of this analysis, the added computer cost did not seem warranted.
tions have proved most illuminating in detection work, they probably should be pursued much more intensively within the RT context.

REFERENCES


APPENDIX A

PEARSONIAN CLASSIFICATION OF EMPIRICAL DISTRIBUTIONS

From the moments of the empirical distributions, values of $\kappa$ were calculated as described in Kendall and Stuart (1961), where

$$\kappa = \frac{[\beta_1(\beta_2 + 3)]}{[4(4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)]}$$

and

$$\beta_1 = \frac{\mu_x^2}{\mu_x^2}; \quad \beta_2 = \frac{\mu_x^4}{\mu_x^4}$$

The value of $\kappa$ is used to classify the empirical distributions as to type of Pearson distribution. For $\kappa$ less than 0 or greater than 1, the distributions are classified as beta (of the first and second kind, respectively); for $\kappa$ between 0 and 1, the distribution
ACCURACY AND VARIABILITY OF REACTION TIMES

TABLE A

| Type of Pearson Distribution as Determined by Value of Statistic $\kappa$ |
|-----------------------------------|------------------|------------------|
| Exp. II                           | Type IV          | Beta I, II        |
| simple                            | 14               | 8                |
| Exp. III                          | 14               | 10               |
| simple                            | recognition      | 7                |
| choice                            | 4                | 8                |
| Exp. IV                           | choice*          | 14               |
| Total                             | 53               | 55               |

* All distributions, including error responses, were analyzed.

is classified as Type IV; and for $\kappa \to \infty$, the distribution is gamma. Table A summarizes the results of this analysis.

Instances of each of the three distributions other than the gamma were found. Approximately equal numbers of our distributions were classified as Type IV and beta and instances of both types of betas appeared. There is a suggestion in Table 5 that more simple distributions are of Type IV and more choice distributions are beta.

It is not surprising that no instances of the gamma distributions were found because $\kappa = \infty$ (the gamma distribution) corresponds to an exact relation among the moments. Unfortunately, however, there is no way of determining, from the value of $\kappa$ alone, how close a distribution approaches the gamma. Sternberg (1964) has recently shown that certain relations must hold among the moments for a distribution to be gamma. Specifically, for the ordinary gamma (equal time constants for each of $c$ exponential steps where $c$ is an integer)

$$\beta_1 = 4/c \quad \text{and} \quad \beta_2 = 3 + 6/c.$$  

This describes a series of points lying on the line satisfying

$$\beta_2 = 3 + (3/2)\beta_1.$$  

For nonintegral $c$ and equal time constants, the distributions lie between these points, on the same line. Sternberg's results are consistent with Pearson's in that when the relation is satisfied, $2\beta_2 - 3\beta_1 - 6 = 0$ and $\kappa = \infty$.4

For the generalized gamma (unequal time constants), he has shown that the following inequalities must be satisfied:

$$0 < \beta_1 \leq 4$$
$$3 + (3/2)\beta_1 \leq \beta_2 \leq 3 + (27/2)\beta_1.$$  

Figure A1 shows the relation between $\beta_1$ and $\beta_2$ for the empirical distributions collected in Exp. II, III, and IV. The insert covers the region near and within the limits for the generalized gamma. The normal distribution lies at the point $\beta_1 = 0, \beta_2 = 3$. It is clear that although a number of distributions lie within or near the region of the generalized gamma, none of them falls on the line for the ordinary gamma. By Sternberg's criteria, only 13 distributions are generalized gamma distrib-

* It is also, of course, true that if $4\beta_4 - 3\beta_2 = 0$, then $\kappa = \infty$; however this relation between $\beta_1$ and $\beta_2$ satisfied neither Sternberg's nor Pearson's restrictions (Kendall & Stuart, 1961, p. 175).
APPENDIX B

Estimation Procedures

1. Double monomial. If \( N \) is the number of observations, the maximum likelihood estimates of the three parameters \( N_0 \) (rank of the RT corresponding to \( t_0 \)), \( \delta \) and \( \epsilon \) are easily shown to satisfy:

\[
N_0 = N \left( \frac{e - 1}{e + \delta} \right) \quad \delta = \frac{1}{\log t_0 - \frac{1}{N_0} \sum_{i=1}^{N} \log t_i} - 1, \quad \epsilon = \frac{1}{N - N_0} \sum_{i=N_0+1}^{N} \log t_i - \log t_0 + 1
\]

where the logarithms are natural ones. We used \( N/2 \) as a first estimate of \( N_0 \), and the three formulas were iterated by computer until the estimate of \( N_0 \) converged or 20 iterations were completed. In very case in which \( N_0 \) did not converge, it oscillated between two integers. In these cases we calculated maximum likelihood and \( \chi^2 \) values for the two sets of three parameter values and their means, and that set of values yielding the highest value of likelihood was chosen.
2. Displaced Gamma. Maximum likelihood of $t_0$, $\lambda$, and $\alpha$ satisfy:

$$\frac{1}{2\alpha} + K = \log \frac{1}{N} \sum_{i=1}^{N} (t_i - t_0) - \frac{1}{N} \sum_{i=1}^{N} \log(t_i - t_0),$$

$$\lambda = (\alpha - 1) \frac{1}{N} \sum_{i=1}^{N} \frac{1}{t_i - t_0}, \quad t_0 = \frac{1}{N} \sum_{i=1}^{N} t_i - \frac{\alpha}{\lambda}$$

where $^5$

$$\Gamma(\alpha) = \alpha^{-\alpha} e^{-\alpha} \sqrt{2\pi} \left[ 1 + \frac{1}{12\alpha} + \frac{1}{288\alpha^2} \right]$$

$$K = \frac{1}{\alpha} \log \left( 1 + \frac{1}{12\alpha} + \frac{1}{288\alpha^2} \right).$$

To solve for $\alpha$, we iterated the first formula, where the $(n - 1)$st value of $\alpha$ was inserted in $K$. The procedure was terminated when the difference between the $n$th and $(n - 1)$st value of $\alpha$ was less than $10^{-7}$.

In contrast to the DM, in all cases the three equations failed to converge on $t_0$. Therefore, we explored the possible range of reasonable values of $t_0$, calculated a likelihood value for each set of parameters, and chose that $t_0$ and corresponding values of $\alpha$ and $\lambda$ for which the likelihood was a maximum. The possible range of $t_0$ was taken to be from $-2,000$ msec. (i.e., the onset of $S_0$) to $t_1 - 1$ msec., where $t_1$ is the value of the shortest RT in the empirical distribution.

(Received August 19, 1966)

$^5$ An abbreviated form of Stirling's approximation for $\Gamma(\alpha)$ was used to obtain $K$: 

$^1$