08773-33-T

HUMAN PERFORMANCE CENTER

DEPARTMENT OF PSYCHOLOGY

The University of Michigan, Ann Arbor

The Elusive Tradeoff: Speed Versus Accuracy in Choice Reaction Tasks with Continuous Cost for Time

RICHARD GAVIN SWENSSON

THE UNIVERSITY OF MICHIGAN · ARTES · SCIENTIA · VERITAS · 1817

Technical Report No. 13

December 1968

THE HUMAN PERFORMANCE CENTER

DEPARTMENT OF PSYCHOLOGY

The Human Performance Center is a federation of research programs whose emphasis is on man as a processor of information. Topics under study include perception, attention, verbal learning and behavior, short- and long-term memory, choice and decision processes, and learning and performance in simple and complex skills. The integrating concept is the quantitative description, and theory, of man's performance capabilities and limitations and the ways in which these may be modified by learning, by instruction, and by task design.

The Center issues two series of reports. A Technical Report series includes original reports of experimental or theoretical studies, and integrative reviews of the scientific literature. A Memorandum Report series includes printed versions of papers presented orally at scientific or professional meetings or symposia, methodological notes and documentary materials, apparatus notes, and exploratory studies.

THE UNIVERSITY OF MICHIGAN

COLLEGE OF LITERATURE, SCIENCE AND THE ARTS

DEPARTMENT OF PSYCHOLOGY

THE ELUSIVE TRADEOFF: SPEED VERSUS ACCURACY IN CHOICE REACTION TASKS WITH CONTINUOUS COST FOR TIME

Richard Gavin Swensson

HUMAN PERFORMANCE CENTER--TECHNICAL REPORT NO. 13

December, 1968

This research was supported in part by the Advanced Research Projects Agency, Department of Defense, monitored by the Air Force Office of Scientific Research, under Contract No. AF(638)-1235 with the Human Performance Center, Department of Psychology, University of Michigan, and in part by the National Aeronautics and Space Administration under Grant NGR 23-005-171 to the Engineering Psychology Laboratory, University of Michigan, monitored by the Ames Research Center, NASA.

PRECEDING PAGE BLANK NOT FILMED

PREFACE

This report is an independent contribution to the program of research of the Human Performance Center, Department of Psychology and also that of the Engineering Psychology Laboratory. It was supported in part by the Advanced Research Projects Agency, Department of Defense, monitored by the Air Force Office of Scientific Research, under Contract No. AF(638)-1235 with the Human Performance Center, Department of Psychology, University of Michigan, and in part by the National Aeronautics and Space Administration under Grant NGR 23-005-171 to the Engineering Psychology Laboratory, University of Michigan, monitored by the Ames Research Center, NASA.

This report was also a dissertation submitted by the author in partial fulfillment of the degree of Doctor of Philosophy (Psychology) in the University of Michigan, 1969. The doctoral dissertation committee was: Professor Ward Edwards, Chairman, Associate Professor Richard W. Pew, Associate Professor Wilfried M. Kincaid, Professor J. E. Keith Smith, and Associate Professor Daniel J. Weintraub.

PRECEDING PAGE BLANK NOT FILMED

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

ABSTRACT

THE ELUSIVE TRADEOFF: SPEED VERSUS ACCURACY IN CHOICE REACTION TASKS WITH CONTINUOUS COST FOR TIME

by

Richard Gavin Swensson

Chairman: Ward Edwards

Human capability to trade accuracy for increased speed of performance underlies all predictions of recently developed random walk models for choice reaction time. Although this predicted tradeoff has been demonstrated in a variety of other choice experiments, Swensson and Edwards (1968) obtained absolutely no trade between speed and accuracy, in spite of every effort to experimentally represent the model's assumptions.

The five experiments reported here attempted to discover the conditions necessary for a speed-accuracy tradeoff. All used a PDP-1 computer to present one of the two equiprobable stimuli and the subject's feedback on a cathode ray tube display. The subjects' entire payoff depended on both their accuracy and response time on each trial. Cost for response time was proportional to that time. Session-to-session changes in the payoff difference between a correct response and an error varied the relative attractiveness of speed and accuracy.

Experiments I, II, and III, using two, five, and seven subjects, varied payoffs over a small range in an attempt to obtain fast responses with intermediate (e.g., 25%) error rates. Tradeoffs clearly did not occur for eleven of the twelve subjects. They used only the same two response strategies found by Swensson and Edwards (1968). Either subjects responded accurately, with 5% or fewer errors, or they ignored the identity of the stimulus and preprogrammed their responses, accepting chance error rates (50%) to decrease their reaction time. With two types of easily discriminable stimuli in Experiments I and II, subjects could decrease their mean time by 15-20 or 45-70 msec. by preprogramming. The difficult stimulus discriminations in Experiment III required longer times for accurate responses, and increased the difference between strategies to 135-210 msec.

Experiments IV and V used two and three different experienced subjects from Experiment III, varying the relative emphasis on speed versus accuracy over a very large range of payoffs. In

Experiment IV one subject's accuracy performance did not change over conditions, but the other displayed a very orderly tradeoff between speed and accuracy of performance. Over 56 conditions on a new, extremely difficult discrimination task, he again traded between speed and accuracy, but accuracy increased more slowly with increases in his reaction time. For both tasks, chance level accuracy would have been reached with discrimination times 90-120 msec. slower than his mean preprogrammed strategy time.

Experiment V modified the linear cost function to charge subjects only for time exceeding a 250-msec. free delay interval. This should have removed the fixed cost for attempting to discriminate the stimuli, resulting from the apparent need for a substantial delay before accuracy could be improved over chance level. All three subjects produced clear tradeoffs over conditions, although they had all failed to do so in Experiment III.

Both optional stopping and fixed sample size versions of the random walk model produced excellent but almost equivalent fits to all five empirical speed-accuracy tradeoff functions in Experiments IV and V. Estimated rates at which the evidence needed for discrimination accumulated over time were slower for the most difficult stimuli, but chance accuracy intercepts seemed unaffected. These intercepts were 66-94 and 92-120 msec. longer than subjects' preprogrammed times for the fixed sample and optional stopping versions of the model.

Reaction times for errors were faster than for correct responses, both in response to easily discriminable stimuli and for the faster, less accurate performance found with difficult stimulus discriminations. Mean error times were slower, however, for highest degrees of accuracy with the difficult stimuli, considerably slower for the most difficult stimulus pairs. To describe these differences between correct response and error times, the random walk models must provide for: 1) unstable sampling or stopping criteria and 2) redundant stimulus observations with difficult discriminations or some maximum effective observation interval.

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Research on the Speed-Accuracy Tradeoff

In a wide variety of tasks maximum speed and maximum accuracy are incompatible. Often there appears to be a tradeoff relationship between speed and accuracy of performance. More specifically, at any given level of practice in tasks that place a premium on both fast and accurate performance, speed can be increased, but only at the cost of more frequent errors, while a reduction in error rate requires slower, more careful performance. One of the earliest statements and demonstrations of a speed-accuracy tradeoff was by Garrett (1922), who hypothesized that "judgment or perception [should] grow in accuracy with the increase in time taken to make it" (p. 5). Using a metronome, he attempted to control the speed with which his subjects (Ss) made various perceptual discriminations and found decreases in accuracy with faster enforced rates of judgment.

Choice reaction time (CRT) tasks, since they emphasize both speed and accuracy of performance while requiring S to make only a single discrete response on each trial, provide strong tests of the generality of the speed-accuracy tradeoff hypothesis. Much of the evidence for a speed-accuracy tradeoff in CRT tasks comes from experiments with more than two stimuli and responses. Generally, these demonstrate: 1) the conditions which produce increased speed also result in higher error rates and 2) manipulations to obtain greater accuracy slow down reaction time (RT).

instructions, which usually define for Ss the relative costs of time and errors, can induce faster performance (with more errors) or reduced error rates (with slower mean RT).

These results have obtained, both between groups and within individual Ss, over a variety of task conditions (Hick, 1952; Howell & Kreidler, 1963, 1964; Leonard & Newman, 1965). The use of explicit payoff matrices, to provide both instructions and incentives for speed or accuracy, can produce substantial changes in tradeoff performance that are highly consistent across Ss (Fitts, 1966; Pachella & Pew, 1968).

The fast reaction times and very low error rates found in most two-alternative CRT tasks with highly discriminable stimuli make any demonstration of a tradeoff between speed and accuracy difficult. By using a forced reaction time procedure, however, Schouten and Bekker (1967) demonstrated a strong and systematic relationship between faster forced reaction times and higher error rates. They used an auditory signal to force Ss' response on cue and presented the visual stimulus that told S what response was appropriate at varying intervals of time before Ss had to respond. Error rates decreased continuously with increases in time between the stimulus presentation and Ss' response.

Models for the speed-accuracy tradeoff.--The recently developed statistical decision or random walk models for CRT performance (Stone, 1960; Edwards, 1965; Fitts, 1966) directly incorporate both time and error costs, whose manipulation have produced most demonstrations of the speed-accuracy tradeoff.

Random walk models treat the CRT experiment as a decision-making situation, in which S has costs associated with both errors and the duration of his decision (response) time. According to these models, the speed and accuracy of S's response depends 1) on the sensory evidence provided by sampled observations of the stimulus input, and 2) on the payoff matrix and sampling cost (either provided to S explicitly or assumed by him). The hypothesized observations are assumed to be short and equal in duration so that a fixed sampling cost per observation corresponds to a continuous linear cost for observing time.

Random walk models assume that S begins each trial with some prior odds that a particular response will be correct on that trial. With the evidence from each sampled observation S modifies his odds about the correct response according to Bayes's theorem. His posterior odds describe a random walk over time as observations accumulate. Unless the stimuli are completely indiscriminable, the expectation of this random walk will be biased and S's odds in favor of the correct response will tend to increase as observation time increases.

The optional stopping version of the random walk models assumes that S has posterior odds cutoffs. Eventually S's posterior odds in favor of one hypothesis about the correct identity of the stimulus become large enough to reach his cutoff criterion; he then stops observing and makes the response appropriate to that stimulus. The values of these cutoff odds are pre-set by S and directly determine his expected error rates. With S's prior odds for a given physical stimulus situation, these cutoffs also determine the expected number of observations and, therefore, S's mean reaction time. A more extreme odds cutoff will decrease the error rate for that response, but will increase the mean reaction time preceding it. This occurs because the more extreme the cutoff odds, the greater are the expected number of observations before S's posterior odds in favor of that response reach the cutoff. More lax criteria will have the opposite effect, increasing errors but producing faster reaction times. For each individual S in a given physical stimulus situation, the value of the odds cutoffs depends upon the loss in payoff for making an error (difference between reward for a correct response and cost for an error) relative to the sampling cost (cost per second of observing time).

Another version of the random walk model is plausible; perhaps human Ss cannot perform the continuous monitoring of their current odds for the optional stopping procedure. A fixed sample size version of the random walk model assumes that sequential sampling is impossible, that Ss must pre-select some fixed number of observations (fixed observation duration) and respond based on whatever evidence it provides. The optimal sample size, for the fixed sample size version, depends on exactly the same factors as the optimal cutoff odds for the optional stopping version, plus S's prior odds. Larger costs for errors--relative to those for observing time--result in larger sized samples (longer observation intervals), thus increasing reaction time while decreasing the error rate. Even though this version of the random walk model predicts a fixed sampling time for all trials under constant prior odds and payoff conditions, a distribution of reaction times will result from fluctuations in other delays associated with the response (e.g., motor and transmission times).

Fundamental to both versions of the random walk model is the prediction of a continuous tradeoff between response speed and accuracy. Prediction of this tradeoff is inherent in their representation of sensory systems--as sequential processors of independent stimulus observations--and in their assumption that human Ss can and will use prior odds, costs and payoffs to control the termination of the sampling process. The two versions differ only in how they assume this control can be exercised. Under the fixed sample size version, S directly controls his sampling time, which indirectly determines his expected error rate; under the optional stopping version, S directly controls his expected error rate, which in turn determines his expected sampling time. Both versions of the random walk model are optimal procedures, given their different assumptions. In general, optional stopping is more efficient than fixed sampling. If sample size is optional, a decision maker can take advantage of unusually definitive early observations to terminate observing sooner.

Ollman (1966) and Yellott (1967) have proposed an entirely different kind of model which also predicts the association of faster mean RTs with higher error rates. The Ollman-Yellott guessing model assumes that S has two possible strategies available on every trial. He can respond based on some internal test of the stimulus identity (with a small expected error rate) or he can elect to guess, with chance accuracy but a faster mean RT. When the stimuli are equiprobable, variation in the proportion of guessing trials from 1.0 to 0 will produce a tradeoff function in which error rates decrease from .50 to that inherent in the test procedure, while mean RTs increase. Although not formally included in this model, both Ollman's (1966) and Yellott's (1967) discussions clearly indicate that the proportion of guessing responses should increase with decreases in the (possibly implicit) cost of an error relative to that of response time.

Predicted tradeoff functions.--Each of the three models predicts a form for the speed-accuracy tradeoff function. Although each could be generalized, the actual functions presented below are limited to two-choice tasks with equiprobable stimuli and symmetric payoff matrices (loss in payoff for an error equal for both stimuli). Each model predicts a linear relationship between mean RT and a different function of Ω, over changes in the relative emphasis on speed versus accuracy. Ω is defined as the odds favoring a correct response (the ratio of the probability of a correct response to the probability of an error). By assumption, a subject is always at least as likely to make a correct response as an error; that is, $\Omega \geq 1.0$.

For the optional stopping version of the random walk model, this function can be adapted from Edwards (1965; eq. 19) as:

$$\left(\frac{\Omega - 1}{\Omega + 1}\right) \ln \Omega = a + b \text{ RT}. \qquad (1)$$

For the <u>fixed sample size</u> version of the random walk model, the function relating mean reaction time to Ω is a trivial consequence of Bayes's theorem and the assumed independence between successive observations. If the two stimuli are equiprobable, the prior odds in favor of a correct response are 1.0, and Ω_N, the posterior odds after a sample of N observations of fixed duration t, is equal to $\Pi_{i=1}^{N} L_i$, where L_i is the likelihood ratio of the ith observation (the ratio of the probability of obtaining that observation if one response is correct, to its probability if the other response is correct). Since each observation is an independent sample from the same distribution, taking logs and expected values reveals that the expected value of ln Ω is a linear function of Nt. So that:

$$\ln \Omega = a + b \text{ RT} \qquad (2)$$

Both versions assume that observing time is equal to the number of observations multiplied by the duration of an observation. Neither assumes any time taken up by calculation or for comparisons between posterior odds or number of observations and the stopping criterion. In the absence of such specific assumptions, the two versions differ only in the type of stopping rule they employ and the constants a and b should not differ between (1) and (2). The best fit estimates of these constants to actual data, of course, will generally differ between the two versions. The constant a represents the expected value of the residual delays associated with the reaction; b represents the average path of the random walk over time--the rate at which evidence accumulates. For the optional stopping model, the expression $\left(\frac{\Omega - 1}{\Omega + 1}\right)$ is always less than 1.0, and indicates the factor by which the optional stopping procedure permits faster response times than the fixed stopping procedure for identical expected error rates. This increased efficiency is greatest when Ω is near 1.0 (error rates are high). For both models, the optimal location on their tradeoff functions, i.e., the performance which will maximize expected payoff (minimize expected cost), depends on the cost of an error relative to that of time. (These statements would have to be much more complicated in less symmetric cases.)

From either the Ollman (1966) or the Yellott (1967) statements of the guessing model, it is possible to obtain the following equation, relating mean reaction time to Ω:

$$RT = g + d\left(\frac{\Omega - 1}{\Omega + 1}\right) \tag{3}$$

The constant g is equal to the expected reaction time for guesses, d is the expected reaction time for responses based on discriminations of the stimulus identity, and $\left(\frac{\Omega - 1}{\Omega + 1}\right)$ can be shown to be equal to the difference between the proportion of correct responses and the proportion of errors. The performances predicted by Equation (3) are always some weighted average of the guessing and discrimination performance. All are suboptimal; S's expected payoff for any such mixed strategy must necessarily fall in between those for the pure guessing and discrimination strategies. The tradeoff functions of Equations (1) and (2) map out sets of optimal performances; each point represents performance which maximizes expected payoff for some set of conditions. Only the end points of Equation (3) represent optimal performances; either guessing or discriminating on all trials will always maximize expected payoff for an S who has only these two strategies available to him.

Empirical tradeoff functions.--Pew (1968) replotted earlier data from Pachella and Pew (1968), Schouten and Bekker (1967), and some unpublished data on individual Ss collected by Professor P. M. Fitts. He found that plots of mean RT against ln Ω, which he called "Speed-Accuracy Operating Characteristic" functions, were very well fit by straight lines in all cases. Additional amounts of practice increased the slopes and decreased the intercepts of these functions, but maintained their linearity--even though the Ss expressed their improvement in faster mean RTs or lower error rates or both. Although strictly linear functions between mean RT and ln Ω are those predicted by Equation (2), and thus Pew's (1968) re-analysis supports the fixed sample size version of the random walk model, the evidence is far from decisive. Pew's (1968) interest was only in describing the empirical relationships, and he did not attempt to fit alternative models. The Schouten and Bekker (1967) forced reaction time procedure might well have artifically constrained Ss to fix the duration of their observation intervals, in order to respond on cue. Pachella and Pew's (1968) dichotomous criterion time procedure, which categorized Ss' responses as fast or slow to provide both payoffs and feedback, might also have forced Ss to fix their sample size to keep their responses fast.

Data also exist that support the Ollman-Yellott guessing model. With variations in speed forced by different criterion payoff times, the model produced good fits to the performance of three Ss in a two-choice auditory frequency discrimination task (Ollman, 1966) and of three Ss in a visual task with two differently colored stimuli (Yellott, 1967). But neither experimenter investigated the fits of alternative models.

The phantom tradeoff.--Swensson and Edwards (1968) attempted to test some of the more controversial predictions of Edwards's (1965) statement of the optional stopping random walk model for two-choice tasks. The procedure used explicit payoffs. Subjects received feedback after every trial about their gross payoff for that stimulus-response combination, their cost for time, and their net payoff, which (summed over trials) completely determined their income from the experiment. Instead of trading off between speed and accuracy over conditions, the three Ss used only two distinct response strategies throughout. Either Ss responded accurately, with error rates of .005-.05, or they preprogrammed a single response on all trials. Since the preprogrammed strategy ignored the identity of the stimulus, it necessarily implied error rates of .50 for equiprobable stimuli, but allowed Ss to decrease their mean RT by 15-20 msec.

The accuracy and preprogrammed response strategies found by Swensson and Edwards (1968) closely resemble those assumed by the Ollman-Yellott model, but there was no sign of even the type of tradeoff predicted by Equation (3). Under most conditions Ss used the same strategy for all trials. In the remaining conditions *Ss* clearly switched between these two strategies, usually for long blocks of trials, apparently uncertain about which response strategy would yield the maximum payoff. These results sharply contradict those of Schouten and Bekker (1967), described earlier, who used similar types of stimuli. This contradiction is even more puzzling, since Swensson and Edwards (1968) took great care to represent the random walk model's assumptions in their procedure. Another marked discrepancy is the much longer tradeoff interval of 180 msec. for Schouten and Bekker's *Ss*. Error rates decreased from .50 to .03 as their stimulus-response interval lengthened from 170 to 350 msec., in contrast to the 15-20 msec. difference in RT found by the Swensson and Edwards study between response strategies with comparable error rates.

Under what conditions can and will *Ss* produce a speed-accuracy tradeoff in two-alternative CRT tasks? Such a tradeoff is implicit in all predictions of both versions of the random walk model. Its absence would render any version of that model useless in CRT tasks.

EXPERIMENTS I AND II

Most of the conditions in the Swensson and Edwards (1968) experiment contained either asymmetric payoff matrices or unequal stimulus relative frequencies or both. They were not specifically designed to discover whether a speed-accuracy tradeoff could in fact be obtained for some payoff conditions. Experiments I and II of this thesis attempted to converge on payoff conditions for which individual Ss would either: 1) switch between preprogrammed and accurate response strategies, or 2) produce intermediate tradeoff performance (e.g. with error rates of .25), as predicted by the random walk models, without resorting to a preprogrammed strategy. Experiment I used apparatus and spatially separated stimuli identical to those used by Swensson and Edwards (1968). Experiment II altered the stimuli and procedure to overcome difficulties that arose in Experiment I.

Experimental pressure for speed versus accuracy.--For symmetric payoff matrices, equiprobable stimuli and a fixed linear time-cost function, the demand for fast or accurate performance depends upon D, the difference in payoff between a correct response and an error. When D is close to zero, correct responses are only slightly more valuable than errors, and S's net payoff depends primarily on his RT. As D increases, the accuracy of S's response becomes increasingly important to his net payoff. If S has only two strategies available, preprogrammed responding and accuracy, the preprogrammed response strategy will have a higher expected payoff for small values of D, while that of the accuracy strategy will be higher when D is large. To maximize his expected payoff, S should switch strategies at some intermediate value of D. The speed-accuracy tradeoff hypothesis predicts continuous variation in mean RT and error rate over intermediate values of D.

Method: Experiment 1

Apparatus and stimulus display.--A PDP-1 computer controlled the stimulus sequence and presented the stimuli on a cathode ray tube, similar in appearance to a television screen (Fig, 1a), mounted at about eye level for a seated S. Subject faced the display and responded by lifting his left or right index finger from one of two micro-switch response keys, mounted on a board before him, and inclined upward toward the display at about a 30° angle.

As Fig. 1a shows, the upper third of the display contained complete information about the payoff matrix, the stimulus relative frequencies, and the cost to S for one second of response time. Note that S receives a payoff even if he makes an error. Otherwise, the cost for time would make small D values too costly for Ss. The stimulus was a small bright square in the center of either the left or right portion of the frame near the bottom of the display. Correct responses (left or right) were compatable with the stimuli. Figures 1a, 1b, and 1c illustrate the feedback sequence, which appeared in the center of the display. Immediately following S's response on each trial, the gross payoff vector appeared (Fig. 1a). After about a one-second delay, the cost vector (Fig. 1b) appeared. The cost for time was equal to the S's reaction time multiplied by the cost

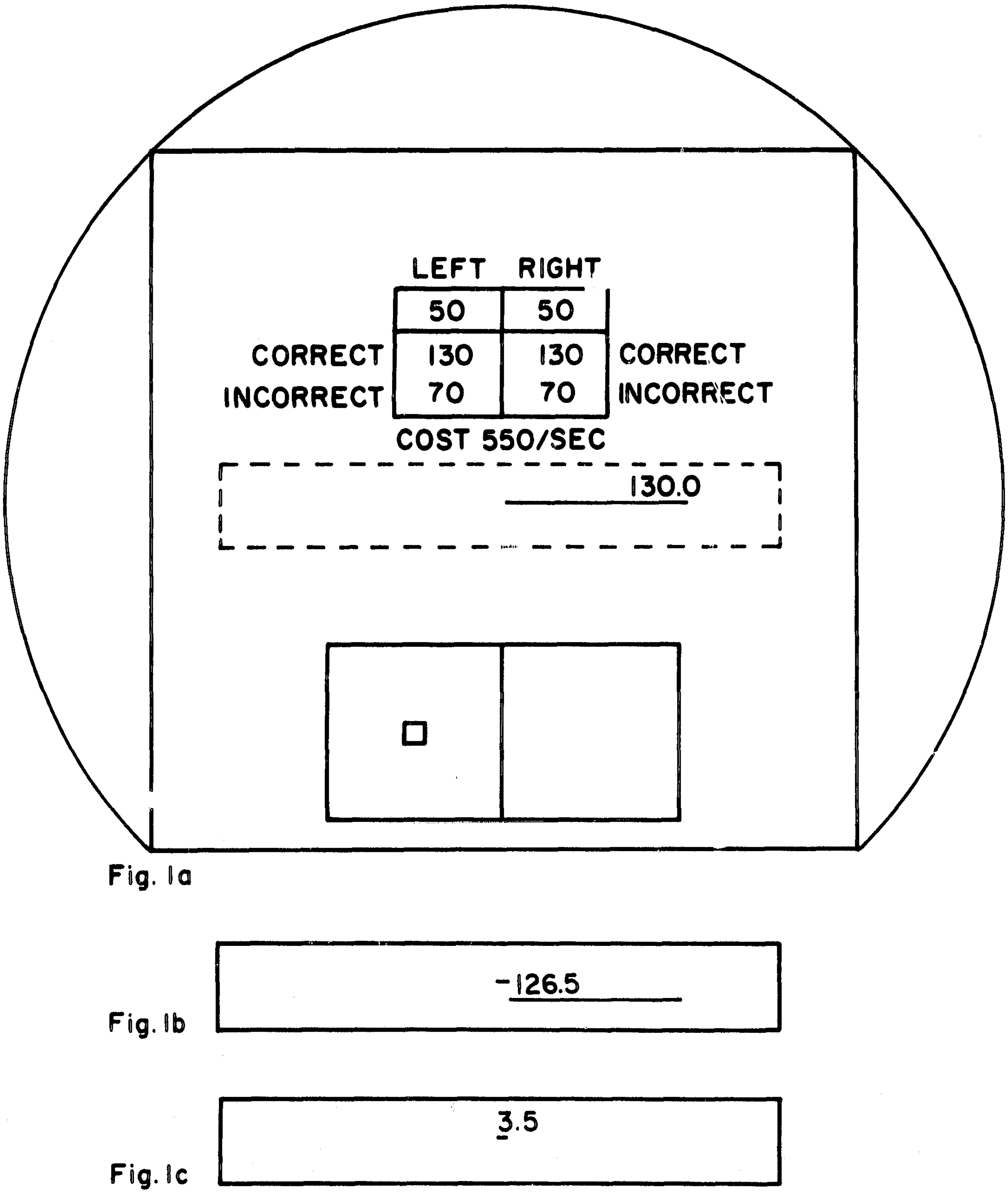

Fig. 1. The stimulus display for Experiment I. Fig. 1a illustrates the stimulus and feedback display as it appeared in Experiment I, immediately after a correct left finger response; the dotted area contained the feedback display. The example illustrates a gross payoff of 130 points for a correct response (Fig. 1a), a cost of 126.5 for an RT of 230 msec. (Fig. 1b), and the resulting net payoff of 3.5 points for that trial (Fig. 1c). All information appeared bright on a dark green ground inside the 10 x 10 inch display area.

of 550 points per second. The vector scale was adjustable; Fig. 1 illustrates the scale which could display payoff and cost vectors between -300 and +300. For long response times, the cost vector was truncated, but the associated number always displayed _S_'s actual cost. To obtain maximum clarity and effect, the time cost part of _S_'s gross payoff vector seemed to "break off" and drift downward for about two seconds. Fig. 1c shows the final stage of feedback, in which only _S_'s net number of points won or lost remained. The rate of the feedback presentation was such that no _S_ had difficulty in reading the exact numbers if he cared to.

Procedure.--The moment of onset of the stimulus was uncertain; it could appear at any time between 1 and 3 seconds after _S_ initiated a trial by pressing down both response keys. The distribution of fore-period intervals was rectangular between these limits. Subject responded by releasing one of the keys after the stimulus appeared. If _S_ responded before the stimulus appeared, the computer suppressed the stimulus and displayed the words, "Please wait for the stimulus." Subject then had to begin the trial again. If _S_ released the second key within 100 msec. (.1 sec.) of the first, his response was considered an error, but he was charged only for the time until his first response. After each double-response error and before the feedback display, the computer displayed the words, "one response per trial will suffice."

Points were worth 1/4 cent each. The cost for time was always 550 points ($1.38) per second. A charge of 500 points

($1.25) for each anticipation response, together with the random foreperiod, effectively discouraged Ss from initiating their preprogrammed responses before the stimulus appeared.

The stimulus sequences had run length distributions that conformed as closely as possible to the expected distributions in a sequence of 450 trials from an independent Bernoulli process with equal stimulus proportions. Once the appropriate number of runs of each length was set, different sequences were constructed by sampling alternately, without replacement, from the sets of runs of the left and right stimulus events until all runs had been chosen. This procedure insured that each individual sequence was representative and contained no fortuitous sequential dependencies. The stimulus sequences were generated prior to the experimental sessions and stored on punched paper tape. Each sequence was used for only a single session.

The Ss were two paid volunteer men who answered an advertisement in the University of Michigan daily newspaper. A third S began the experiment at the same time, but left after the second week of testing.

The experiment had three phases. Phase 1 consisted of four training sessions, during which Ss worked with two different payoff conditions on alternate blocks of 50 or 100 trials. One condition had a large value of D, demanding accurate performance to maximize payoffs; the other condition had a very small value of D and ... demanded speed. During these initial training sessions Ss were paid at a fixed rate of $1.50 per hour. The purpose of these sessions was to permit Ss to develop response strategies appropriate to each of the extreme values of D before the real money sessions began. Phase 2 consisted of eight sessions with less extreme values of D to allow performance to stabilize. In these and later sessions, Ss' pay depended entirely on the total number of points they won. Each set of four sessions in Phase 2 combined high and low values of D with both favorable and unfavorable levels of payoffs, as Table 1 shows. Each individual S's training performance determined the actual payoff matrix entries for the first four conditions; his performance on these determined the payoffs for the second four.

During Phase 3 D varied systematically between conditions, to induce either switching between the accuracy and preprogrammed strategies or performance with intermediate levels of errors. For each S, D for the initial condition was the value having equal expected payoffs for both the accuracy and preprogrammed strategies, based on his most recent performance. D varied across conditions according to the following procedure. If S responded accurately, D decreased in the next condition; if S preprogrammed his response, D increased. If S switched back and forth between the two strategies or used an intermediate strategy, the same value of D was repeated to a limit of three sessions and then changed. The original intention was to use a fixed set of rules for gradually decreasing the size of the change in D. These had to be abandoned for a more informal procedure because of gradual, but continuous, improvements in Ss' performance. Given the value of D used, the actual numerical

Table 1

PAYOFFS FOR CORRECT RESPONSES, ERRORS, AND THEIR DIFFERENCE (D)

USED FOR BOTH SUBJECTS IN EACH CONDITION OF EXPERIMENT I

Condition	Payoffs in Points					
	Subject LP			Subject RZ		
	Correct	Error	D	Correct	Error	D
1	122	63	59	140	105	35
2	129	102	27	118	114	4
3	114	87	27	132	128	4
4	137	78	59	126	91	35
5	128	108	20	144	113	31
6	136	84	52	123	120	3
7	122	70	52	137	134	3
8	114	94	20	130	99	31
9	129	93	36	134	118	16
10	129	93	36	131	121	10
11	129	93	36	131	121	10
12	135	93	42	128	124	4
13	131	93	38	128	124	4
14	127	93	34	125	121	4
15	127	93	34	122	118	3
16	127	93	34	121	119	2
17	123	93	30	119	117	2
18	125	93	32	119	117	2
19	124	93	31	118	117	1
20	122	93	29	118	117	1
21	120	93	27	116	115	2
22	123	93	30	117	114	3
23	123	102	21	118	113	5
24	123	106	17	119	112	7
25	123	110	13	119	112	7
26	133	118	15	121	110	11
27				132	117	15
28				134	115	19
29				137	112	25

entries in the payoff matrix were based upon S's recent performance so that, without improvement, use of the strategy having the higher expected payoff would result in about $3.00 for each two-hour session. Table 1 contains values of D and the payoff matrices for each S over all conditions.

Method: Experiment II

Apparatus and stimulus display.--Figure 2a illustrates the display and stimuli for Experiment II. Other features were identical to those in Experiment I. When S initiated a trial, the computer displayed a small cross to mark a fixation point, over which the stimulus appeared following the 1-3 second random delay. The stimulus was one of two small rectangles with its longer axis slanted either northwest-southeast or northeast-southwest. As seen in Fig. 2a, the two stimuli were easily discriminable; their length-width ratio was 30:1, so that the stimulus appeared as one of the diagonal cross-bars of an X, superimposed over the small fixation point. Retracting the left index finger was the correct response to the rectangle whose top was oriented 45° to the left of vertical, and retracting the right index finger was correct when the upper end of the rectangle was inclined 45° to the right of vertical. Since the phosphor of the cathode ray tube faded fairly slowly, traces of the previous stimulus were removed prior to the next trial by illuminating all points in the square area around the stimuli seen in Fig. 2a. This insured that the stimulus for the next trial would appear on a faintly illuminated, homo-

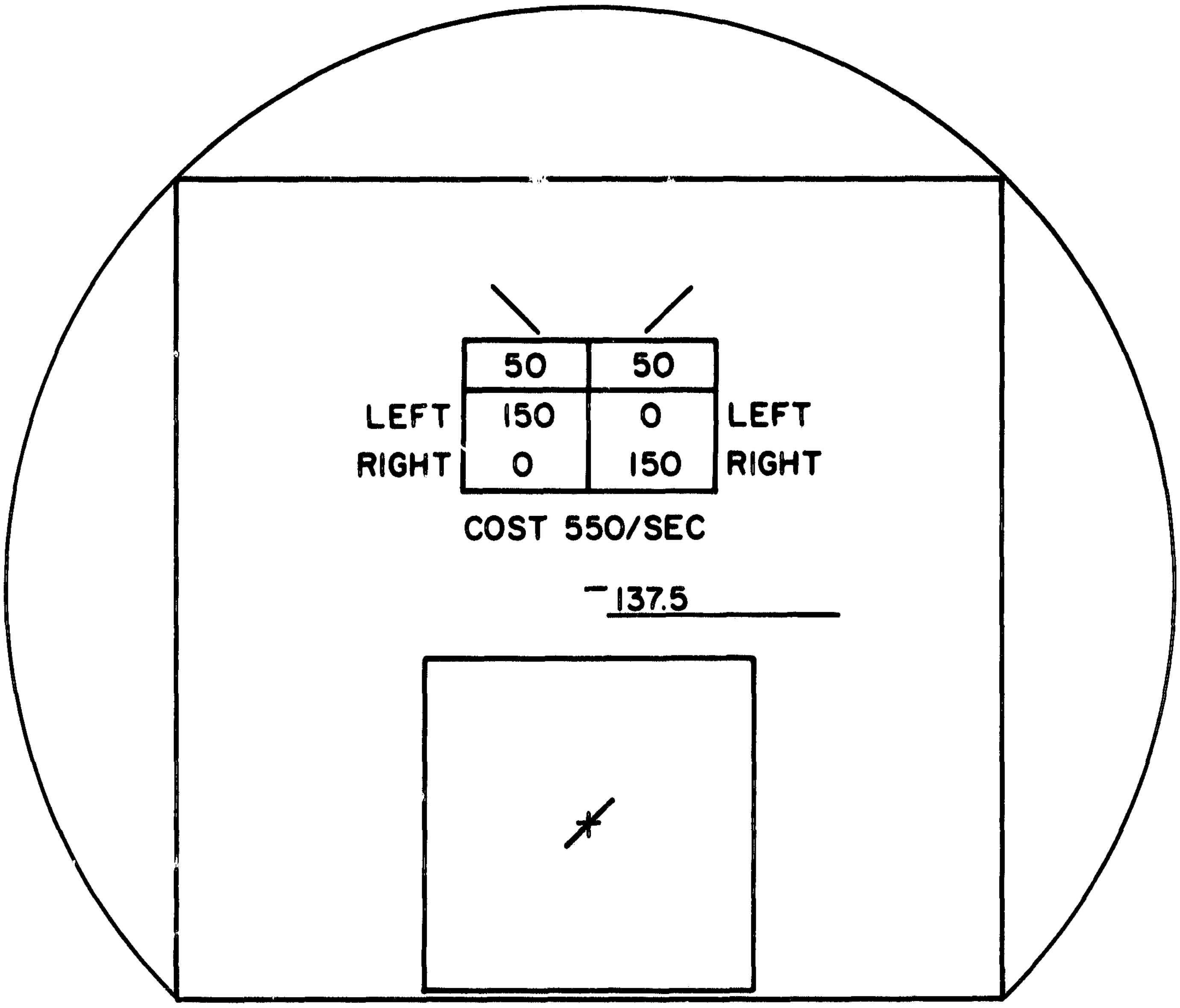

Fig. 2a

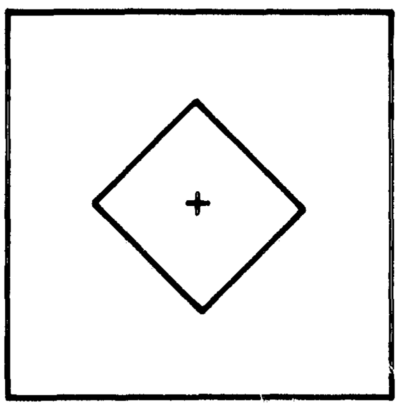

Fig. 2b

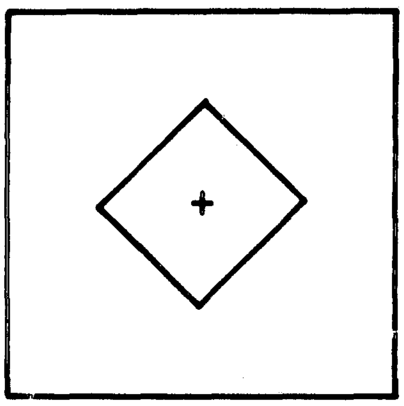

Fig. 2c

Fig. 2. The stimulus display for Experiments II-V. Fig. 2a illustrates the stimulus and display as it appeared in Experiment II, midway through the feedback sequence after a correct right-finger response with an RT of 250 msec. (cost = 137.5). Figures 2b and 2c illustrate the easiest pair of stimuli used in Experiment III (length-width ratio = 55:53). Figure 2b required a left-finger response and Fig. 2c required a response with the right finger.

geneous field. The location of the fixation point also varied slightly along the horizontal axis from trial to trial, as an additional precaution against effects due to traces of the previous stimulus.

Procedure.--The Ss included the two experienced Ss from Experiment I and three naive Ss, selected in the same manner. The experienced Ss received two training sessions with the new stimuli and procedure and then began the test sessions. The three new Ss received four training sessions divided into blocks of 50 or 100 trials and received information about total points earned after each block. Two of the Ss alternated between payoff matrices with large and small values of D (150 and 10), which emphasized accurate or fast responding respectively. The third S, JF, received training on only the payoff matrix that demanded accuracy.

After this initial training, Ss were paid only for the number of points they earned. Over the first 10 sessions, D decreased systematically until S used the preprogrammed strategy in two successive conditions, and then increased. The E emphasized the fact that, as D decreased, speed became more important to their net payoff. Since JF received training under accuracy conditions only, he had had no previous opportunity to discover the preprogrammed strategy. These first 10 sessions provided E with an estimate of JF's preprogrammed response time and also provided additional opportunity for the accuracy performance of all Ss to stabilize.

During the test sessions D varied according to S's performance on the previous condition, much as it did in Experiment I. However, the amount by which D changed depended upon the amount of change in S's performance. Following a large change in S's performance, each new value of D yielded approximately equal expected payoffs for both the accuracy and preprogrammed strategies. Actual values of D used for each S appear in Table 2.

In all conditions of Experiment II, D was displayed directly in the payoff matrix. The payoff for an error was zero points and S's gross payoff was equal to D each time he made a correct response. The number of points S lost for his response time was, of course, subtracted from his gross winnings on each trial. As D varied, so did the expected payoff of both response strategies and the S's total net points for that session.

To compensate for the change in net payoff with D, S received a stake at the beginning of each session of a specified number of trials. The amount was based on his most recent performance and biased in his favor. At the end of the session S's total net points was added to or subtracted from this stake. If S equaled his previous performance on either the accuracy or preprogrammed strategy (whichever had the higher expected value for that condition), he earned approximately $1.50 per hour. Improved performance, fewer errors or faster mean RT, increased S's earnings for that session but entered into the calculation of his stake for the next session. A poor performance resulted in reduced winnings or a loss for that session, but produced a more favorable stake

Table 2

DIFFERENCE IN PAYOFF BETWEEN CORRECT RESPONSES AND ERRORS (D)

USED FOR EACH SUBJECT AND CONDITION IN EXPERIMENT II

Condition	Subject				
	LP	RZ	JS	WW	JF
1	100	130	150	150	150
2	90	120	125	125	100
3	80	110	100	100	50
4	70	40	75	75	25
5	60	60	50	50	25
6	65	80	25	25	50
7	70	100	25	25	100
8	75	100	50	50	150
9	65	100	75	75	200
10	55	40	100	100	250
11	58	100	60	40	200
12	63	160	55	60	150
13		100	50	80	100
14		160	50	70	50
15			65	70	80
16			60	65	100
17			55	70	100
18			52	75	90
19			65	75	95
20			65	70	92
21			63	70	95
22			60	70	95
23				65	90
24				65	92
25				57	95
26				63	92
27				58	92
28				54	

for the next session. Thus, while it was advantageous for S to maintain and, if possible, to improve his performance, losses occasioned by a poor performance could be overcome. The basis for calculating the stake was thoroughly explained to each S and rigidly adhered to by E. Although the variance in S's payoffs for individual sessions was large and all lost on some sessions, no S averaged less than $1.25 per hour over the entire experiment.

Sessions lasted 1 1/2 to 2 hours and included 300-450 trials, depending upon the S. All other procedural details were identical to those in Experiment I.

Results: Experiments I and II

Response strategies.--The most important question in both experiments was whether a speed-accuracy tradeoff could be obtained for some values of D. For four of the five Ss in both experiments the answer was unambiguously negative; each appeared to use either the preprogrammed or the accuracy strategy as had all Ss in the previous experiment (Swensson & Edwards, 1968). The data for the remaining S, RZ, were ambiguous for both experiments.

Figures 3a, 3b, 4a, 4b, and 4c present RT and error data for each of the four typical Ss in Experiments I and II. Open figures are for sessions in which S used either the accuracy (circles) or the preprogrammed strategy (triangles) throughout the session. Filled figures are for sessions in which S switched strategy at least once. In such a mixed session, two error rates and two RTs per session are plotted, one for each strategy. Which

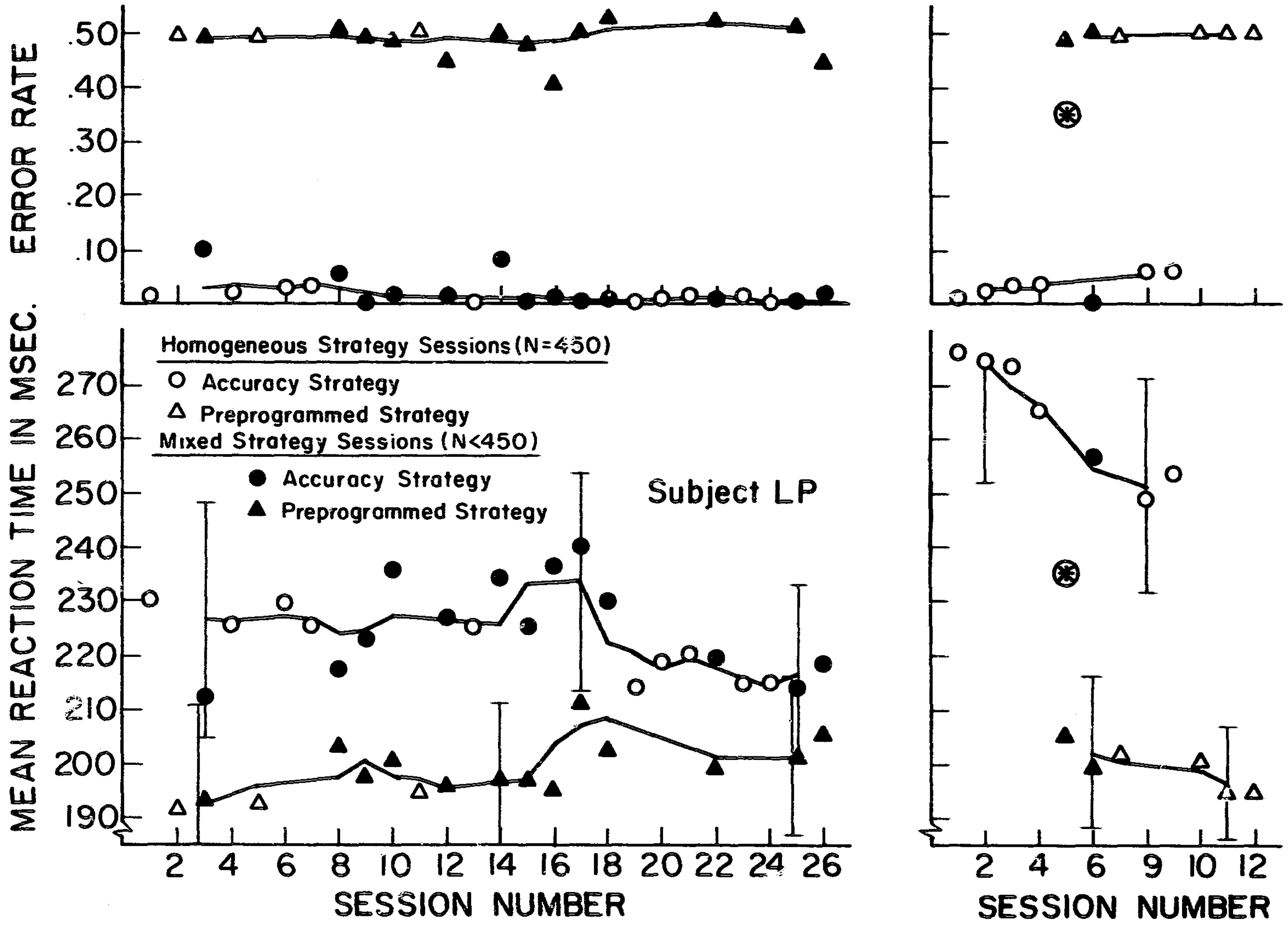

Fig.3a Experiment I

Fig.3b Experiment II

Fig. 3. Subject LP's mean RT (lower panels) and error rates (upper panels) for performance under both the accuracy (circles) and preprogrammed strategies (triangles) in Experiments I and II. Open points represent conditions for which LP used the same strategy on all 450 trials; filled points (one for each strategy per condition) represent mixed sessions where he switched strategies at least once. Solid lines indicate moving averages over adjacent sessions and vertical bars indicate variability. The performance indicated by a circled asterisk did not separate easily into the two strategies (see text).

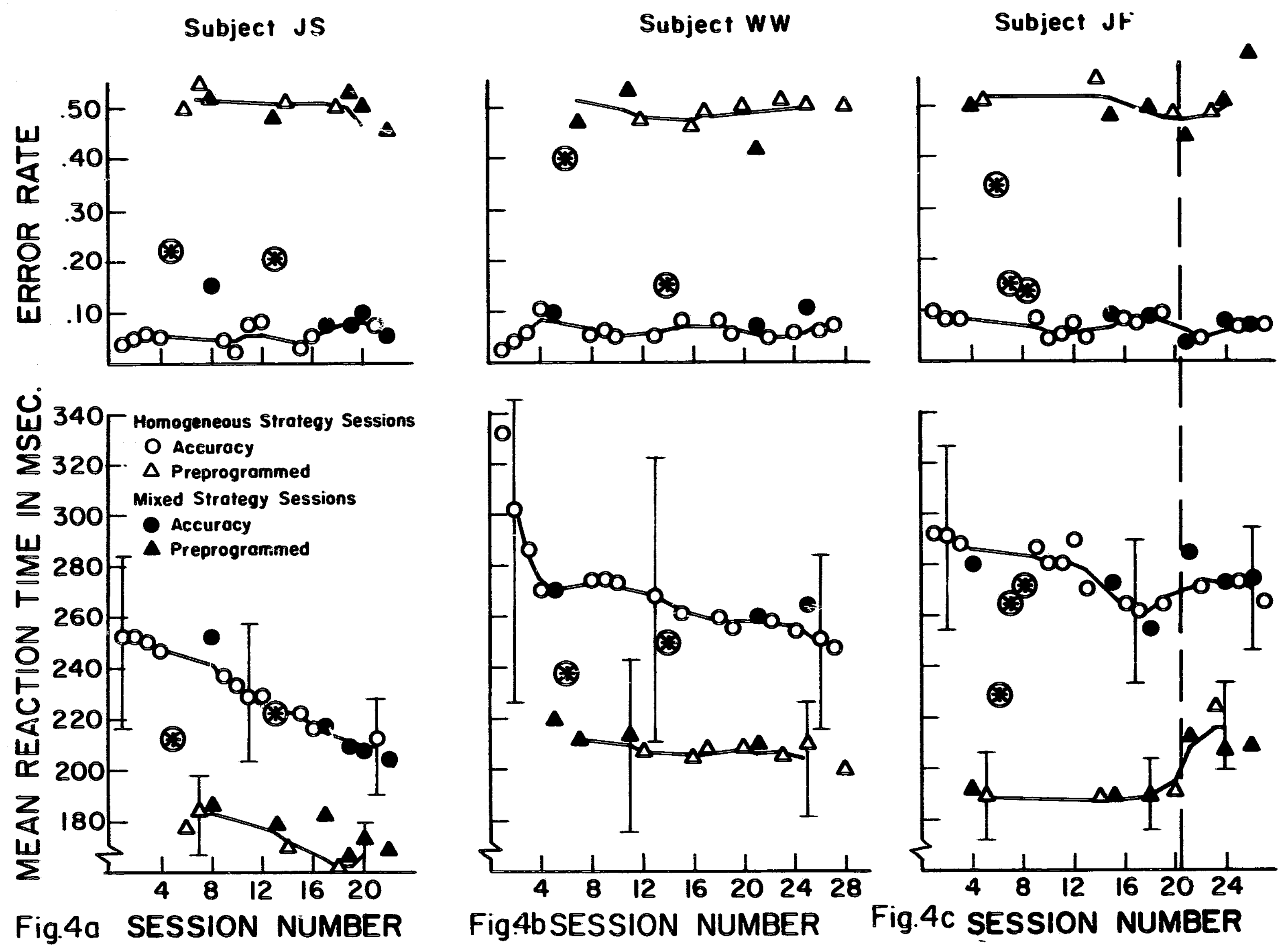

Fig. 4. Mean RT (lower panels) and error rates (upper panels) for each of the three Ss' accuracy (circles) and preprogrammed strategy (triangles) performance in Experiment II. Open points represent conditions for which Ss used the same strategy throughout; filled points (one for each strategy per condition) represent mixed sessions where S switched strategies at least once. Solid lines indicate moving averages over adjacent sessions and vertical bars indicate RT variability. Performance indicated by circled asterisks did not separate easily into the two strategies (see text).

strategy had been used on a given trial was usually obvious; the reason why is discussed below. Solid lines indicate the running average RT and error rate for each strategy over adjacent sets of three conditions. Vertical bars periodically indicate the RT variability. Error rates and RTs for the five entire sessions and the part of another in which Ss' strategies were not obvious, are indicated by circles with asterisks inside them. They were not included in the running averages.

The figures show that in many conditions, these four Ss used either an accuracy or a preprogrammed strategy on all trials. This occurred even during the later conditions, when only small changes in D reversed the difference in expected payoff between the two strategies. When preprogramming, Ss adopted highly systematic response patterns, such as using the same response on all trials or always using the response that had been correct on the previous trial. As in the earlier experiment (Swensson & Edwards, 1968), these stereotyped patterns allowed E to dichotomize each S's mixed sessions into blocks of trials during which he used only the preprogrammed or accuracy strategy. Successive strategy switches within a session were usually separated by long blocks of trials. The mean block length, the average number of consecutive trials for which S used the same strategy before switching, ranged from 72 to 132 for the accuracy strategy and from 66 to 127 for the preprogrammed strategy for the four Ss. The Ss' fortunate tendency to persevere with the same strategy for a long sequence of trials made it easy to determine objectively exactly where the strategy switching occurred.

The points in Figs. 3a, 3b and 4a, 4b and 4c represent various numbers of trials. Most of the deviant points are variable simply because they contain relatively few observations; they have little effect on the running averages. Vertical bars indicate the variability of each S's RT distributions by including plus and minus one average standard deviation about the running mean RT. They represent the square root of the weighted mean variance over the three adjacent conditions and, of course, refer only to the RT variability for individual responses. Because of the large number of observations, standard errors of the mean for most points were less than one msec.

All five figures clearly show Ss' dichotomous performance. Average preprogrammed strategy error rates hovered around .50, while those for the accuracy strategy remained low and relatively constant. Mean RT under the accuracy strategy was consistently slower than mean preprogrammed RT, and decreased over sessions.

Figures 3b, 4a, 4b, and 4c show very similar data for all four Ss in Experiment II. Mean RTs for the accuracy strategy were slower and error rates were higher than those obtained by Swensson and Edwards (1968) with the spatially separated stimuli of Experiment I (and seen for LP in Fig. 3a). With the exception of the last seven sessions for subject JF (Fig. 4c) all three showed a consistent improvement in accuracy strategy RTs with little change in their error rates. Those seven sessions to the right of the broken vertical line in Fig. 4c followed a six-week delay in the experiment for subject JF, because of final examinations and a break between semesters. His slower RTs for both the accuracy and preprogrammed strategies after the long interruption show that he could not continue to perform well without practice.

The difference in mean RT between the accuracy and preprogrammed strategies for LP (Fig. 3b), JS (Fig. 4a) and WW (Fig. 4b) was 45-50 msec. by the end of Experiment II; for JF (Fig. 4c) this difference was about 70 msec. before the six-week interruption and about 60 msec. afterward. These differences in Experiment II were much larger than the 15-20 msec. previously obtained (Swensson & Edwards, 1968) with the spatially separated stimuli used in Experiment I.

Although most sessions for these four Ss could be segregated into blocks of trials on which S used either preprogrammed or accuracy strategies, some could not. The circled asterisks in Figs. 3b, 4a, 4b, and 4c indicate blocks of trials with fairly high error rates and fairly fast RT in which no clear preprogrammed strategy could be identified. The smallest number of such anomalous trials was a single block of 28 trials for LP (Fig. 3b); the largest was three entire sessions for JF (Fig. 4c). Almost all instances occurred in the early experimental sessions. Inspection of the RT distributions indicated that some, although perhaps not all, of these atypical data were caused by rapid switching between strategies in short blocks or on individual trials. This is clearest for subject JF, who produced the largest number of these unusual sessions. The proportion of correct responses in his atypical sessions began to exceed that of errors only for reactions slow enough to fall within the distributions obtained under the accuracy strategy. In his adjacent accuracy sessions, less than 1% of JF's responses were faster than 225 msec., while only about 5% of his responses in adjacent preprogrammed sessions were slower. In his atypical sessions, of the 22% of JF's responses that were faster than 225 msec., 48.5% were errors; the error rate was only 14.7% for all responses slower than 225 msec.

A speed-accuracy tradeoff?--All but a small part of the data for the four typical Ss could be easily attributed to one of the two distinct response strategies. Subject RZ, however, used an identifiable preprogrammed strategy only during the four sessions in Phase 2 of Experiment I that had a small value of D. During the remainder of Experiment I and throughout Experiment II his error rates varied reasonably continuously between .02 and .50. Did this performance result from a true tradeoff between speed and accuracy? Or did he use only two distinct strategies, but vary the proportion of trials on which he inserted preprogrammed guesses? It would seem that the question could be answered by fitting both random walk models and the Ollman-Yellott model to RZ's data, and comparing the fits.

Figure 5 presents RZ's speed-accuracy operating characteristic plots for both Experiment I (solid points) and Experiment II (open points). They plot RZ's inferred Ω (number of correct responses divided by the number of errors), on a logarithmic scale, as a function of his mean RT for each condition. The data from

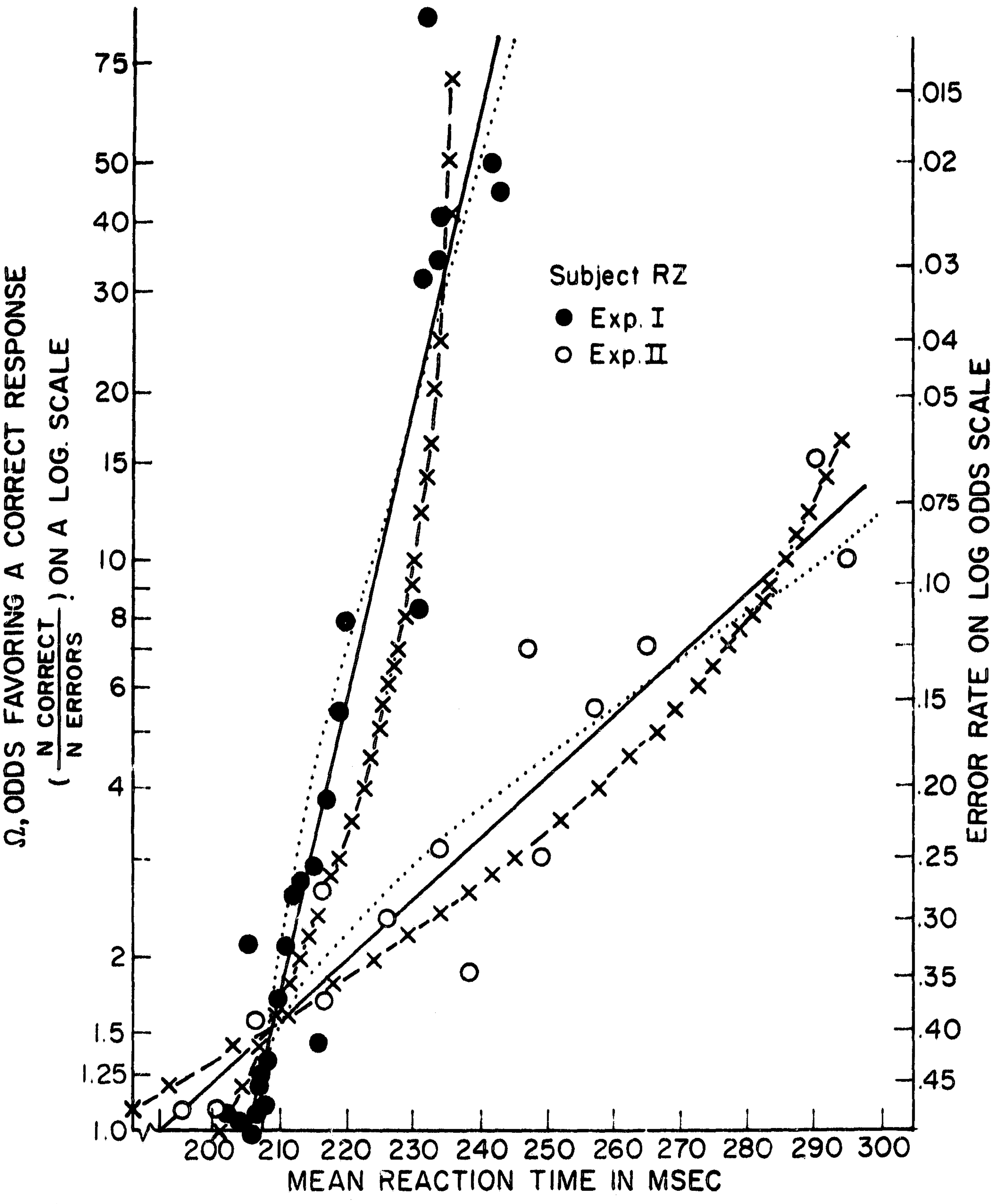

Fig. 5. Speed-accuracy operating characteristic functions for subject RZ in both Experiments I and II. These plot Ω (ratio of number of correct responses to number of errors) on a log scale against mean RT. Solid straight lines, dotted lines, and broken lines marked with crosses indicate the best fit functions for the fixed sample size and optional stopping versions of the random walk model and the guessing model.

Experiment I exclude the early preprogrammed sessions in Phase 2, which had mean RTs about 20 msec. slower than those of later sessions when his error rates exceeded .47. The three models were fitted to RZ's performance in each experiment by minimizing the squared deviations about the appropriate transformations of Ω. Correlations between RZ's mean RTs and the appropriate transformation of his error rates were .943, .952, and .926 for the optional stopping, fixed sample size and guessing models in Experiment I, and .905, .921 and .905 in Experiment II. Each model accounted for 80-90% of the variance between conditions in both experiments. Although distinct conceptually, they provided about equally good descriptions of RZ's performance; there was little basis for discriminating among them.

Attempts to differentiate between the strategy switching and tradeoff hypotheses by examining the RT distributions also failed. Particularly in Experiment I, RT variability was large compared with the rather small difference in mean RT between sessions with high and low error rates. Comparisons between non-adjacent conditions in Experiment II were questionable, since data from the four other Ss indicated substantial decreases in accuracy RT over at least the first 14 sessions.

Stimulus characteristics.--A major reason for changing the stimuli between Experiments I and II was an effect which may be unique to spatially separated stimuli. Using stimuli identical to those in Experiment I, Swensson and Edwards (1968) obtained slower preprogrammed RTs on trials when the stimulus failed to match the response S had preselected (errors). These and other data suggested that Ss may fixate the expected stimulus location or the location matching the response to which they are committed. Preprogrammed errors were slower than correct preprogrammed responses in their data. This also occurred in Experiment I. The upper part of Table 3 presents correct and incorrect preprogrammed responses for both Ss in Experiment I[1], together with comparable data for three earlier Ss (Swensson & Edwards, 1968).

Preprogrammed response times for the five Ss in Experiment II, in which both stimuli appeared over a common fixation point, appear in the lower part of Table 3. Although mean RT differences between erroneous and correct preprogrammed responses ranged between 2.0 and 16.2 msec. for the spatially separated stimuli, no such differences appeared with the new stimuli in Experiment II. For Ss LP and RZ, who served in both experiments, Table 3 shows slower mean RTs for incorrect preprogrammed responses to spatially separated stimuli than for their other preprogrammed mean times. This supports Swensson and Edwards's (1968) conclusion that uncertainty about the spatial location of the stimulus increases the time necessary to detect its presence only when it appears in the less favored location.

Discussion

Speed versus accuracy.--With both types of easily discriminable stimuli and the procedures of Experiments I and II, four of

[1]Although subject RZ never used an identifiable preprogrammed response strategy later in the experiment, his responses in sessions with error rates exceeding .47 seemed sufficiently independent of the stimulus sequence to warrant inclusion in Table 3.

Table 3

RT IN MSEC. FOR CORRECT AND ERROR PREPROGRAMMED RESPONSES FOR BOTH TYPES OF EASILY DISCRIMINABLE STIMULI

Type of Stimuli	Subject	Correct Responses (N)	Correct Responses Mean RT	Errors (N)	Errors Mean RT	Difference in Mean RT
Spatially separated (Exp. I; Swensson & Edwards, 1968)	JM*	(1181)	186.6	(1208)	188.6	2.0
	JD*	(577)	197.0	(547)	210.1	13.1
	TW*	(404)	183.7	(393)	189.5	5.8
	LP**	(1530)	198.0	(1481)	201.4	3.4
	RZ**†	(1152)	198.0	(1097)	214.3	16.2
	Mean <u>S</u>		192.7		200.8	8.1
Not spatially separated (Exp. II)	LP	(1333)	199.1	(1328)	199.5	.4
	RZ†	(378)	198.4	(355)	197.9	-.5
	JS**	(646)	168.2	(634)	167.8	-.4
	WW**	(1348)	206.0	(1316)	206.7	.7
	JF**	(1183)	196.5	(1186)	196.9	.4
	Mean <u>S</u>		193.6		193.8	.2

*Data from Swensson & Edwards (1968)

**Data from the first 8 or 10 sessions are not included.

†Data for RZ includes only sessions with error rates greater than .47.

the five Ss exhibited no trace of a speed-accuracy tradeoff in their performance, in spite of every effort to create conditions favorable to such a tradeoff. Almost all of the data from these four Ss could be attributed to the use of either an accuracy or a preprogrammed response strategy. Under the preprogrammed strategy Ss adopted some stereotyped rule to select their response, and thus obtained error rates close to .50 but decreased their mean reaction times. This decrease was larger for the stimuli in Experiment II, primarily because of longer RTs in that experiment under the accuracy strategy. Although accuracy response times improved considerably, the terminal mean difference in RT between the two strategies was still 45-70 msec., compared to 15-20 msec. obtained with the spatially differentiated stimuli used in Experiment I (Swensson & Edwards, 1968).

The performance of the fifth Ss, who participated in both Experiment I and II, could not be neatly categorized into two discrete response strategies. Although his data were quite orderly it was unclear whether the inverse relationship between his mean RTs and error rates was caused by the speed-accuracy tradeoff as predicted by the random walk models, or by varying amounts of intermittent preprogrammed guessing. Two different predicted speed-accuracy tradeoff functions obtained from the optional stopping and the fixed sample size versions of the random walk model, as well as the function predicted by the Ollman-Yellott guessing model, all provided equally excellent fits to the S's data for both experiments. The models look alike because of the relatively short additional time necessary for S to improve his accuracy from chance to near perfect with both types of easily discriminable stimuli. This is frustrating, but carries the methodological moral that models ought always to be evaluated against their competitors, never alone. An advocate of any of these models might defend its merits vigorously with these data--if he never tried to fit the other two.

Stimulus characteristics.--The stimuli used in Experiment II, both of which appeared at the same spatial location, removed the tendency for preprogrammed responses to be faster when the stimulus matched S's preprogrammed response. This phenomenon occurred in both Experiment I and an earlier study (Swensson & Edwards, 1968) where the identity of the stimulus (and thus of the correct response) depended on the spatial location in which it appeared. Coding stimuli by their spatial location has been a common practice in CRT experiments, which frequently use the onset of one of a set of lights as the stimulus. With such stimuli Bernstein, Schuman and Forester (1967) obtained slower mean RTs when the location of the stimulus was uncertain, even though the response had been predetermined.

Although Bernstein et al. (1967) interpreted these results as the effect of stimulus uncertainty, the data from Experiments I and II strongly suggest that uncertainty about the location, rather than about the identity of the stimulus is the critical factor. Data from the two Ss who performed with both types of stimuli in the present experiments indicate that slower RTs to the spatially uncertain stimuli are caused by longer latencies in detecting the presence of the stimulus when it did not match the response S had predetermined. Apparently, Ss fixate one of the stimulus locations and are slower to detect the presence of the stimulus when it appears elsewhere. The availability of this strategy for spatially separated stimuli could also account for the phenomenon of "behavioral hypotheses" (Bernstein & Reese, 1965; Bernstein et al., 1967). These studies demonstrated faster RTs when S's pre-trial prediction about the identity (location) of the stimulus was confirmed.

The tradeoff hypothesis.--None of the data in Experiments I and II provided compelling evidence for the tradeoff hypothesis. Most of the data, in fact, unequivocally argued that Ss had only two response strategies available--to use accurately or to ignore the identity of the stimulus in responding. But these results only sharpen the contradiction between the absence of a speed-accuracy tradeoff under the present conditions and its robust ubiquity in other experiments. One other major difference between these results and those demonstrating a tradeoff was the very short additional time these Ss required to decrease their errors from 50% to below 10%. Search for a speed-accuracy tradeoff with stimuli that are hard to discriminate should precede abandonment of the random walk models. Such stimuli should also permit more sensitive differentiation between the dichotomous preprogrammed and accuracy strategies.

EXPERIMENT III

This experiment replicated the procedures of Experiment II, using highly similar stimuli to provide greater differentiation between Ss' accurate and preprogrammed RT distributions. More difficult stimulus discriminations should not affect preprogrammed RTs, but should require longer times for Ss to achieve high levels of accuracy.

Method

Apparatus and stimulus display.--Except for the stimuli, the experimental situation was identical to that of Experiment II. The stimuli were of the same type as in Experiment II, but instead of two easily discriminable rectangles having length-width ratios of 30:1, their length-width ratio was much smaller. As seen in Figs. 2b and 2c, both rectangles looked very much like a hollow square, tilted at a 45° angle and centered on the small fixation cross. The length-width ratio (and therefore the difficulty of the discrimination) was established individually for each S. It represented the smallest length-width ratio (most difficult discrimination) that would allow the S to obtain practically zero error rates during intial sessions, provided that he took as much time as necessary to respond. The length-width ratios varied from 55:53 to 63:62, and appear in Table 4, together with the actual dimensions.

Table 4

LENGTH-WIDTH RATIOS AND ACTUAL DIMENSIONS IN INCHES OF THE STIMULUS RECTANGLES USED FOR EACH SUBJECT IN EXPERIMENT III

Subject	Length-Width	Ratio	Actual Dimensions (Inches) Length	Width
TM	55:53	1.038	1.520	1.464
EK	35:34	1.029	.966	.940
RP	40:39	1.026	1.104	1.078
GO	40:39	1.026	1.104	1.078
MC	40:39	1.026	1.104	1.078
MMcL	50:49	1.020	1.380	1.354
WL	63:62	1.016	1.740	1.712

Procedure.--The seven naive Ss were selected by the procedure described in Experiment I and each received $1.50 per hour during the training and initial sessions used to select the stimuli. The training terminated with two 300-trial sessions under a payoff matrix having a large value of D to encourage accuracy. The Ss were told that after the training sessions, their pay would depend entirely upon the number of points they won. The subsequent procedures were identical to those used for subject JF in Experiment II. None of the Ss had any opportunity to use or to discover the preprogrammed strategy before D systematically decreased during the first "real money" conditions. Each of the Ss received between 20 and 37 sessions of 350 trials. Table 5 presents the actual values of D used for each condition.

Table 5

DIFFERENCE IN PAYOFF BETWEEN CORRECT RESPONSES AND ERRORS (D)

USED FOR EACH SUBJECT AND CONDITION IN EXPERIMENT III

Condition	Subject						
	TM	EK	RP	GO	MC	MMcL	WL
1	800	800	800	400	800	800	800
2	500	600	600	200	600	500	500
3	200	400	400	200	400	200	200
4	200	200	200	300	200	200	200
5	300	200	200	400	200	300	300
6	400	300	300	500	300	400	400
7	500	400	400	600	400	500	500
8	600	500	500	500	500	600	600
9	800	600	600	400	600	800	800
10	300	600	600	300	600	1000	1000
11	350	300	300	350	200	1000	250
12	300	250	200	350	200	300	225
13	275	250	225	400	150	250	275
14	290	225	225	350	150	275	260
15	280	225	250	300	175	260	280
16	270	250	250	275	200	260	300
17	260	275	235	290	225	240	310
18	260	250	230	305	200	220	325
19	240	250	200	300	185	220	275
20	240	250	225	310	185	210	290
21	225	250	215	350	195	225	280
22	250	235	220	350		215	270
23	250	225	225	310		215	260
24	240	225	220	305		220	230
25	230	210	225	280		225	220
26	235	200	222			222	225
27	235	190	220			224	230
28	232	180	221			224	235
29	230	170	221			220	240
30	225	170	200			210	250
31	220	170				210	260
32	225	180				210	270
33	235					200	270
34	230					205	265
35	235					215	255
36	240					215	245
37	250					225	

Results

Mean RTs and error rates over sessions for the accuracy and preprogrammed modes of responding appear for each of the seven Ss in Figs. 6a-6f and 7a. Open points represent sessions in which S used either an accuracy (circles) or preprogrammed strategy (triangles) throughout. Solid points indicate mean RT (lower panels) and error rate (upper panels) for each type of performance in sessions that included some preprogrammed responding. Triangular and solid square points represent the two different criteria for identifying preprogrammed responses, explained below. Use of these two criteria for all sessions meant that each point for S's accuracy strategy performance (circles) contains only responses with RTs of 250 msec. or slower. Solid lines indicate running average RT and error rates for each strategy, over adjacent sets of three sessions. Vertical bars periodically indicate RT variability. Part of a single session for four of the Ss did not clearly separate into accuracy and preprogrammed strategies when the two criteria were applied. Performances on these blocks of trials appear in the figures as circles with an asterisk inside; they were not included in the running averages.

Forced to discriminate between two highly similar stimuli, all seven Ss had substantially slower mean RTs under the accuracy strategy than the Ss in Experiment II, but their preprogrammed RTs were about the same. They did not display speed-accuracy tradeoffs. All Ss easily discovered and used the preprogrammed strategy when the value of D in the payoff matrix decreased until their

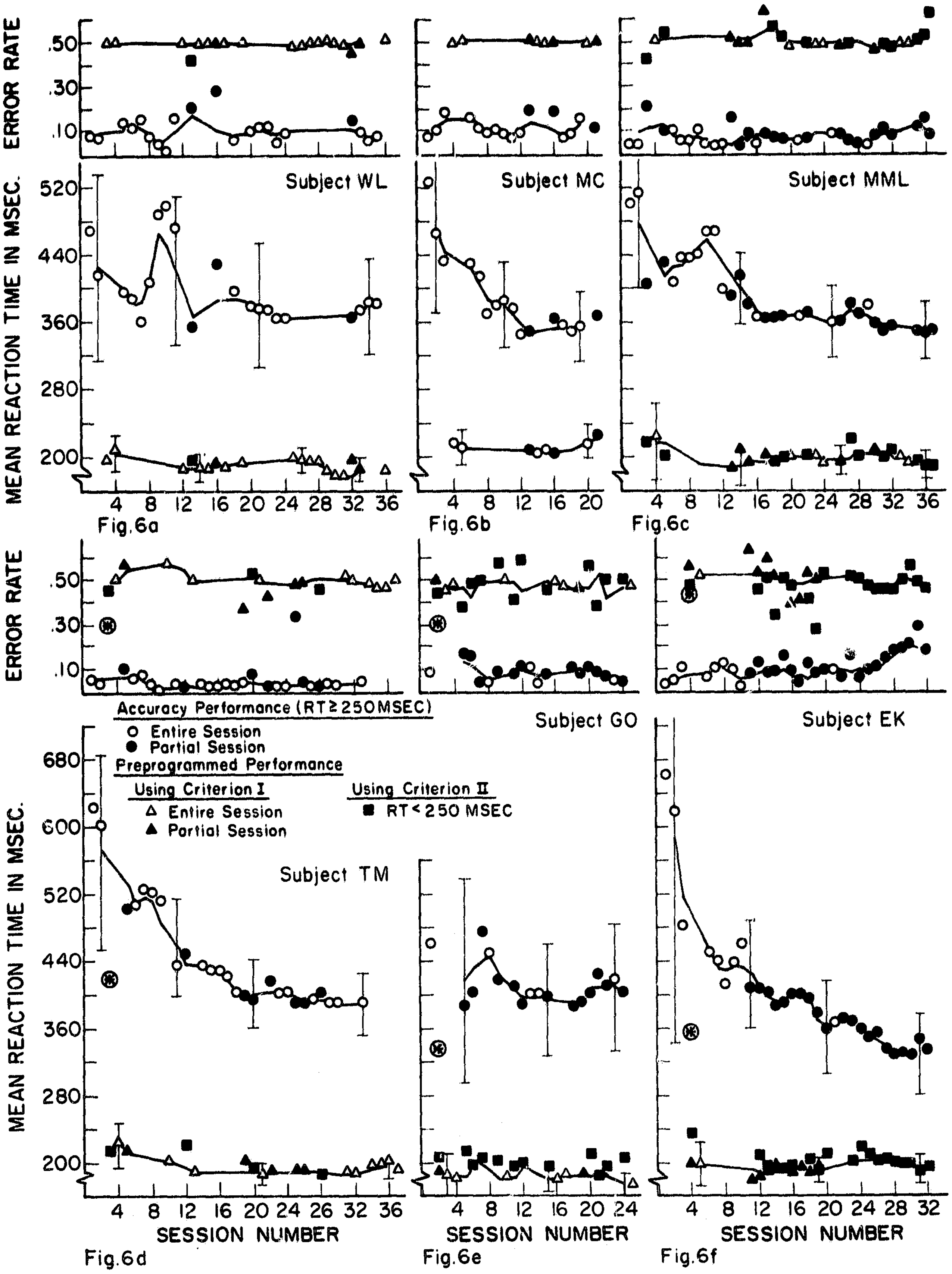

Fig. 6. Mean RT (lower panels) and error rates (upper panels) for accuracy (RT ≥ 250 msec.) and preprogrammed strategy performance of six Ss in Experiment III. Open and filled points differentiate entire sessions under the same strategy from mixed sessions, in which S used both strategies. Triangular points represent preprogrammed trials identified by the blocking procedure (Criterion 1); solid square points represent all trials faster than 250 msec. (Criterion 2). Solid lines indicate moving averages over adjacent sessions and vertical bars indicate RT variability. The single performance indicated by a circled asterisk for three Ss did not separate easily into the two strategies (see text).

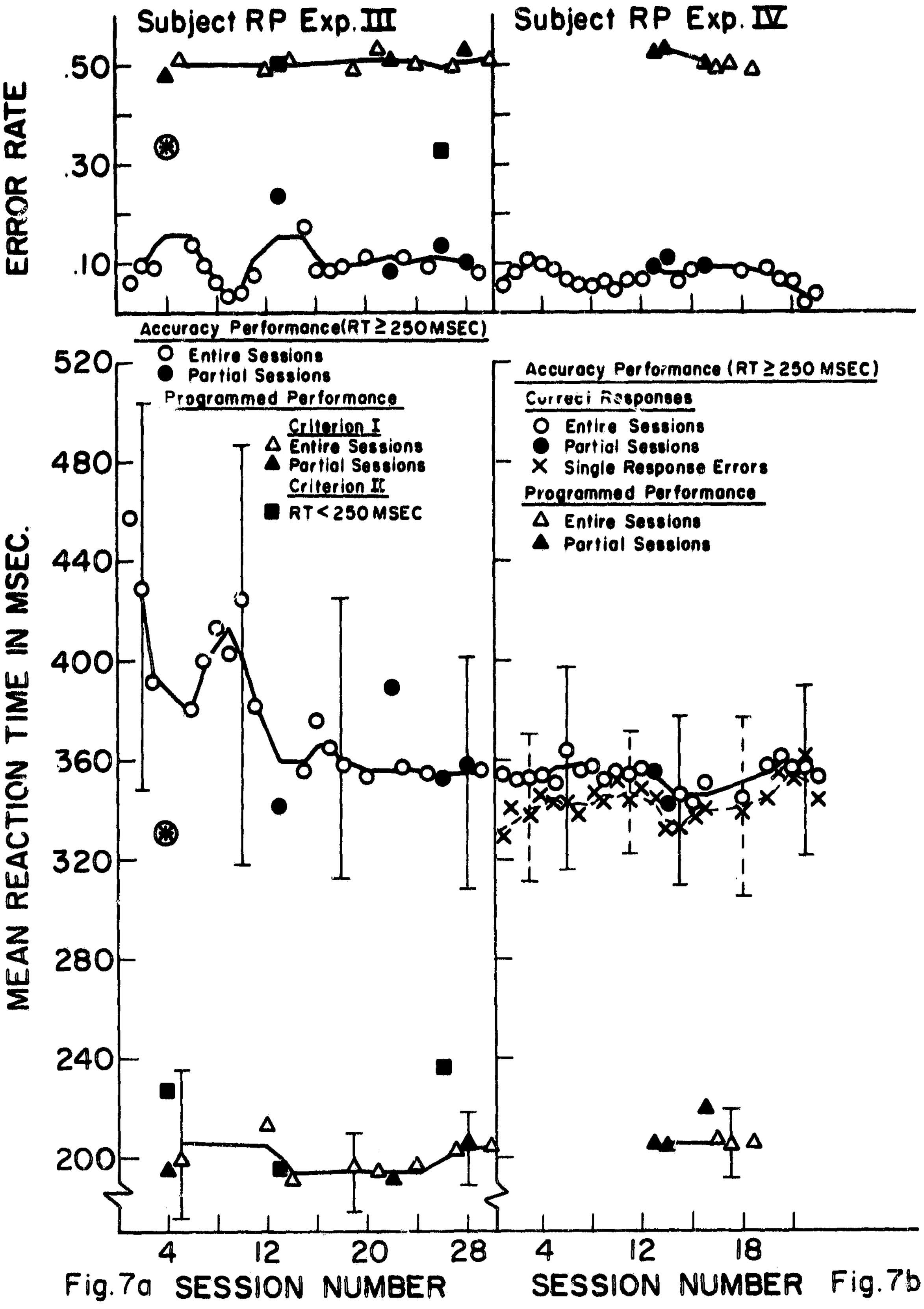

Fig. 7. Subject RP's mean RT (lower panels) and error rates (upper panels) for accurate (circles) and preprogrammed strategy (triangles) performance in both Experiments III(Fig. 7a) and IV (Fig. 7b). Open and filled points differentiate conditions in which RP used the same strategy on all 350 trials from mixed sessions, in which S used both strategies. Fig. 7b separates mean RT for correct responses (circles) and single response errors (crosses). Solid lines indicate moving averages for error rates and mean RT; in Fig. 7b solid and dotted lines differentiate moving average RT for correct responses and single response errors under accuracy. The single performance indicated by a circled asterisk did not separate easily into the two strategies (see text).

response speed primarily determined their net payoffs. Because of the slower accuracy response times in this experiment, there was practically no overlap in the RT distributions for preprogrammed responses and responses based on the identity of the stimulus. This was fortunate, since only two Ss confined their use of the preprogrammed strategy entirely to long blocks of trials with a simple rule for selecting which response to preprogram. The other five Ss at least occasionally embedded very fast responses (with a high frequency of errors) among attempts to respond accurately, even in sessions having low overall error rates.

This intermittent guessing required the use of two criteria to identify preprogrammed responses. The first, indicated in Figs. 6a-6f and 7a by open and filled triangles, was identical to that used in Experiments I and II. It consisted of segregating relatively long blocks of trials within which S used a consistent and mechanical rule to select his responses.

Criterion 2, applied for every condition, separated all trials on which S's RT was less than 250 msec. Since very few of any S's preprogrammed responses based on the first criterion were longer than 250 msec., this second criterion identified almost all of those trials on which Ss took a guess and preprogrammed their response. The large number of entire sessions plotted under the accuracy strategy (open circles), however, indicates that much of the Ss' accurate performance was not affected by this 250 msec. criterion time. Filled square points in Figs. 6a-6f and 7a denote mean RT and error rate within each session for responses faster than 250 msec. These points tend to be variable, because they frequently contained relatively few observations. The general levels of error rates and mean RTs for preprogrammed responses identified by Criterion 2, using the 250 msec. criterion time, however, were comparable to those isolated by the blocking method used in the earlier experiments.

Solid lines in Figs. 6a-6f and 7a represent the weighted running averages of mean RTs and error rates for preprogrammed and accuracy performance, over every three adjacent sessions. Most of the deviant points have little effect on the running averages because they usually represent only a small number of responses. Vertical bars, which periodically indicate plus and minus one average standard deviation of the individual RT distributions, are the square root of the weighted average variance for the adjacent three sessions under each type of performance. Circled asterisks indicate the single data point for four of the seven Ss that could not be clearly segregated by the two criteria into preprogrammed and accuracy strategies. Error rates and mean RTs for these few sessions fall in between those for the two types of performance. Since these intermediate performances occurred only on the first occasion each S experienced a payoff matrix with a small value of D (demanding very fast responding), they probably reflect the Ss' inefficient and sporadic first attempts to apply a preprogrammed strategy.

Mean RTs for accurate performance decreased over sessions, as in Experiment II. With the possible exception of the final sessions for subject EK (Fig. 6f), this decrease reflects absolute improvement in performance, since Ss' error rates showed little change over sessions. However, the figures show that Ss' error rates, as well as their RT means, were generally higher for accuracy strategy performance with these more difficult stimuli than they were in Experiment II, with easily discriminable stimuli. The small amount of improvement and the general magnitude of Ss' preprogrammed RT was comparable to that found in both Experiments I and II. The RT variability decreased over sessions for both types of performance, but the decrease was much greater for accuracy strategy times, whose original variability was much greater. By the end of the 20-37 sessions of the experiment, differences in mean RT between preprogrammed and accurate performance varied from about 135 to 210 msec. for the seven Ss.

Discussion

The data from this experiment were unambiguous. All seven Ss blatantly contradicted the tradeoff hypothesis. Each either preprogrammed his response or responded much more slowly, with relatively low error rates. The differences in mean RT between these two modes of performance were about three times as great as those obtained with the easily discriminable stimuli in Experiment II, even after large amounts of practice. Five of the seven Ss did not confine their preprogramming to long blocks of trials, but the clear separation between RT distributions permitted identification of preprogrammed guesses, embedded among trials on which Ss attempted to respond accurately.

As in Experiment II, the procedure varied the payoff matrix in an attempt to converge upon intermediate values of D that would either produce high error rates without complete abandonment of accuracy, or force Ss to switch between accurate and preprogrammed strategies. As a consequence, the difference in payoff between correct responses and errors varied over only a small range for each S during the later conditions, as he switched between strategies or sharply varied the proportion of preprogrammed trials. Only subject EK (Fig. 6f) consistently performed with error rates exceeding 15% in the later sessions, when responding to the identity of the stimulus.

This experiment completely failed to demonstrate that Ss could, if necessary, process only intermediate amounts of evidence, as predicted by the speed-accuracy tradeoff hypothesis of the random walk models. The tradeoff phenomenon remains ephemeral; and the conflict in results between these experiments and others (e.g. Pachella & Pew, 1968; Schouten & Bekker, 1967) remains unresolved.

However, Ss' error rates under the accuracy strategy were consistently higher in Experiment III than they had been in Experiments I and II. This was puzzling, since each S could and did achieve error rates as low as .01-.02 when he ignored his response times during the initial training. Apparently, the Ss refused to respond as accurately as they could in Experiment III. Although the tradeoff hypothesis seems unable to cope with the absence of very high error rates, it may explain Ss' refusal to respond more accurately than they did. Perhaps achieving very low error rates would have demanded prohibitively expensive increases in RT for the intermediate values of D during most of Experiment III. The use of larger values of D would increase the cost for errors relative to the cost for time. These conditions should force Ss to respond more accurately in order to maximize their payoffs.

EXPERIMENT IV

In this experiment two experienced Ss responded to the same stimuli as in Experiment III, but payoff matrices varied over a large range of values of D. The principal question was whether the Ss, in response to pressure from the payoff matrix to achieve very low error rates, would be forced to respond more slowly. Following this, the experiment further investigated the effects of stimulus discriminability by giving one S extensive training on an even more difficult pair of stimuli.

Method

The two Ss used, RP and EK, had worked for 30 and 32 sessions in Experiment III with stimuli having length-width ratios of 40:39 (1.026) and 35:34 (1.029). During the final 10-15 sessions of Experiment III, Fig. 7a shows that subject RP's accuracy performance had remained fairly stable, with about 10% errors and mean RTs of 350-360 msec. Subject EK (shown in Fig. 6f) displayed the greatest change in accuracy performance during Experiment III. His initial accuracy sessions had low error rates and the slowest mean accuracy RTs, while his final sessions had the fastest mean times (325 msec.) and highest error rates (up to .20) of all seven Ss.

Except for the variation in D over conditions, the apparatus, procedures and stimuli were identical to those used for these two Ss in Experiment III. Over successive conditions D increased and decreased systematically, from a value small enough to produce preprogrammed responding (or about 50% errors) to values as high as 1 1/2 times as large as S's cost per second of response time. Under these large values of D, a response as slow as 1450 msec. was preferable (in terms of payoff) to making an error; one error more or less, could change Ss' net payoff by as much as 800 points ($2.00). The 26 conditions for subject RP and 32 conditions for subject EK traversed this range of D in both ascending and descending order, as Table 6 shows.

Following these 32 conditions and after a break between semesters, subject EK performed in 56 conditions with a much more difficult pair of stimuli. These had a length-width ratio of 85:84 (1.012) and measured 2.348 x 2.320 inches. Table 6 presents the values of D over these 56 conditions. D increased and decreased systematically from values low enough to produce 50% errors, to values almost 20 times as large as the cost per second of observing time. Even though the exchange rate for points was deflated from 1/4 to 1/10 cent per point, these extreme values of D resulted in variations in net payoff of as much as $10.00 for a single error. Although S's income from individual sessions varied a great deal (and he frequently lost money) he averaged about $2.25 per hour over the entire 56 sessions.

Results

Subject RP.--Figure 7b presents the RT and error rate data from the 26 conditions for subject RP. The display format is identical to that for Fig. 7a and the earlier figures, except for

Table 6

DIFFERENCE IN PAYOFF BETWEEN CORRECT RESPONSES AND ERRORS ($\overline{D}$) USED IN EXPERIMENT IV FOR BOTH SUBJECTS WITH THE SAME STIMULI AS IN EXPERIMENT III AND USED FOR THE 56 CONDITIONS WITH THE MORE DIFFICULT (85:84) STIMULI FOR SUBJECT EK

Stimuli Used in Experiment III			85:84 Stimuli for Subject EK			
Condition	RP	EK	Condition	D	Condition	D
1	210	175	1	800	33	250
2	250	175	2	600	34	225
3	300	175	3	400	35	200
4	300	165	4	300	36	175
5	300	160	5	200	37	150
6	500	150	6	100	38	125
7	600	175	7	300	39	100
8	700	200	8	500	40	75
9	700	225	9	700	41	50
10	400	250	10	700	42	100
11	250	300	11	900	43	200
12	225	500	12	900	44	300
13	210	700	13	1100	45	400
14	200	400	14	1100	46	600
15	190	250	15	1100	47	900
16	180	210	16	1500	48	900
17	170	190	17	2000	49	1200
18	160	180	18	3000	50	1200
19	175	170	19	3000	51	2000
20	185	160	20	5000	52	2000
21	180	150	21	8000	53	4000
22	250	140	22	10000	54	4000
23	400	125	23	10000	55	6000
24	600	100	24	900	56	6000
25	800	75	25	500		
26	500	150	26	300		
27		175	27	250		
28		200	28	900		
29		250	29	900		
30		350	30	500		
31		500	31	400		
32		700	32	300		

RTs under the accuracy strategy. Circular points represent mean RT for only RP's correct responses under the accuracy strategy. Crosses indicate his mean RT for single response errors, trials on which he retracted only the incorrect finger (excluding trials when he made both responses less than 100 msec. apart). These comprised 88% of his total errors under the accuracy strategy. Solid and dotted lines separately indicate the running average RT over each adjacent set of three conditions for the correct responses and the single response errors. Solid and dotted vertical lines differentiate periodic indications of the RT variability for the two types of responses.

In conditions with small values of D RP used the preprogrammed strategy for long blocks of trials or the entire session, with performance identical to what it had been in Experiment III. When RP did not preprogram, however, his performance still remained almost constant, even for the largest values of D. The decrease in error rate for accuracy performance during the final few sessions was not accompanied by any change in his mean RT. Figure 7b shows that single response errors were consistently 10-15 msec. faster than his correct responses, but their times were only slightly less variable.

Subject EK: Mean performance.--While the accuracy performance of subject RP was almost constant over the entire range of D, subject EK displayed unmistakable evidence of a speed-accuracy tradeoff for both the moderately difficult (35:34) and very difficult (85:84) stimulus discriminations. Over part of the range of small D values, EK interspersed preprogrammed guesses with accuracy trials, as he had in Experiment III. Since clearly bimodal RT distributions resulted from mixing these two strategies with both the 35:34 and the 85:84 stimuli, Criterion 2 of Experiment III separated preprogrammed guesses (RT < 250 msec.) from trials on which S attempted to respond accurately (RT ≥ 250 msec.). Preprogrammed responses identified in this way had error rates and mean RT of .475 and 199.2 msec. for the 35:34 stimuli and .502 and 196.1 msec. for the 85:84 stimuli, both comparable to S's preprogrammed response performance in Experiment III.

Figure 8 presents subject EK's speed-accuracy operating characteristic functions for responses longer than 250 msec. These plot Ω on a logarithmic scale against mean RT for each condition. Circular points indicate his performance over the 32 conditions with the 35:34 stimuli, and diamond-shaped points indicate performance over the final 27 sessions with the 85:84 stimuli. Open points represent conditions in which all of his responses were longer than 250 msec.; solid points represent conditions in which some preprogramming occurred.

The data in Fig. 8 clearly demonstrate that, whenever subject EK did not choose to preprogram, he could and did vary his speed and accuracy in response to changes in the payoff matrix. Moreoever, the error rates and mean RTs varied in a tradeoff relationship; low error rates required slower mean RTs, while high error rates permitted faster mean times. With the easier 35:34 stimuli, EK could reduce his error rate to less than .01, but there was some

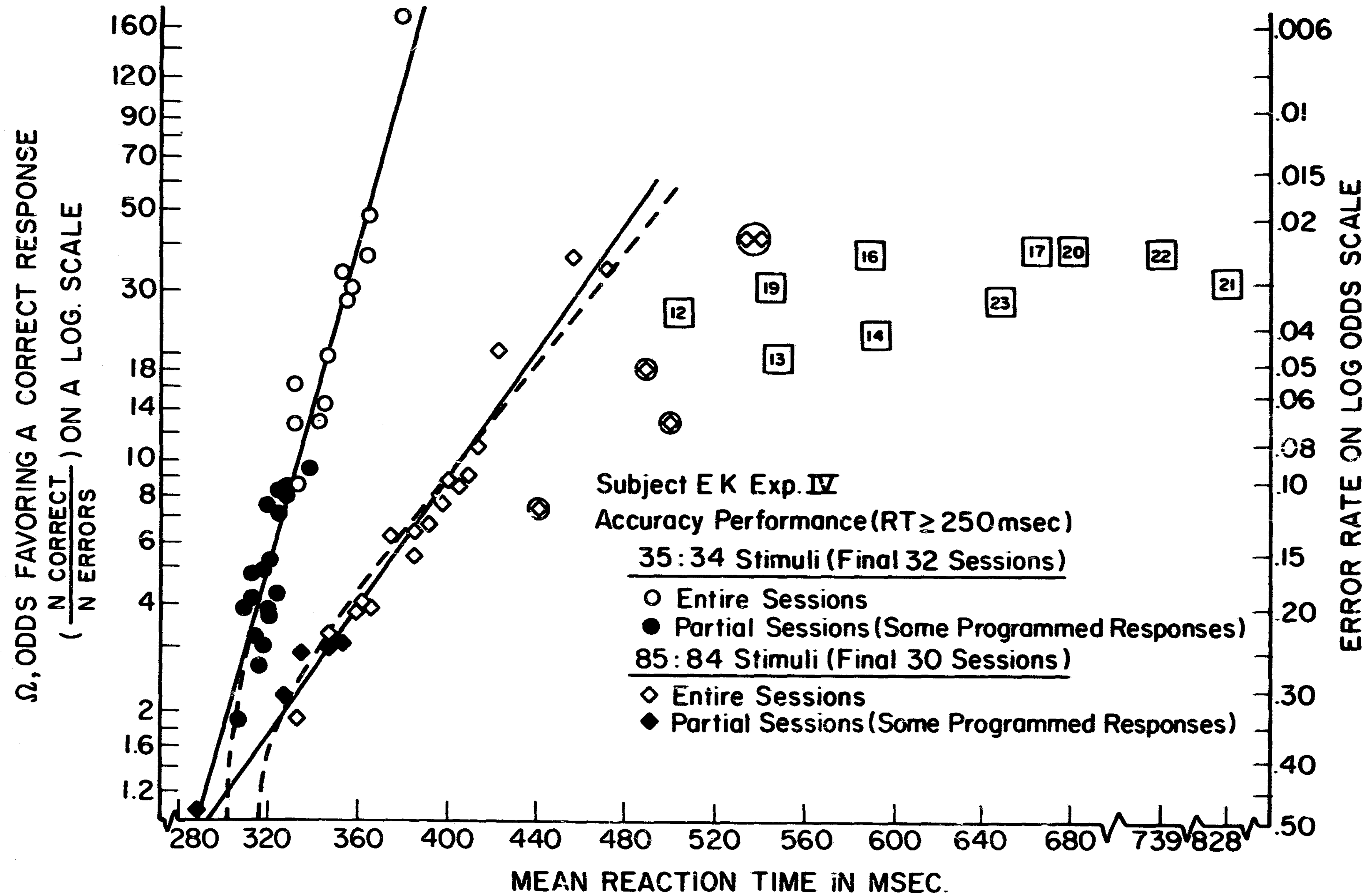

Fig. 8. Speed-accuracy operating characteristic functions for subject EK's performance with both discrimination tasks in Experiment IV. These plot Ω (ratio of number of correct responses to number of errors) on a log scale as a function of mean RT for discrimination performance (RT ≥ 250 msec.) with both 35:34 stimuli (sessions 33-64) and 85:84 stimuli (sessions 27-56). Open points represent conditions for which RT ≥ 250 msec. on all 350 trials; filled points represent conditions in which some preprogramming occurred. Solid lines and dotted curves indicate best fit functions of the fixed sample size and optional stopping versions of the random walk model. Numbered square points show EK's highly accurate performance (5% errors or less) in earlier sessions with 85:84 stimuli. Neither these data nor those indicated by circled points were used in fitting the models (see text).

evidence of a ceiling on his accuracy for the more difficult 85:84 stimuli. The locations of the numbered squares in Fig. 8 indicate EK's performance in the conditions prior to session 34 which had error rates less than .05. Although in some of these earlier sessions EK drastically slowed his mean RTs (to 830 msec.), he was unable to reduce his error rate below .03. The ordering of the session numbers indicates that these very slow performances cannot be dismissed as the result of insufficient practice. They occurred during his first experience with payoff matrices placing such extreme demands upon accuracy that obtaining a reaction time of 5-20 seconds was preferable (in terms of net payoff) to making an error. Many of these individual reaction times were longer than 2-3 seconds. Since EK could achieve error rates as low as .03 with a much faster mean RT, he rarely allowed his individual response times to extend beyond one second during later conditions with even the most extreme values of <u>D</u>.

Solid and dotted lines in Fig. 8 show the best fit predictions of the optional stopping and fixed sample size versions of the random walk model. These were fit by least square methods to $\left(\frac{\Omega - 1}{\Omega + 1}\right) \ln \Omega = \alpha + b$ RT and $\ln \Omega = \alpha + b$ RT. Conditions represented by circled points for the 85:84 stimuli in Fig. 8 were not used to fit the models, since these reflected either unusually poor performance or the apparent accuracy ceiling effect. Both models provided good fits to these data. For the 35:34 stimuli (N = 32) product-moment correlations between mean RTs and the appropriate error rate transformations were .953 and .950 for the optional stopping and fixed sample size versions of the random walk model. For the 85:84 stimuli (N = 22), the correlations were .969 and .986. The reason the data provide little basis for preference between the two versions is that both predict almost identical functions when error rates are less than .33 ($\Omega > 2.0$)--the region of speed-accuracy tradeoff performances that contained almost all of the data.

The predicted mean RT intercept for error rates of .50 ($\Omega =$ 1.0), based on the fixed sample size model, was almost identical for both the 35:34 stimuli (289 msec.) and the 85:84 stimuli (292 msec.). Extrapolating the best fit optional stopping models predicted a somewhat slower intercept for chance performance with the more difficult 85:84 stimuli (317 msec.) than with the 35:34 stimuli (302 msec.).

These predicted intercept times imply that S's accuracy would have dropped to chance level if he had attempted to respond as quickly as 290-320 msec., with either discrimination task--90-120 msec. slower than his mean RT under the preprogrammed strategy. Such a dead time interval, with no improvement in accuracy unless responses were delayed considerably longer than preprogrammed RT (e.g. for 90-120 msec.), might explain the absence of any tradeoff in the earlier experiments. This point will be developed more fully later, but the fastest predicted tradeoffs for EK would have cost him much more time than preprogramming, for only slightly improved error rates.

The fitted slopes of the functions relating mean reaction times to $\ln \Omega$ or $\left(\frac{\Omega - 1}{\Omega + 1}\right) \ln \Omega$ can both be interpreted as the rate at which S obtained evidence to discriminate between the two stimuli.[2] For both models these estimated rates were considerably slower for the more difficult 85:84 stimuli, even though by then EK had practiced for more than 150 hours on this and the easier task. In natural log units, the estimated rates of increase in evidence per 10-msec. increase in mean RT for the fixed sample size and optional stopping models were .517 and .591 with the 35:34 stimuli and .203 and .213 with the 85:84 stimuli.

Subject EK: RT for correct responses versus errors.--In addition to a speed-accuracy tradeoff, both versions of the random walk model predict that correct responses and errors should have the same RT distribution, although no present version specifies its exact form. Figure 9 shows subject EK's mean RT for correct responses and single response errors, plotted against the odds favoring a correct response on a log scale for both the 35:34 and 85:84 stimuli. To obtain more stable estimates of error RTs, the data were combined across conditions that had similar error rates within each of the two levels of stimulus difficulty. Single response errors accounted for 92% and 98% of all the errors EK made within the 35:34 and the 85:84 stimulus conditions. Figure 9 shows that errors were substantially faster than correct responses over all levels of accuracy for the easier 35:34 stimulus discrimination.

[2]Since the prior odds were 1.0, ln Ω and $\left(\frac{\Omega - 1}{\Omega + 1}\right)$ ln Ω can be used as each model's estimate of the expected log likelihood ratio (base e) for the total evidence that S obtained to discriminate between the two stimuli on each trial. Green and Swets (1966, pp. 171-173) have shown that this quantity is equal to $(d')^2$. The slopes quoted in the text are thus numerically equal to the rate of increase in the square of the familiar d' parameter of signal detectability theory, per 10-msec. increase in mean RT.

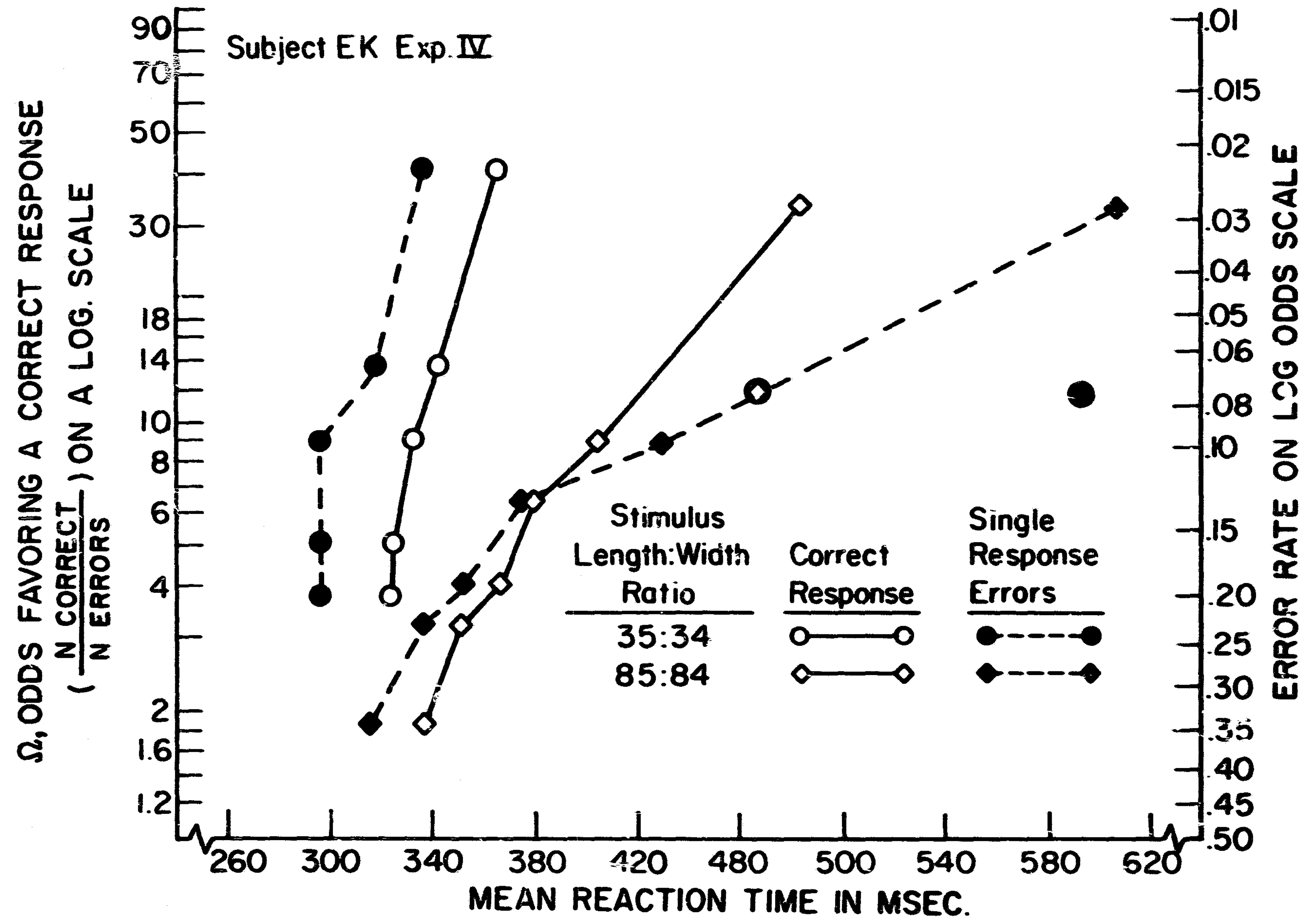

Fig. 9. Subject EK's speed and accuracy for correct responses and single response errors in grouped sessions within each discrimination task. These plot mean RT for each type of response against Ω for the pooled sessions on a log scale. Circled points for the 85:84 stimuli represent data from the five circled sessions in Fig. 8.

With the very difficult 85:84 discrimination, however, error times were faster than correct responses only when both were fast and his error rates were high. As EK's accuracy increased this difference decreased and reversed, until mean error RTs became 125 msec. slower than mean correct response RT when his performance was most accurate.

The considerably slower RTs for errors with the most difficult 85:84 stimulus discrimination appeared related to the apparent ceiling of EK's accuracy, noted for the data presented in Fig. 8. The data from the earlier, very slow, sessions with error rates less than .05 were not included in Fig. 9, but as his overall mean RT continued to increase, so did the difference between mean RT for errors and correct responses. In the five slowest of these earlier conditions (not the earliest, but those which exerted the greatest pressure for very accurate performance) EK's error rate was .030 and his mean error RT was 775 msec. slower than his mean time for correct responses.

Discussion

One of the two Ss displayed clear evidence of a speed-accuracy tradeoff, as changes in the payoff matrix varied the relative importance of speed and accuracy. The form of his tradeoff function suggested that Ss in Experiment III (and perhaps also Experiment II) may have refused to exhibit performance with high error rates if they made any attempt to respond accurately, although they could have done so. Any reasonably continuous extrapolation of subject EK's tradeoff functions indicated that chance error rates would have been reached with discrimination times still considerably slower than his mean preprogrammed RT. For both versions of the random walk model and two levels of stimulus difficulty, the estimated additional delay was 90-120 msec. before EK's accuracy could exceed chance level.

Since the time-cost function charged Ss an amount proportional to their total response time, such a fixed delay meant that Ss would have had to accept a fixed cost for any attempt to respond based on evidence about stimulus identity. This fixed cost may have forced Ss to preprogram their responses, avoiding the additional processing necessary to obtain evidence about the stimulus identity until the payoff matrix placed a high premium on accuracy. Thus, processing intermediate amounts of evidence about the stimulus identity may never have been an optimal strategy for Ss. Only under high values of D, demanding the acquisition of considerable amounts of information to obtain low error rates, would an accuracy strategy produce a higher expected payoff than the preprogrammed strategy. In this experiment, subject EK's fastest and least accurate performances actually resulted in lower payoffs than he could have obtained by preprogramming. Since Experiments II and III used values of D just large enough to produce accurate performance, slow performance with very low error rates may also have been suboptimal.

If the cost function charged S only for his actual observing time, the substantial savings in time enabled by the preprogrammed strategy would be irrelevant to his payoffs. Removing the high fixed cost for initial stimulus observations would make S's cost per observation independent of the number of observations taken. This, of course, is the assumption made by both versions of the random walk model.

EXPERIMENT V

This experiment tested the fixed cost hypotheses suggested by the results of Experiment IV, by modifying the linear time-cost function to include a free delay interval. With this modification, S was charged a linear function of only that portion of his total response time exceeding t msec. (the length of the free delay). If the hypothesis is correct, S should trade accuracy for increased speed for an appropriate fixed delay and low values of D. Ss' willingness to exhibit formerly suboptimal tradeoffs, together with payoff matrices that vary D over a wide range of values, should enable mapping of their speed-accuracy operating characteristic functions, as done for subject EK in Experiment IV.

Method

The experiment used three experienced Ss, whose performance had been stable for the final 10-15 sessions of Experiment III and devoid of any speed-accuracy tradeoff. The apparatus, stimuli, and most of the procedural details were identical to those used for these Ss in Experiment III. Under the free delay modification, Ss received feedback about the length of their RT only when this time exceeded t, the free delay duration. If their RT was less than t, the feedback displayed only zero cost for time. The cost for additional time was 550 points per second, the same rate as in the earlier experiments.

Since subjects WL and MMcL had error rates of about .10 during their final sessions in Experiment III, their payoff conditions included both large and small values of D. The maximum value of D used was 5.5 times the cost per second for response time in excess of t. Because subject TM had responded with the lowest accuracy error rates in Experiment III (.01-.02), he received only free delay conditions with small values of D to encourage faster and less accurate performance. Table 7 presents the actual values of D used.

Table 7

DIFFERENCE IN PAYOFF BETWEEN CORRECT RESPONSES AND ERRORS (D) USED FOR EACH SUBJECT AND CONDITION IN EXPERIMENT V

Condition	Subject TM	Subject MMcL	Subject WL
1	200	200	200
2	175	150	175
3	160	125	185
4	160	125	175
5	170	100	225
6	170	175	300
7	150	250	400
8	125	350	600
9	125	500	900
10	125	500	900
11	80	1,000	900
12	65	1,500	1,200
13	65	3,000	1,800
14	55	3,000	3,000
15		300	300
16		200	250
17		150	200
18		100	150
19		75	

All three Ss received one training session with $t = 250$ msec. at an hourly rate of pay to accustom them to the new time-cost feedback. This session contained three blocks of 100 trials each under a high, moderate and low value of D. It became obvious to all three Ss that preprogrammed responding would result in zero cost for time and a guaranteed average net payoff of D/2 points per trial. The E then pointed out that performance just fast enough to obtain zero cost for time, but with any fewer than 50% errors, would always earn a higher net payoff than the preprogrammed strategy. To minimize Ss' temptation to preprogram their responses, E reduced D gradually, so that one of their previous accuracy performances always had a higher expected payoff than the preprogrammed strategy. Subjects WL and MMcL received 18 and 19 test sessions of 350 trials each, all with a free delay of 250 msec. For subject TM, t was set at 250 msec. for the first five sessions of 350 trials, but increased to 300 msec. for sessions 6-14.

Results

Preprogrammed responses.--Some preprogrammed responding occurred in the data for all three Ss. As he had done in Experiment III, subject TM always used a single response for relatively long blocks of trials under the preprogrammed strategy. Although these responses were obviously made independently of the stimulus identity and had an error rate of .481, their mean RT was 218.4--more than 20 msec. slower than TM's mean preprogrammed RT at the end of Experiment III. This increase probably resulted from the absence of any pressure to respond much more quickly than the free delay interval (250 and 300 msec.), together with the lack of feedback about how fast or slow these responses actually were. The same phenomenon occurred in the data for Ss MMcL and WL, whose preprogrammed responses had to be identified by their speed, using Criterion 2 of Experiment III. Of all responses faster than 250 msec. for Ss MMcL and WL, error rates were .512 and .493, but their mean RT was 211.7 and 224.4 msec. These mean times were about 20 and 40 msec. slower than their mean preprogrammed RT in the final sessions of Experiment III.

Mean performance.--Even after preprogrammed trials were removed, the free delay conditions successfully obtained faster and less accurate performance than any of the three Ss had been willing to exhibit in Experiment III. Figures 10a-10c present their speed-accuracy operating characteristic functions for responses slower than 250 msec. (mean RT against odds favoring a correct response on a logarithmic scale). Figures 10a and 10b indicate that the large variations in D, combined with the free delay, induced subjects MMcL and WL to trade accuracy for speed over a wide range of performance. The procedure for subject TM encouraged and obtained substantially higher error rates than he produced in any of his final nine accuracy sessions during Experiment III (indicated by triangular points in Fig. 10c).

Table 8 presents the results of fitting the fixed sample size and optional stopping versions of the random walk model to

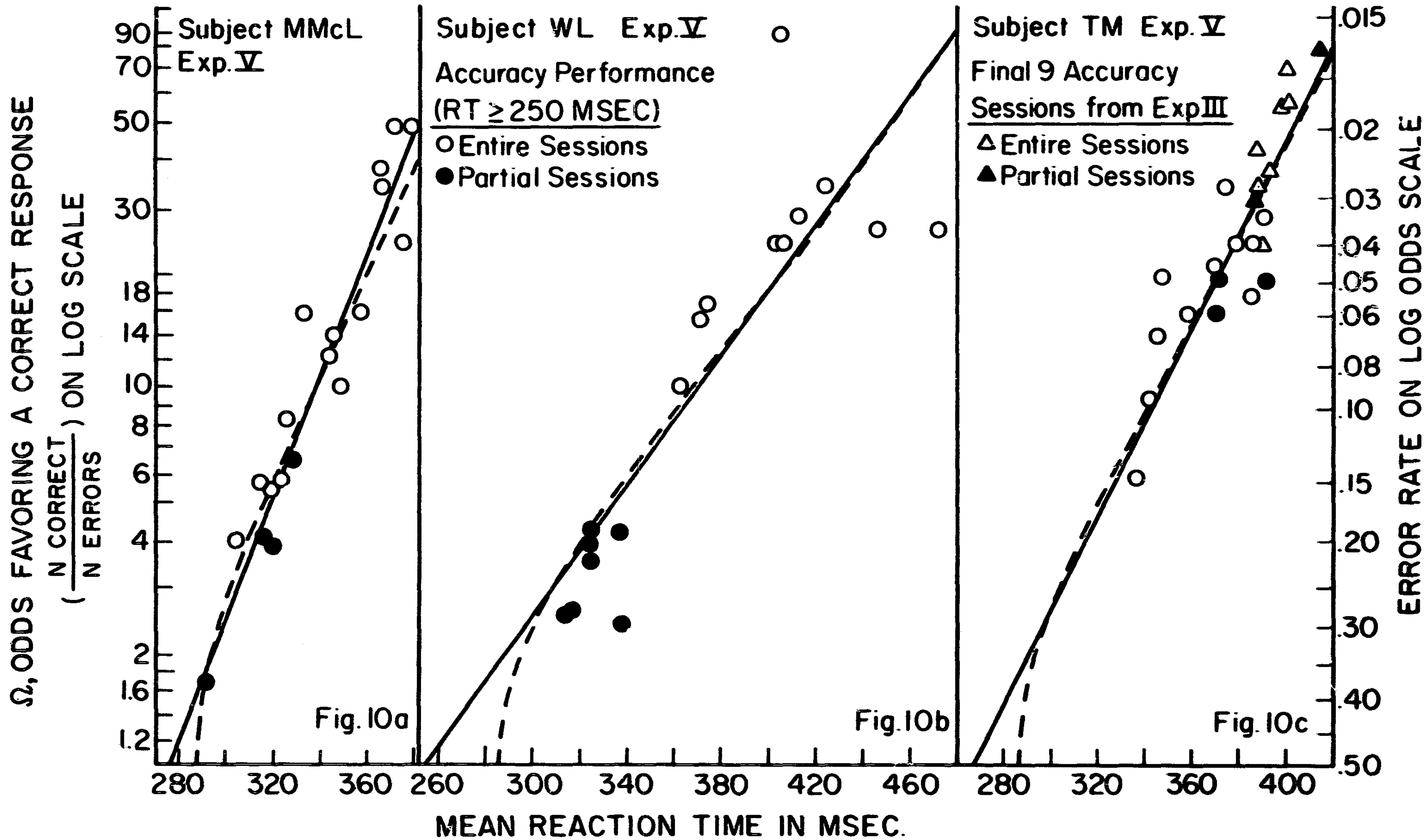

Fig. 10. Speed-accuracy operating characteristic functions for each of the three Ss in Experiment V. These plot Ω (ratio of number of correct responses to number of errors) on a log scale as a function of mean RT for each S's discrimination performance (RT ≥ 250 msec.). Open points represent conditions with RT ≥ 250 msec. on all 350 trials; filled points represent conditions in which some preprogramming occurred. Triangular points in Fig. 10c indicate subject TM's final nine sessions in Experiment III. Solid lines and dotted curves represent the best fit functions of the fixed sample size and optional stopping versions of the random walk model.

Table 8

RESULTS OF FITTING BOTH VERSIONS OF THE RANDOM WALK MODEL TO THE SPEED-ACCURACY PERFORMANCE OF THE THREE SUBJECTS IN EXPERIMENT V AND SUBJECT EK IN EXPERIMENT IV. SUBJECTS RANKED BY INCREASING SIMILARITY (DIFFICULTY) OF THE STIMULI

Subject	Stimulus Length-Width Ratio	Fixed Sample Size Model				Optional Stopping Model			
		r*	b**	a†	Est. Fixed Delay††	r*	b**	a†	Est. Fixed Delay††
TM	55:53	.853	.283	267	76	.864	.316	286	95
EK	35:34	.950	.517	289	92	.953	.541	302	104
MMcL	50:49	.939	.366	274	78	.986	.375	288	92
WL	63:62	.867	.198	254	66	.873	.226	285	97
EK	85:84	.986	.203	292	94	.969	.213	317	120

*Correlation between mean RT and either $\ln \Omega$ or $\left(\frac{\Omega - 1}{\Omega + 1}\right) \ln \Omega$

**Rate of increase in either $\ln \Omega$ or $\left(\frac{\Omega - 1}{\Omega + 1}\right) \ln \Omega$ per 10-msec. increase in mean RT

†Predicted RT intercept in msec. when $\Omega = 1.0$ ($\ln \Omega = 0$)

††Estimated difference between _a_ and _S_'s mean preprogrammed times in Experiment III (Experiment IV for EK)

the data for each of the three Ss (shown by solid lines and dotted curves in Figs. 10a-10c), together with those obtained for subject EK in Experiment IV. The ordering of Ss in Table 8, from top to bottom, indicates increasing difficulty of the required stimulus discrimination (decreasing length-width ratio). Although the best fits in this experiment for both models were to the data from subject MMcL, correlations between mean RT and the appropriate transformation of Ω for all Ss were higher than .85. Even the poorest fit accounted for more than 72% of the total variance in speed and accuracy over conditions. The optional stopping version accounted for 1%, 1.8% and 9% more of the total variance for these three Ss than did the fixed stopping version--hardly a decisive improvement. Performance of all three Ss was in the range for which the two versions of the random walk model predict almost identical functions.

The estimated rates of increase in evidence for the discrimination per 10-msec. increase in mean RT (*b* in Table 8) were slightly higher for the optional stopping version, as they had been for subject EK in Experiment IV. As for subject EK's data also, the estimated RT intercepts for chance accuracy performance ($\Omega = 1.0$) were 10-20 msec. slower under the assumption of optional stopping. These intercept times resulted in fixed delay estimates of 65-95 msec. longer than each of the three Ss' mean preprogrammed response times in Experiment III. Such fixed delays represent the estimated additional time, over that required for stimulus detection and response execution, necessary before S began to obtain any evidence about the stimulus identity.

<u>RT for correct responses versus errors</u>.--Although both models provide an excellent description of the speed-accuracy tradeoff functions found for mean performance, neither can account for the RT differences between correct responses and errors observed for all three <u>S</u>s. The direction and magnitude of these differences depended on how accurately <u>S</u> performed, as it had for subject EK with the difficult 85:84 stimulus discrimination in Experiment IV. Figure 11 presents mean RT for correct responses and single response errors, plotted against Ω on a logarithmic scale, for grouped sessions having similar error rates within each <u>S</u>. For all three <u>S</u>s, mean RT for single response errors was faster than that of correct responses within their faster and less accurate sessions. This difference reversed for the most accurate and slowest sessions of all three <u>S</u>s. The largest difference in both directions occurred for subject WL, who had the most difficult (63:62) discrimination task of the three. His single response errors had a mean RT of 41 msec. faster than correct responses in the fastest sessions which had an average error rate of .289, but mean error RT was 167 msec. slower than mean correct RT in sessions whose error rates varied around .029.

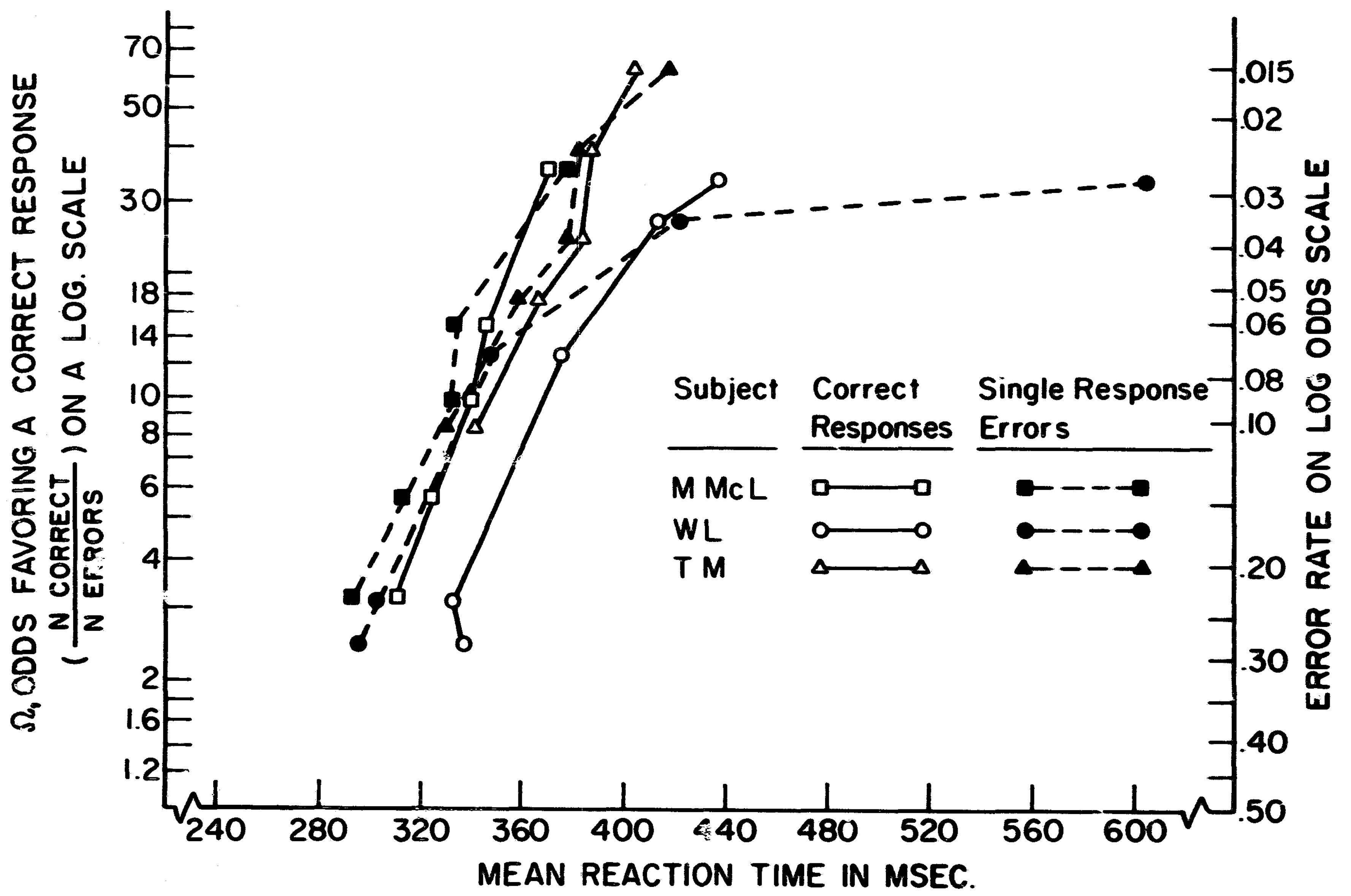

Fig. 11. Speed and accuracy for correct responses and single response errors in grouped sessions for each of the three Ss in Experiment V. These plot mean RT for each type of response against Ω for the pooled sessions on a log scale.

DISCUSSION

The phantom observed.--The final two experiments seem to explain the mysterious disappearance of a speed-accuracy tradeoff under the earlier conditions intentionally designed to produce it. Both experiments demonstrate that human Ss can trade accuracy for increased speed, and Experiment V specifies conditions under which they will do so. The refusal of most (or all) Ss to exhibit a tradeoff in the earlier experiments apparently resulted, not from lack of capability to control the amount of evidence they processed, but from their adaptive adjustment to the particular payoff matrices and time-cost function used.

The use of a preprogrammed response strategy was an option easily discovered and available to all Ss. It permitted a considerable increase in the speed of their reaction times, at the cost of chance error rates. Since Ss' responses under this strategy were independent of the stimulus identity, they appeared to reflect a lower level detection response to the simple presence of the stimulus. The time-cost function used in Experiments I-IV was proportional to Ss' total response times, therefore, any reduction in time increased Ss' net payoffs, provided that it did not involve higher error rates. Extrapolations of all five different speed-accuracy operating characteristic functions in Experiments IV and V displayed the same phenomenon. All functions indicated that Ss' accuracy would drop to chance levels for performance considerably slower than they could obtain by preprogramming their responses.

This fixed cost in time, for any attempt at higher-order processing of evidence about the stimulus identity, meant that Ss could obtain a higher expected payoff for the preprogrammed strategy until the cost in payoff for making an error became fairly large.

The apparent lag time for any higher processing of the stimulus information explains the absence of performances with intermediate error rates for most Ss' later sessions in the earlier experiments. Certainly, the reduction or elimination of the fixed processing cost in Experiment V induced all three Ss to adopt considerably faster and less accurate performances. The Ss' determination to avoid performance with high error rates, together with procedures designed to converge on the payoff matrix which would just induce them to switch from preprogramming to an accuracy strategy, probably accounts for the high stability in RT and error rates for their accuracy performance during the later sessions of Experiments II and III. The control over Ss' performance by costs and payoffs was not complete, however. In Experiment IV, subject RP refused to slow his mean RT to obtain the very low error rates he demonstrated in his initial training, even when the payoff matrix exerted considerable pressure for high accuracy. Also in Experiment IV, subject EK produced fast but inaccurate performances that were dominated in expected payoff by his preprogrammed strategy --although he exhibited considerable amounts of preprogrammed responding during these sessions as well.

Evaluation of tradeoff models.--Both the necessity for and success of using the 250 msec. criterion time to separate out Ss' preprogrammed responses provides a direct verification of one of the Ollman-Yellott guessing model's basic assumptions (Ollman, 1966; Yellott, 1967). Many fast responses in CRT tasks--at least under the conditions used in these experiments--result from fast guesses not controlled by the identity of the stimulus. But the Ollman-Yellott model, like most others, cannot account for a speed-accuracy tradeoff once these preprogrammed guesses have been identified and removed.

The optional stopping and fixed sample size versions of the random walk model do imply such a tradeoff, however, and make strong predictions about the rate of increase in accuracy with increases in mean RT. Increases in mean RT should produce linear increases in $\ln \Omega$ for the fixed sample size version and linear increases in $\left(\frac{\Omega - 1}{\Omega + 1}\right) \ln \Omega$ for the optional stopping version, where Ω is the ratio of the correct response probability to the probability of an error. But Figs. 8 and 10a-10c show that when the two linear parameters of both models were fitted to the data of Experiments IV and V, both provided almost identically good fits--in spite of the conceptual difference between them. The shape of the speed-accuracy operating characteristic function differs substantially between the two versions only for performance with error rates exceeding .33 ($\Omega < 2.00$). Even the poorest of the 10 fitted functions accounted for more than 72% of the variance in S's speed and accuracy performance over sessions. Although goodness of fit comparisons did not, and are not likely to, differentiate between the two versions, they provide impressive evidence for the basic random walk model, particularly in the absence of any competitors.

The two versions do yield markedly different predictions about the effects of asymmetric payoff matrices and unequal stimulus probabilities, which were not investigated in these experiments. As Smith (1968) observed in a recent review of CRT models, since the fixed sample size version assumes a predetermined observation interval, RT cannot depend on the probability or value of particular stimuli. The optional stopping version predicts faster RT for the response appropriate to the more frequent stimulus, and higher error rates for that appropriate to the stimulus with the smaller loss in payoff for an error. Predictions are exact, once the form of the speed-accuracy tradeoff function has been fitted (Edwards, 1965).

Smith (1968) considered the fixed sample size version to be untenable because of evidence that the RT to particular stimuli does depend upon their probability and value. His conclusion, however, rests mainly on data from experiments using spatially separated stimuli. The results of Swensson and Edwards (1968) imply that stimuli coded by their spatial location allow Ss to fixate on a particular location, producing differences in RT merely to detect the presence of the stimulus. Their choice (accuracy strategy) RT differences due to stimulus relative frequency could all be explained by differences in time necessary for stimulus detection (preprogrammed RT). Even with conclusive demonstrations that Ss can vary the amount of evidence processed depending on which response they make, the fixed sample size model still deserves serious consideration as an approximation and for special cases. Its predicted speed-accuracy tradeoff function is simpler and almost identical in form to that predicted by optional stopping, and generalizes easily to more than two stimuli and responses. In addition, experimental procedures may either force Ss to fix their observation intervals (Schouten & Bekker, 1967) or may make such a strategy optimal (Pachella & Pew, 1968).

Parameters of the tradeoff function.--The slope of the fitted tradeoff function, for both versions of the random walk model, estimates the rate at which S accumulated evidence to discriminate the stimuli over successive observations. It reflects how rapidly S's accuracy increased with increases in his mean RT. Theoretically, aside from differences between individual Ss, this slope should decrease as the task of discriminating between the stimuli becomes more difficult. Estimated slopes for both versions of the model were lowest for subjects EK and WL, who had the two most similar pairs of stimuli (length-width ratios of 85:84 and 63:62). The single within-subject comparison for subject EK showed a much slower rate of increase in accuracy for the more similar 85:84 stimuli than for the 35:34 pair.

Theoretical presentations of the random walk model have devoted scant attention to the RT intercept when $\Omega = 1.0$. Probably because of the absence of data to argue otherwise, the tacit assumption has been that this intercept was equal to S's "simple" reaction time (Stone, 1960; Edwards, 1965). Experiments IV and V demonstrate that this is certainly not true, at least for reasonably difficult stimulus discriminations. Both versions of the random walk model estimated the tradeoff function intercept to be considerably slower than S's preprogrammed detection time. Estimates of this fixed delay ranged between 66 and 120 msec. over Ss and both versions of the model, and were 10-30 msec. longer under the optional stopping assumptions.

The absence of a tradeoff in Experiment III can be attributed to the fixed cost resulting from this fixed processing delay combined with the original time-cost function. Identical results with the same procedure for the easily discriminable stimuli in Experiment II argue for the existence of a similar, but shorter, fixed delay in processing any evidence about the identity of these easily discriminable stimuli. Thus, the tradeoff function intercept, the discrimination time for which accuracy drops to chance level, seems to be neither identical to nor predictable from S's simple RT for the same physical stimuli. A plausible, but highly speculative interpretation of intercept differences with identical responses, is that they may reflect differences in the lowest neurological level of organization whose output over time relates systematically to the physical stimulus differences.

What variables affect the slope and intercept parameters of the tradeoff functions? Decreasing discriminability, by increasing the physical similarity between the stimuli, clearly lowered the rate of increase in accuracy over time for subject EK in Experiment IV. Intercept differences, however, were small or absent. Unlike the functions' slopes, they showed no systematic variation between the generally difficult pairs of stimuli in Experiments IV and V.

But since Ss in Experiment II could achieve error rates of .05-.10 with a mean RT only 45-70 msec. longer than their preprogrammed times, any speed-accuracy function for these easily discriminable stimuli (length-width ratio of 30:1) must have both a faster intercept and steeper slope. This evidence, of course, is indirect, and assumes that a speed-accuracy tradeoff could be obtained with the easily discriminable stimuli by using the procedures of Experiment V.

Pew (1968) found that plots of log Ω against RT for means over Ss had similar slopes for both a two-alternative forced reaction time task (Schouten & Bekker, 1967) and a 15-alternative task (Pachella & Pew, 1968). The stimuli in both experiments were easily discriminable and identified by their spatial location. It is tempting to conclude from this that the slope of the speed-accuracy tradeoff function may be unaffected by the number of alternative stimuli, provided that all are well differentiated. However, the 180 msec. tradeoff interval for Schouten and Bekker's (1967) Ss was long compared to the 15-20 msec. difference in mean RT obtained by Swensson and Edwards (1968) for comparable error rates with similar types of stimuli. This makes suspect any inference that depends on the exact slope of the fitted tradeoff function for Schouten and Bekker's data. Since chance accuracy differed considerably between the two and 15-alternative tasks, direct intercept comparisons would have been dubious, but for similar values of Ω (the ratio of all correct responses to errors) Pew (1968) found that the 15-alternative mean RT values were much slower.

Pew (1968) also presented data indicating that practice both increases the slope and decreases the intercept of the tradeoff functions. In Experiment IV, only the slope of subject EK's tradeoff function appeared to change over the 56 sessions on the new, 85:84 stimulus discrimination, but he had already practiced for 64 sessions with similar stimuli and identical responses on the easier 35:34 task. Neither version of the random walk model attempts to account for the effects of stimulus-response compatibility (Fitts, 1959; Smith, 1968). Any such effects, caused by the specific pairing of stimuli and responses, would probably be absorbed primarily by the tradeoff function's intercept. Increased compatibility with practice (Fitts, 1959) may explain some of the decrease in the tradeoff functions' intercepts for Pew's (1968) data.

Modification of the random walk model assumptions.--Both Fitts (1966) and Stone (1960) showed that the optional stopping model predicts identical RT distributions for both correct responses and errors; in particular, their means should be the same. The fixed sample size model makes the same prediction, since the observation interval should be identical for all trials in a particular condition. Although Fitts' (1966) data did not contradict the prediction of equal mean error and correct RT, he noted that an unstable response criterion would result in faster mean error RT for the optional stopping model. A higher proportion of the total errors would occur with faster RT, on trials when S responded with more lax criterion odds. The same type of argument predicts faster mean error RT for the fixed sample size version if S's observation duration, although fixed prior to each trial, varied over trials.

All 12 Ss in Experiments II-V displayed evidence of variations, either systematic or random, in their stopping criteria (cutoff odds or observation duration) from trial to trial within conditions. Each S's single response errors[3] were consistently faster than correct responses, either thoughout all conditions, or for those conditions in which performance was fast and error rates high (in Experiments IV and V). Single response errors were 16-24 msec. faster than correct responses for Ss' accurate performance with the easily discriminable stimuli in Experiment II. With the difficult stimulus discriminations, both the direction and magnitude of this RT difference depended on how accurately S attempted to respond. Figures 9 and 11 showed that the speed of single response errors, compared to that of correct responses, was greatest for Ss' fastest and least accurate performance.

Both versions of the random walk model assume that successive stimulus observations are independent. This assumption implies

[3]By far the larger proportion of errors, particularly with the more difficult discriminations in Experiments III-V, occurred when S made the inappropriate response only (single response errors). Virtually all double response errors were occasions on which Ss made the inappropriate response first. Most double response errors in Experiments III-V occurred in sessions having the highest error rates. Mean RT to the first response, for these double response errors, was consistently slower than RT for single response errors for all 12 Ss in Experiments II-V. This result is mysterious, since the only difference between them was the subsequent occurrence of the correct response within 100 msec. for the double response errors.

that S can accumulate evidence about the stimulus identity at the same rate, indefinitely, permitting his error rates to become arbitrarily small. Several features of the results argue that later observations cannot be regarded as independent for the difficult stimulus discriminations. With the most difficult 85:84 stimulus discrimination, subject EK appeared unable to reduce his error rates under .03. Figure 8 showed that, under extreme pressure to achieve high accuracy, EK increased his mean RT by more than 400 msec. (almost three times the amount of time necessary to decrease his error rates from .30 to .03) with no reduction in his error rate. In the five slowest of these conditions, which had the strongest demands for accuracy, his mean error RT was 775 msec. slower than his mean time for correct responses. Apparently, after some point, additional observations of this difficult stimulus contributed no further to S's discrimination.

All three of the Ss in Experiment V also displayed slower mean RT for errors than correct responses in their most accurate sessions.[4] For subject WL, whose discrimination task (63:62) was nearly as difficult as EK's, mean error RT was 167 msec. slower than correct RT for sessions with an average error rate of .029. Although most recently performed experiments have not separated mean RT for errors and correct responses, several earlier studies reported slower response times associated with errors for difficult stimulus discriminations. Kellogg (1931) reported longer response times when Ss made errors in discriminating which half of a visual field was darker. His five Ss received instructions to concentrate

[4]For all Ss, errors consisted almost entirely of single responses when error rates were less than .10.

only on their accuracy (they did not know responses were timed), but the discrimination tasks were difficult enough to produce error rates of .08-.20. Henmon (1911), for 10 Ss in two experiments who discriminated the longer of two lines (20.3:20.0 mm), found higher proportions of errors for both the fastest and the slowest discrimination times within Ss. These early results, particularly Henmon's (1911), are very similar to those obtained with the difficult discriminations in Experiments IV and V.

For the random walk models, redundant stimulus observations would result in decreasing increments of evidence, and would effectively limit the maximum amount of evidence (the minimum error rate) which could be obtained. This kind of dependence between successive observations could account for both an accuracy ceiling and slower RT for errors. To cope with redundant observations, an optional stopping procedure would set less stringent criterion odds for responding as the number of observations increased (Birdsall & Roberts, 1968; Swets & Birdsall, 1967). This would permit termination of the sampling process when additional observations could not provide evidence sufficient to change S's decision. An increasingly lax criterion would associate less accurate responses with longer observation (reaction) times, producing a slower mean RT for errors. The decrease in criterion odds with increases in observing time would have little or no effect when S operated with a low-accuracy criterion. Because the decrements in incoming evidence from redundant observations would initially be small, criterion odds would decrease only gradually over the first few observations. An optimal adjustment to redundancy might be practically indistinguishable from a fixed criterion if almost all observation sequences terminated early.

It might be possible to modify the fixed sample size version to account for the magnitude and reversal of the difference between mean error and correct RT, as well as the possible accuracy ceiling effect with very difficult discriminations. But the additional assumptions necessary to do this would be numerous and cumbersome. The very lack of flexibility that the fixed sample size version attributes to Ss in terminating their observing process leads to its inflexibility in accounting for the more detailed features of these data.

In spite of their failure to account for some details of the data, both versions of the independent-sample random walk model provided excellent fits to all five of the empirical speed-accuracy tradeoff functions. Their success implies that, although the process is undoubtedly more complicated, observations of the stimulus identity may be treated as sequential, discrete, and independent over a substantial range of speed-accuracy performances. The range in values of mean RT, from predicted intercepts at Ω = 1.0 to those of the most highly accurate performance obtained, was less than 200 msec. for even the most difficult stimuli.

This relatively short maximum tradeoff interval may reflect only lots of practice and the particular type of stimuli used, but suggests a modification of the models as an alternative to consistently redundant observations. There may be some critical duration, similar to that identified by studies on time-intensity reciprocity in stimulus identification tasks (Kahneman & Norman, 1964; Kahneman, Norman & Kubovy, 1967). Within some interval (perhaps up to 200 msec.) the visual system might obtain essentially independent evidence from additional stimulus observations. After this critical duration, which could vary depending on the type of discrimination required, further observations might produce little or no increment in evidence.

REFERENCES

Bernstein, I. H., & Reese, C. Behavioral hypotheses and choice reaction time. Psychon. Sci., 1965, 3, 259-260.

Bernstein, I. H., Schurman, D. L., & Forester, G. Choice reaction time as a function of stimulus uncertainty, response uncertainty and behavioral hypotheses. J. exp. Psychol., 1967, 74, 517-524.

Birdsall, T. G., & Roberts, R. A. Theory of signal detectability: Deferred decision theory. J. acoust. Soc. Amer., 1965, 37, 1064-1074.

Edwards, W. Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. J. math. Psychol., 1965, 2, 312-329.

Fitts, P. M. Information handling in speeded tasks. IBM Research Report RC-109, 1959.

Fitts, P. M. Cognitive aspects of information processing: III. Set for speed versus accuracy. J. exp. Psychol., 1966, 77, 849-857.

Garrett, H. E. A study of the relation of accuracy to speed. Arch. Psychol., 1922, No. 56.

Green, D. M., & Swets, J. A. Signal detection theory and psychophysics. New York: Wiley, 1966.

Henmon, V. A. C. The relation of time of a judgement to its accuracy. Psychol. Rev., 1911, 18, 186-201.

Hick, W. E. On the rate of gain of information. Quart. J. exp. Psychol., 1952, 4, 11-26.

Howell, W. C., & Kreidler, D. L. Information processing under contradictory instructional sets. J. exp. Psychol., 1963, 65, 39-46.

Howell, W. C., & Kreidler, D. L. Instructional sets and subjective criterion levels in a complex information-processing task. J. exp. Psychol., 1964, 68, 612-614.

Kahneman, D., & Norman, J. The time-intensity relation in visual perception as a function of the observer's task. J. exp. Psychol., 1964, 68, 215-220.

Kahneman, D., Norman, J., & Kubovy, M. Critical duration for the resolution of form: Centrally or peripherally determined? J. exp. Psychol., 1967, 73, 323-327.

Kellogg, W. N. The time of judgment in psychometric measures. Amer. J. Psychol., 1931, 43, 65-86.

Leonard, J. A., & Newman, R. C. High speed and high accuracy in a keyboard task. Ergonomics, 1965, 8, 281-304.

Ollman, R. Fast guesses in choice reaction time. Psychon. Sci., 1966, 6, 155-156.

Pachella, R. G., & Pew, R. W. Speed-accuracy tradeoff in reaction times: Effect of discrete criterion times. J. exp. Psychol., 1968, 76, 19-24.

Pew, R. W. Current status of speed-accuracy tradeoff in reaction time. Paper presented at Donders Centennial Conference on Reaction Time. Eindhoven, The Netherlands, August, 1968.

Schouten, J. F., & Bekker, J. A. M. Reaction time and accuracy. In A. F. Sanders (Ed.), Attention and performance, Proceedings of Driebergen Symposium, Acta Psychol., 1967, 27, 143-153.

Smith, E. E. Choice reaction time: An analysis of the major theoretical positions. Psychol. Bull., 1968, 69, 77-110.

Stone, M. Models for choice-reaction time. Psychometrika, 1960, 25, 251-260.

Swensson, R. G., & Edwards, W. Response strategies in a two-choice reaction time task with a linear cost for time. Unpublished manuscript, 1968.

Swets, J. A., & Birdsall, T. G. Deferred decision in human signal detection: A preliminary experiment. Percept. Psychophys., 1967, 2, 15-28.

Yellott, J. I., Jr. Correction for guessing in choice reaction time. Psychon. Sci., 1967, 8.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Michigan, Human Performance Center, Department of Psychology, Ann Arbor, Michigan	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

The Elusive Tradeoff: Speed Versus Accuracy in Choice Reaction Tasks with Continuous Cost for Time.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report, 1968.

5. AUTHOR(S) (First name, middle initial, last name)

Richard G. Swensson

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
December 1968	86 plus viii	25

8a. CONTRACT OR GRANT NO. AF49(638)-1736 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 13 08773-33-T 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Qualified requesters may obtain copies of this report from DDC. Others may obtain copies of this report from Office of Technical Services, Department of Commerce.

11. SUPPLEMENTARY NOTES This report is a Doctoral Dissertation on file in the University of Michigan Library	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency, Assistant Director, Behavioral Sciences, Command and Control Research, monitored by Air Force Office of Scientific Research.

13. ABSTRACT

This study attempted to discover the conditions necessary for a speed-accuracy tradeoff in choice reaction time. The subjects' entire payoff depended on both their accuracy and response time on each trial. Cost for response time was proportional to that time. Session-to-session changes in the payoff difference between a correct response and an error varied the relative attractiveness of speed and accuracy. The first three experiments used both very easy and difficult stimulus discriminations, and varied payoff over a small range in an attempt to obtain fast responses with intermediate (e.g., 25%) error rates. Tradeoffs clearly did not occur for eleven of the twelve subjects; they used only two response strategies throughout. Either subjects responded accurately, with 5% or fewer errors, or they ignored the identity of the stimulus and preprogrammed their responses, accepting chance error rates to decrease their reaction time. Experiments IV and V varied the emphasis on speed versus accuracy over a wide range, and Exp V used a free delay interval that charged subjects only for time exceeding 250msec. After removing preprogrammed guesses, these procedures obtained clear speed-accuracy tradeoffs. All five empirical tradeoff functions with these difficult stimuli reached chance accuracy with response times 65-120 msec slower than each subjects' preprogrammed response time. Two different verions of the random walk model both fit these empirical tradeoff functions but neither could account for the systematic differences between times for correct responses and errors.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification