

# Instructional and probability manipulations of bias in multiletter matching

ROBERT W. PROCTOR and DANIEL J. WEEKS  
Auburn University, Auburn, Alabama

Ratcliff (1985) performed fits of his diffusion model to the results of multiletter-matching experiments conducted by Ratcliff and Hacker (1981) and Proctor, Rao, and Hurst (1984), in which bias to respond "same" or "different" was manipulated by instructions and probabilities, respectively. The fits showed that both bias manipulations affected settings of a goodness-of-match criterion, whereas instructions also affected sensitivity. Evaluations of the experimental procedures and of Ratcliff's model-fitting procedures were performed in the present study. Three experiments showed that instructions and probabilities had similar effects, regardless of whether the *different* pairs were blocked or randomized according to the number of mismatching positions. The most salient feature of the results—that "same" reaction times were traded off more than were "different" reaction times, with no corresponding asymmetry in the error rates—was evident in all situations. The evaluation of Ratcliff's model-fitting procedures indicated that the apparent influence of instructions on sensitivity likely is an artifact of unequal variance for the sets of *same* and *different* pairs. Moreover, the effects of bias can be explained in terms of settings of response criteria, rather than of the goodness-of-match criterion, as in Ratcliff's fits.

Manipulations of bias have been used extensively in psychophysical research since the development of signal-detection theory and methods (Green & Swets, 1966; Swets, 1964). Bias often is manipulated in detection tasks by varying stimulus probabilities across blocks of trials or by holding probabilities constant and varying instructions or payoffs. Measures of the subject's sensitivity to the pertinent sensory information, such as  $d'$  (Green & Swets, 1966), then are computed. These measures of sensitivity, which are considered to be independent of bias, are based on the assumption that bias manipulations influence only the settings of response criteria (the *response-criteria assumption*; e.g., Baird & Noma, 1978). Moreover, instructional and probability manipulations of bias typically have been considered to be qualitatively equivalent procedures for affecting the response criteria (the *equivalency assumption*; e.g., Gescheider, 1985; Green & Swets, 1966).

## Speeded Choice Reactions

Bias also has been manipulated in experiments for which subjects were required to make speeded choice reactions (e.g., Hyman, 1953; Snodgrass, Luce, & Galanter, 1967). As in detection tasks, bias effects customarily have been assumed to conform to the response-criteria and

equivalency assumptions (e.g., Laming, 1968; Link, 1975). Acceptance of these assumptions for speeded choice reactions is illustrated clearly by random-walk models of response latency and accuracy (Laming, 1968; Link, 1975; Stone, 1960). For a two-choice task, such models assume that evidence about the stimulus accumulates gradually toward either of two response criteria by a random-walk process.

The random walk begins from a starting point that is located between the response criteria and drifts probabilistically in discrete steps until one of the two criteria is reached (see the lower half of Figure 1). When the walk reaches the criterion, the corresponding response is executed. The latency of the response is determined by the time to reach the criterion and the accuracy by whether the criterion (and, thus, the response) is appropriate for the stimulus that was presented. Biases for specific responses are modeled by setting the distances between the respective response criteria and the starting point to be asymmetric (e.g., Link, 1975). Moreover, instructions and probabilities typically have been treated as equivalent means of affecting the response criteria (e.g., Link, 1975; Ratcliff, 1978).

Ashby (1983) and Townsend and Ashby (1983) recently questioned both the response-criteria and equivalency assumptions for choice-reaction tasks. In their extensive review of stochastic models of elementary psychological processes, Townsend and Ashby noted that little evidence has been obtained that is relevant to either of the two assumptions and concluded that "there seems to be no good reason, at this point at least, to expect the same sorts of bias to result from such different experimental manipulations" (p. 314). Townsend and Ashby suggested further that whereas instructions may affect response crite-

---

This research was supported in part by Grant AFOSR-88-0002 from the Air Force Office of Scientific Research. We would like to thank Mei-Shio Jang for programming the experiments; Janet Sugg for preparing the figures; Teri Woolard, Carl Honeycutt, Derek Bagwell, and Julie Gaut for assistance in collecting the data; and Bart Farell, Lester Krueger, and Roger Ratcliff for helpful comments on an earlier draft. Reprint requests should be sent to Robert W. Proctor, Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907.

ria, stimulus probabilities may affect the accumulation of information on which the decision is based.

### Multiletter Matching

One specific type of choice-reaction task in which the effects of bias manipulations have been examined is multiletter matching. In multiletter-matching tasks, the latency and accuracy with which subjects classify pairs of letter strings as "same" or "different" are measured. In the most common form of the task, a pair is classified correctly as "same" if the two strings contain the same letters in the same positions and "different" if the strings contain mismatching letters at one or more positions (e.g., Bamber, 1969).

In the matching task, both instructions (Ratcliff & Hacker, 1981) and probabilities (Proctor & Healy, 1987; Proctor, Rao, & Hurst, 1984) have been shown to effectively manipulate the relative bias to respond "same" or "different." When subjects are biased by either manipulation to respond "same," rather than "different," correct "same" responses are relatively faster and false "same" errors (incorrectly responding "same" to *different* pairs) are relatively more numerous. "Different" responses show opposite patterns of results, with correct "different" responses being relatively slower and false "different" errors (incorrectly responding "different" to *same* pairs) being relatively less numerous. These effects of bias manipulations are not limited to multiletter pairs, having also been found for "same"-"different" responses to pairs of single letters (Krueger & Shapiro, 1981; Proctor & Rao, 1983), pairs of multidimensional forms (Downing, 1971), and pairs of tones (Coltheart & Curthoys, 1968). Consistent with both the response-criteria and the equivalency assumptions, the effects of the instructional and probability manipulations of bias in multiletter matching typically have been attributed to the relative settings of response criteria (e.g., Proctor & Rao, 1982; Proctor et al., 1984; Ratcliff & Hacker, 1981).

Recently, however, Ratcliff (1985, 1987) proposed an account that does not conform to the response-criteria and equivalency assumptions. This account is based on his diffusion model (Ratcliff, 1981), which combines a continuous random-walk decision process with a comparison process based on signal-detection theory (see Figure 1). The comparison process determines the rate at which information accumulates toward the respective response criteria in the decision process. For the comparison process, *same* and *different* pairs are assumed to result in overlapping, normal distributions on a goodness-of-match continuum. The relative rates of accumulation of information in the decision process are determined by (1) the separation between the means of the *same* and *different* distributions on the goodness-of-match continuum (i.e., the sensitivity, as indicated by asymptotic  $d'$ ) and (2) the goodness-of-match (or comparison) criterion. This criterion determines whether a momentary value on the continuum drives the decision process toward either the "same" or the "different" response criterion. For ex-

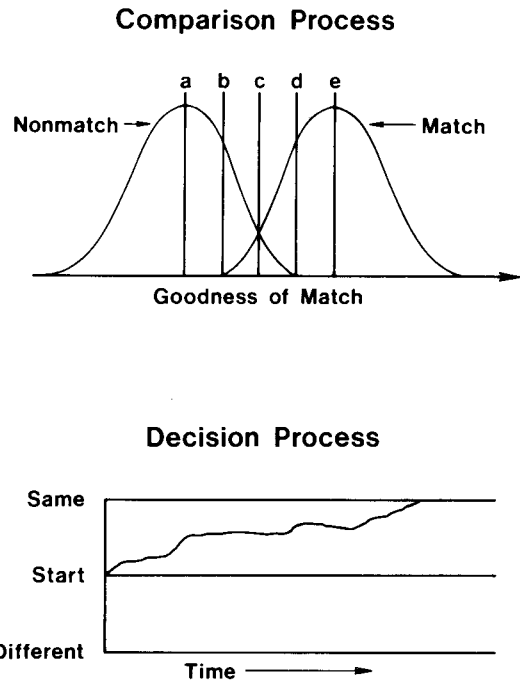


Figure 1. Depictions of the signal-detection comparison process and the random-walk decision process in Ratcliff's (1981, 1985) diffusion model. (The comparison process shows multiple criteria, a-e, and the decision process shows an example accumulation of evidence to the "same" response criterion.)

ample, as the comparison criterion is shifted from a to e in Figure 1, information in the decision process builds up more rapidly toward the "different" response criterion than toward the "same" response criterion.

Ratcliff (1985) demonstrated that the diffusion model could produce close fits to the instructional-bias data obtained by Ratcliff and Hacker (1981) and to the probability-bias data obtained by Proctor et al. (1984). These fits were accomplished by allowing the sensitivity and criterion of the comparison process, as well as the response criteria of the decision process, to vary. As a consequence, the specific parameter settings for the fits deviated from both the response-criteria and the equivalency assumptions. First, the probability and instructional manipulations exerted their primary effects on the comparison criterion, rather than on the response criteria. Second, the instructional manipulation of Ratcliff and Hacker, but not the probability manipulation of Proctor et al., also affected the sensitivity of the comparison process.

Thus, although the specific patterns generated by Ratcliff (1985) differ from those proposed by Townsend and Ashby (1983), the fits of his diffusion model provide some suggestive evidence that bias manipulations may affect the accumulation of information on which a decision is based and that instructions and probabilities may have distinct effects. However, the experimental procedures used by Ratcliff and Hacker (1981) and by Proctor et al. (1984)

differed in several ways in addition to the type of bias manipulation. Furthermore, the specific model-fitting procedures used by Ratcliff may have been problematic. The purpose of the present study was to systematically evaluate the experimental and model-fitting procedures, with the intent of determining whether instructional- and probability-bias manipulations in multiletter matching have qualitatively similar effects and whether these effects can be explained in terms of response criteria.

## EVALUATION OF EXPERIMENTAL PROCEDURES

Ratcliff's (1985) fits of the diffusion model to the instructional-bias data of Ratcliff and Hacker (1981) showed  $d'$  to vary with bias, whereas his fits of the model to the probability-bias data of Proctor et al. (1984) did not. For the instructional manipulation, sensitivity was greater when the bias was to respond "same" than when it was to respond "different." Ratcliff interpreted this effect of the instructional manipulation on  $d'$  as reflecting systematically better performance when subjects were biased to respond "same." This better performance was evident in that reaction times (RTs) were faster overall when the bias was to respond "same" rather than "different," with similar error rates for the two bias conditions. In turn, the overall RT advantage for the "same" bias was due to the fact that "same" RTs benefited relatively more from this bias than "different" RTs benefited from a "different" bias. Thus, in terms of Ratcliff's interpretation, the feature of Ratcliff and Hacker's results that is reflected in the  $d'$  shift is the asymmetric tradeoff for "same" and "different" RTs.

Three experiments were conducted in the present study to determine the situations that produce the asymmetric tradeoff between "same" and "different" RTs. Experiment 1 employed an instructional manipulation of bias similar to that used by Ratcliff and Hacker (1981), and Experiment 2 employed a probability manipulation similar to that used by Proctor et al. (1984). The instructional manipulation of Experiment 1 produced a greater tradeoff than did the probability manipulation of Experiment 2. Because Ratcliff (1985) suggested that a similar difference in the strength of the bias manipulations could account for differences in the results obtained by Ratcliff and Hacker and by Proctor et al., a third experiment was conducted that used a more extreme probability manipulation than that of Experiment 2.

Ratcliff and Hacker (1981) and Proctor et al. (1984) both used four-letter strings, but the *different* pairs were much more heterogeneous in Ratcliff and Hacker's experiment; that is, the *different* pairs mismatched at either 1, 2, 3, or 4 positions in Ratcliff and Hacker's experiment, whereas they mismatched at only 1 position in Proctor et al.'s experiment. Thus, heterogeneity of difference could be the cause of the asymmetric tradeoff for "same" and "different" RTs. That heterogeneity could cause the asymmetric tradeoff is suggested by the external noise principle of Krueger's (1978, 1979, 1986) noisy-operator

theory. According to this principle, increasing the heterogeneity of the *different* pairs should slow "same" RTs relative to "different" RTs and also possibly increase false "same" errors relative to false "different" errors. Moreover, when the *different* pairs are heterogeneous, shifts in the "same" and "different" response criteria that produce approximately symmetric changes in the rates of false "same" and false "different" errors seemingly should produce a greater change in "same" RTs than in "different" RTs. Each of the present experiments thus included a manipulation of heterogeneity of difference to determine whether it causes the asymmetric RT tradeoff.

## Method

### Subjects

Forty students at Auburn University participated as subjects. Ten subjects were tested in Experiment 1, 10 in Experiment 2, and 20 in Experiment 3. Four additional subjects were omitted from Experiment 1, 2 for excessive RT tradeoffs of approximately 1 sec and 2 for having error rates that exceeded 20%. One additional subject was omitted from Experiment 2 for having a substantial number of responses (>20%) below 200 msec. Each subject in Experiments 1 and 2 was tested in five sessions, whereas each subject in Experiment 3 was tested in four sessions.

### Designs

The designs for Experiments 1 and 2 were entirely within subjects, with the following independent variables: pair type (*same* or *different*), bias ("same" or "different"), number of positions at which the two strings in the *different* pairs mismatched (1, 2, 3, or 4), and method of presentation (whether all *different* pairs mismatched at the same number of positions—*blocked presentation*—or whether the number of mismatching positions varied randomly—*random presentation*). In each session, a subject received either random presentation or blocked presentation involving one of the four degrees of mismatch; one set of trials was conducted with a "same" bias, and the other with a "different" bias. The experiments differed only in that bias was manipulated by instructions in Experiment 1, but by the relative probabilities of *same* and *different* pairs in Experiment 2.

Experiment 3 was similar to Experiment 2 in that both used a probability manipulation of bias, but the probabilities were more extreme in Experiment 3. Because the extreme probabilities produced relatively few trials in a set for the infrequent pair, the blocked/random variable was manipulated between subjects in Experiment 3 to enable collection of sufficient data. Otherwise, the design was similar to those of Experiments 1 and 2.

### Apparatus

The experiments were conducted on a Radio Shack TRS-80 Model 4 microcomputer, with members of the standard alphabetic character set presented in white on the dark background of the CRT display screen. Responses were made by pressing one of two response keys on the computer's keyboard (a standard typewriter keyboard). Half of the subjects in each experiment responded "same" by pressing the Z key with the left index finger and "different" by pressing the ? key with the right index finger. For the other half of the subjects, the left finger was used for "different" and the right finger for "same."

### Stimuli

Stimuli were pairs of four-letter strings composed from the uppercase consonants, excluding Y. Within a given string, no letters were repeated. Pairs of strings were either *same*, with the two strings containing the same letters in the same positions, or *different*, with

the strings mismatching at 1, 2, 3, or 4 positions. For the *different* pairs, the second string was generated by replacing the appropriate number of letters in the first string with new letters that were different from each other and from any in the first string.

The subjects viewed the display screen from a distance of approximately 50 cm. Individual letters subtended a visual angle of approximately  $.57^\circ$  in height and  $.34^\circ$  in width. Letter strings were displayed horizontally, with spaces of approximately  $.06^\circ$  between each letter, thus subtending a visual angle of  $1.54^\circ$ . The pair of strings on each trial was centered about a pair of adjacent fixation asterisks, with the first string presented  $.68^\circ$  above the asterisks and the second presented  $.68^\circ$  below the asterisks.

For Experiment 1, 10 orders of trials were constructed. Each order included a set of 16 practice trials (8 *same* and 8 *different* pairs) and 96 test trials (48 *same* and 48 *different* pairs). The orders were constructed by first specifying two four-letter strings, with different letters in each, for each trial number. These root strings were prepared by a program that used the computer's random number generator, and they were used to generate the pairs actually seen by the subject. A distinct random ordering of *same* and *different* classifications then was specified, with the constraint that an equal number of *same* and *different* trials occurred in each half of the practice and test sets. For *same* trials, the subject saw only the first of the two root strings, presented twice. For *different* trials, the first of the two root strings was presented initially to the subject. The second string seen by the subject was constructed by taking the letters from the corresponding positions of the first string for all matching positions and the letters from the corresponding positions of the second root string for the mismatching positions. For example, if the root strings were DXBN and RTFC, and the trial involved a *different* pair that mismatched at only the last position, the subject would see DXBN and DXBC.

With this method, each order could be used in any session, regardless of whether the number of mismatching positions was held constant at a value of 1 to 4 (blocked presentation) or varied randomly (random presentation). Thus, across subjects, each order could be used approximately equally often for each of the 10 conditions created by the five methods of presentation (random, or blocked by 1, 2, 3, or 4 mismatching positions) and bias ("same" vs. "different"). When presentation was random, 12 of the 48 *different* pairs mismatched at 1, 2, 3, or 4 positions. When presentation was blocked, all 48 *different* pairs mismatched at the same number of positions. For each order, regardless of whether presentation was random or blocked, the number of *different* pairs that mismatched at a particular number of positions were divided approximately equally among all of the possible specific positions of difference.

For Experiment 2, each set of 16 practice trials and 96 test trials included either 75% *same* pairs (12 practice and 72 test trials) and 25% *different* pairs (4 practice and 24 test trials) or vice versa. The specific orders were constructed by modifying those used in Experiment 1; for the 75%-*same* orders, half of the *different* pairs from that experiment were changed to *same* pairs, whereas for the 25%-*same* orders, half of the *same* pairs were changed to *different* pairs.

For Experiment 3, the number of pairs in each set of 96 test trials was changed to 88 of the predominant type and 8 of the lesser type, and the 16 practice trials were changed to 14 of the predominant type and 2 of the other. These changes were accomplished by altering the appropriate number of pairs from the orders used in Experiment 2.

### Procedures

In all experiments, a trial consisted of the following sequence of events: First, the fixation asterisks were presented for 500 msec, followed by a 500-msec blank interval. Then, the string above fixation was presented for 500 msec, followed after an interval of

500 msec by the string below fixation. The second string remained in view until the subject responded. One second later, the next trial began. RTs were recorded from the onset of the second string. RTs shorter than 200 msec or longer than 1,500 msec (less than 1% of the responses) were omitted from analysis.

In Experiment 1, each subject was tested in five sessions, each of which included two sets of trials (16 practice and 96 test trials). In each session, the subject was biased to respond "same" for one set, by being instructed to respond "same" as quickly as possible, without simply guessing, and to respond "different" only when sure. For the other set, the subject was biased to respond "different," by being instructed to respond "different" quickly and to respond "same" only when sure. For a given subject, the order of the bias conditions was the same for each session. Across subjects, the two possible orders of the bias conditions were used equally often.

The five sessions of Experiment 1 differed in the specific types of *different* pairs that were presented, with the same type used for both sets in a session. For one session, the number of mismatching positions for *different* pairs varied randomly. For each of the other four sessions, the number of mismatching positions was held constant at 1, 2, 3, or 4 positions. The specific sessions in which the random and the four blocked conditions were tested were counter-balanced for order, according to a Latin square.

The procedure for Experiment 2 was similar to that of Experiment 1, with the exception that, prior to each set, subjects were told the proportions of *same* and *different* pairs that would occur, but were not told explicitly to adopt a bias toward responding either "same" or "different." This procedure also was used for Experiment 3, with the only change being that each subject participated in only four sessions, receiving either random presentation in all sessions or blocked presentation in all sessions.

In all experiments, prior to the first test session, each subject received training on five sets of 20 trials (10 *same* and 10 *different* pairs), for which the instructions were to respond "same" or "different" as rapidly and accurately as possible. For the first four sets, the number of positions at which the *different* pairs mismatched was held constant at 4, 3, 2, and 1, in that order. In the fifth set, the number of mismatching positions for the *different* pairs varied randomly. The training sets were intended to familiarize the subjects with the task and with the presentation conditions for the *different* pairs that would occur in the test sessions.

## Results

For each experiment, mean correct RTs and proportions of errors were obtained for each subject, as a function of pair type (*same*, or *different*), method of presentation (blocked or random), bias ("same" or "different"), and number of mismatching positions (1, 2, 3, or 4; this variable was not applicable for *same* pairs when presentation was random). For one set of analyses, the data for the *different* pairs and for the *same* pairs in the blocked condition were collapsed across the number of mismatching positions, whereas the number of mismatching positions was considered as a factor in a second set of analyses.

### Same Versus Different

Analyses of variance (ANOVAs) for each experiment showed the following effects for pair type (see Figure 2).<sup>1</sup> First, in all experiments, the bias  $\times$  pair type interaction was significant both for RTs [ $F(1,9) \geq 24.7$ ,  $ps < .001$ ] and for errors [ $F(1,9) \geq 9.47$ ,  $ps < .02$ ], indicating that

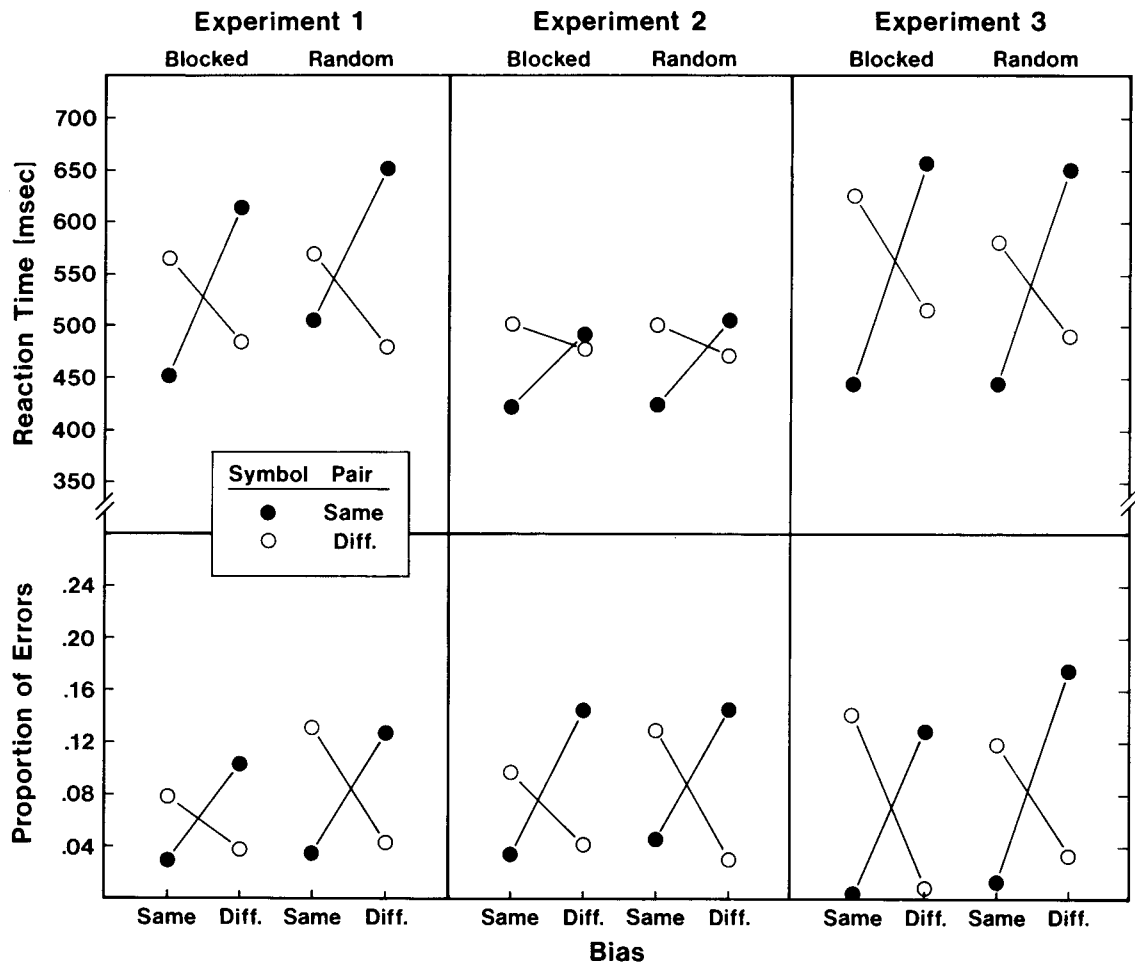


Figure 2. Mean reaction times and proportions of errors in Experiments 1, 2, and 3, as a function of pair type, bias, and method of presentation.

the bias manipulations were effective. Responses to *same* pairs were faster and more accurate when a “same” bias was in effect, with the reverse relation holding for responses to *different* pairs.

Second, the main effect of bias was significant for RTs in all three experiments [ $F(1,9) \geq 5.87, ps < .05$ ], indicating that responses were faster overall when the bias was to respond “same” than when it was to respond “different.” This outcome was due to the tradeoffs being of greater magnitude for “same” RTs than for “different” RTs (see the upper row of Figure 2). The error data did not show a comparable main effect of bias in any experiment ( $F_s < 1.0$ ), indicating that the tradeoff for errors was approximately symmetric. Thus, the tradeoff patterns for all three experiments are similar to those of Ratcliff and Hacker (1981) in being asymmetric for RTs but symmetric for errors.

Third, no terms involving method of presentation were significant for RTs in any of the experiments. Most importantly, method of presentation had neither a main effect nor an interaction with bias condition ( $F_s < 1.0$ ),

indicating that the asymmetry in tradeoff for “same” and “different” RTs was of similar magnitude for blocked and random presentation. In other words, heterogeneity of difference did not affect the asymmetric-tradeoff pattern.

Although method of presentation did not significantly affect RTs, nonsignificant trends were present for the method  $\times$  pair type interaction in Experiment 1 [ $F(1,9) = 4.15, p = .07$ ] and in Experiment 3 [ $F(1,18) = 2.98, p = .10$ ]. These trends reflect a tendency for “same” RTs to be slowed relative to “different” RTs, by 45 and 36 msec in Experiments 1 and 3, respectively, when presentation was random as opposed to when it was blocked. A similar tendency of 11 msec also was present in Experiment 2, but did not approach significance. Moreover, even when the RT data for Experiments 1 and 2 were pooled, a comparison of “same”-“different” disparity scores for the random and blocked conditions showed only a nonsignificant trend [ $t(19) = 1.93, p > .05$ ]. The tendency for “same” RTs to increase relative to “different” RTs when number of mismatching po-

sitions is random is consistent with Krueger's (1978, 1979) external noise principle.

Method of presentation also affected the proportion of errors in Experiment 1: errors were more numerous when presentation was random than when it was blocked [ $F(1,9) = 20.9, p < .01$ ]. Moreover, the method  $\times$  bias  $\times$  pair type interaction showed a nonsignificant trend in that experiment [ $F(1,9) = 3.92, p = .08$ ]. This trend was due to the fact that proportions of errors were similar for blocked and random presentation when a "different" bias was in effect, but were considerably lower for blocked presentation when a "same" bias was in effect. Subjects tended to trade off on incorrect responses to *same* pairs to an approximately equivalent extent, regardless of whether presentation was random or blocked. However, the error tradeoff for *different* pairs tended to be less when the number of mismatching positions was blocked than when it was random; that is, when subjects were biased to respond "same," they made more false "same" responses to *different* pairs if the number of mismatching positions was unknown than if it was known.

A similar pattern of errors was evident in Experiment 2, although only the method  $\times$  bias interaction approached significance [ $F(1,9) = 3.51, p = .09$ ]. The method  $\times$  bias interaction also showed a nonsignificant trend in Experiment 3 [ $F(1,18) = 3.03, p = .10$ ], but the pattern was opposite to that of Experiments 1 and 2. The tendency in Experiments 1 and 2 for a relative increase in false "same" errors when number of mismatching positions was random also is consistent with Krueger's (1978, 1979) external noise principle.<sup>2</sup>

### Number of Mismatching Positions

For each experiment, two different ANOVAs were performed with number of mismatching positions as a factor. First, only the *different* pairs were analyzed, with blocked versus random presentation also included as a factor. Second, for blocked presentation, responses to both *same* and *different* pairs were analyzed, as a function of the number of mismatching positions for the *different* pairs in the block.

**Responses to *different* pairs.** In all experiments, RTs and proportions of errors for the *different* pairs decreased as the number of mismatching positions increased (see Tables 1 and 2) [ $F_s(3,27) \geq 13.5$  and  $8.78, p_s < .001$ , respectively]. Responses also were faster when subjects were biased to respond "different" than when they were biased to respond "same" [ $F_s(1,9) \geq 14.0, p_s < .01$ ] in both Experiments 1 and 3, but not in Experiment 2. In addition, the responses were less accurate with the "same" bias in Experiments 2 and 3 [ $F_s(1,9) \geq 11.9, p_s < .01$ ], with a similar nonsignificant tendency apparent in Experiment 1 [ $F(1,9) = 3.50, p = .09$ ].

The method  $\times$  number of mismatches interaction was significant for errors in Experiments 1 and 3 [ $F_s(3,27) \geq 3.63, p_s < .03$ ], as was the method main effect in Experiment 1 [ $F(1,9) = 6.79, p < .03$ ]. The method main

**Table 1**  
Mean Reaction Times (in msec) for Experiments 1, 2, and 3,  
by Pair Type, Bias, Presentation Method, and  
Number of Mismatching Positions

Bias and Pair Type	Blocked				Random			
	1	2	3	4	1	2	3	4
Experiment 1								
"Same" Bias								
<i>Same</i>	512	478	422	390	—	—	—	—
<i>Different</i>	631	574	566	492	644	553	548	536
"Different" Bias								
<i>Same</i>	694	643	590	527	—	—	—	—
<i>Different</i>	549	493	457	442	519	485	476	433
Experiment 2								
"Same" Bias								
<i>Same</i>	477	419	400	395	—	—	—	—
<i>Different</i>	573	508	471	453	560	506	468	463
"Different" Bias								
<i>Same</i>	579	503	439	445	—	—	—	—
<i>Different</i>	565	496	421	432	525	469	442	450
Experiment 3								
"Same" Bias								
<i>Same</i>	471	461	445	410	—	—	—	—
<i>Different</i>	709	634	616	539	625	590	557	553
"Different" Bias								
<i>Same</i>	740	623	636	605	—	—	—	—
<i>Different</i>	640	502	487	452	568	492	454	447

effect indicates again that more false "same" errors were made in Experiment 1 when the number of mismatching positions was random than when it was blocked. The interaction shows that the increase in false "same" errors with random presentation occurred primarily for the strings that differed at only one position. In Experiment 3, the bias  $\times$  number of mismatches interaction also was significant [ $F(3,54) = 5.76, p < .01$ ], with the effect of number of mismatches on the error rate being greater when subjects were biased to respond "same" rather than "different." However, this interaction apparently was due to a floor effect on accuracy when subjects were biased to respond "different." Thus, except that the functions were less affected by bias in Experiment 2 and were noisier in Experiment 3, the patterns of errors were similar across the three experiments.

**Responses to *same* and *different* pairs with blocked presentation.** Both RTs and errors decreased with increases in the number of mismatching positions in all experiments [ $F_s(3,27) \geq 7.97$  and  $4.43, p_s < .02$ , respectively]. Also, the bias main effect was significant for RTs in all experiments [ $F_s(1,9) \geq 7.49, p_s < .03$ ] and for errors in Experiments 1 and 2 [ $F_s(1,9) \geq 6.64, p < .03$ ]. The bias  $\times$  pair type interaction was significant for RTs [ $F_s(1,9) \geq 23.2, p_s < .001$ ] and errors [ $F_s(1,9) \geq 10.2, p_s < .02$ ] in all three experiments. These terms show only the effect of bias and the asymmetry for responses to *same* and *different* pairs shown previously.

Two additional terms were significant for RTs in Experiment 2: the pair type main effect [ $F(1,9) = 29.4$ ,

**Table 2**  
**Mean Proportions of Errors for Experiments 1, 2, and 3,**  
**by Pair Type, Bias, Presentation Method, and**  
**Number of Mismatching Positions**

Bias and Pair Type	Blocked				Random			
	1	2	3	4	1	2	3	4
Experiment 1								
“Same” Bias								
<i>Same</i>	.042	.033	.023	.019	—	—	—	—
<i>Different</i>	.144	.081	.050	.042	.242	.100	.083	.058
“Different” Bias								
<i>Same</i>	.121	.094	.113	.088	—	—	—	—
<i>Different</i>	.073	.029	.017	.031	.142	.050	.025	.000
Experiment 2								
“Same” Bias								
<i>Same</i>	.061	.025	.026	.024	—	—	—	—
<i>Different</i>	.138	.088	.088	.079	.267	.150	.033	.017
“Different” Bias								
<i>Same</i>	.204	.117	.143	.129	—	—	—	—
<i>Different</i>	.076	.036	.032	.026	.078	.017	.006	.022
Experiment 3								
“Same” Bias								
<i>Same</i>	.007	.005	.002	.001	—	—	—	—
<i>Different</i>	.225	.088	.150	.100	.288	.113	.063	.013
“Different” Bias								
<i>Same</i>	.212	.112	.050	.125	—	—	—	—
<i>Different</i>	.020	.006	.002	.001	.113	.001	.007	.003

$p < .001$ ] and the bias  $\times$  number of mismatches interaction [ $F(3,27) = 4.03, p < .02$ ]. As in the analysis that collapsed across number of mismatches, the former term reflects faster “same” responses than “different” responses. The latter term reflects slightly different functions when the “same” bias was in effect than when the “different” bias was. In Experiment 3, errors showed an interaction of bias  $\times$  pair type  $\times$  number of positions [ $F(3,27) = 4.31, p < .02$ ]. This interaction reflects the opposing, deviant points for the *same* pairs, with a “different” bias, and the *different* pairs, with a “same” bias, for blocks in which the strings mismatched at three positions. This pattern likely is not meaningful since it was not apparent in Experiments 1 and 2.

**Discussion**

The primary results of the experiments can be summarized briefly. Responses were faster overall when subjects were biased to respond “same” than when they were biased to respond “different,” and there was little difference in the corresponding proportions of errors. The overall faster RTs with a “same” bias were a function of subjects’ trading off more for the “same” RTs than for the “different” RTs. In general, the RTs and proportions of errors varied as a decreasing function of the number of mismatching positions, with these functions affected little by the bias manipulations. Most importantly, the asymmetrical pattern of RT tradeoffs occurred regardless of whether bias was manipulated by instructions (Experiment 1) or by probabilities (Experiments 2 and 3).

Although the tradeoffs produced by the 75%/25% probability manipulation were not as substantial as those produced by the instruction manipulation, the tradeoffs were at least as great when the percentage of the predominant pair type was increased to approximately 90%.

Comparison of the results obtained when number of mismatching positions was blocked with those obtained when it was random indicates that the asymmetric RT tradeoff was equally apparent in both situations. Thus, the tradeoff is not a function of heterogeneity of difference. Although blocked versus random presentation did not influence the asymmetric tradeoff, it tended to affect the RT and errors in other ways generally consistent with Krueger’s (1978, 1979) external noise principle: For the more heterogeneous situation (i.e., random presentation), correct “same” RTs tended to be slower than correct “different” RTs. Additionally, when subjects were biased to respond “same,” relatively more false “same” errors tended to be made in the more heterogeneous situation.

**EVALUATION OF MODEL-FITTING PROCEDURES**

The fits of Ratcliff’s (1985) model to Ratcliff and Hacker’s (1981) data showed  $d'$  to be greater when subjects were biased to respond “same” than when they were biased to respond “different.” According to Ratcliff’s interpretation, this difference in  $d'$  reflects the asymmetric tradeoff for “same” and “different” RTs. Experiments 1-3 indicate that neither heterogeneity of difference nor type of bias manipulation is crucial to the asymmetric tradeoff pattern. “Same” RTs are affected more by bias manipulations than are “different” RTs, regardless of whether bias is induced by instructions or by probabilities and of whether the number of mismatching positions is blocked or random. These findings seemingly conflict with Ratcliff’s fits, in that the variables examined in the present study are those that most obviously distinguished Ratcliff and Hacker’s experiment, which showed an effect of bias on  $d'$ , from Proctor et al.’s (1984) experiment, which did not. However, the apparent discrepancy may reside in Ratcliff’s model-fitting procedures, rather than in the data.

**Does  $d'$  Reflect the Reaction-Time Asymmetry?**

The apparent conflict between the results of the present study and the fits performed by Ratcliff is resolved by examining Table 3, which presents the mean RTs and proportions of errors to *same* and *different* pairs for the experiments of Ratcliff and Hacker (1981) and Proctor et al. (1984). The important point to note is that the asymmetric tradeoff for RTs was present not only in the experiment of Ratcliff and Hacker, but also in that of Proctor et al. Moreover, the magnitude of the asymmetry actually was greater in the latter experiment than in the former. Because the fits of Ratcliff’s model to the data

**Table 3**  
**Mean Reaction Times (in msec) and Proportions of Errors (PE)**  
**for the Multiletter-Matching Experiments Fit by Ratcliff (1985)**

Pair	"Same" Bias		"Different" Bias	
	RT	PE	RT	PE
Ratcliff and Hacker (1981)				
Same	472	.033	573	.109
Different	582	.138	516	.075
M	527	.085	545	.092
Proctor, Rao, and Hurst (1984)				
Same	559	.071	697	.166
Different	674	.204	690	.123
M	617	.138	694	.145

Note—The tabled conditions for Proctor, Rao, and Hurst (1984) are the two most extreme conditions from their experiments, with 80% of the pairs being the predominant type and 20% the lesser type.

of Proctor et al. did not show a systematic effect of bias on  $d'$ , whereas the fits to the data of Ratcliff and Hacker did, the  $d'$  shift apparently does not reflect the asymmetric RT tradeoff.

If the  $d'$  measure does not reflect primarily the asymmetric RT tradeoff, the question remains as to why the fits of Ratcliff's (1985) model to Ratcliff and Hacker's (1981) data showed an effect of bias on  $d'$ . The answer apparently lies in the fact that  $d'$  is a parametric measure of sensitivity that is based on the assumptions that the underlying distributions are normal and of equal variance (Green & Swets, 1966; McNicol, 1972). Violations of these assumptions can cause  $d'$  to vary artifactually as bias is manipulated. The artifactual effect of bias on  $d'$  can be sizable, as illustrated by an example from McNicol (1972, p. 87) in which the signal and noise distributions are normal, but of unequal variance. In this example, the standard deviation of the signal distribution is only 1.5 times that of the noise distribution, yet  $d'$  varies artifactually from 0.17 at a low criterion setting to 1.50 at a high setting.

Although heterogeneity of difference did not have much influence on performance in Experiments 1–3, it can affect derived measures, such as  $d'$ , that are based on the assumption of equal variance; that is, when *different* pairs mismatch at 1, 2, 3, or 4 positions, as in Ratcliff and Hacker's (1981) experiment, the equal-variance assumption is violated because the *different* distribution is considerably more heterogeneous than the *same* distribution. Regarding the *same* distribution as signal and the *different* distribution as noise, the heterogeneous situation corresponds to one in which the standard deviation of the noise distribution is greater than that of the signal distribution. In such situations,  $d'$  will be greater when the criterion is low (i.e., shifted toward the noise distribution) than when it is high (i.e., shifted toward the signal distribution; Green & Swets, 1966; McNicol, 1972), which is the pattern that Ratcliff (1985) obtained when fitting Ratcliff and Hacker's data.

That the violation of the equal-variance assumption could have produced the  $d'$  shift in the fits to Ratcliff and Hacker's (1981) data is suggested by a comparison of  $d'$  to another sensitivity measure,  $A'$  (Gescheider, 1985), which is not based on the assumption of equal variance. When calculated from the mean error data for each of the bias conditions in Ratcliff and Hacker's experiment (see Table 3), the  $d'$  values are 2.93 for the "same" bias and 2.67 for the "different" bias. However, when  $A'$  is calculated from the same data, the values are .95 for both bias conditions. Thus, of the two sensitivity measures, only the one that assumes equal variance shows a difference between the "same" and "different" bias conditions.

Ratcliff (1985) did not treat the *different* pairs as a single distribution when calculating the asymptotic  $d'$  values for Ratcliff and Hacker's (1981) experiment. Rather, he attempted to remedy the problem of unequal variance by computing four different  $d'$  values for each bias condition; that is, he treated the *different* pairs that mismatched at 1, 2, 3, and 4 positions as distinct normal distributions, each with a variance equal to that of the *same* distribution. An initial  $d'$  value was determined for each bias condition based on responses to the *same* pairs and to the *different* pairs that mismatched at one position. These initial fits fixed the distance between the mean for the *same* distribution and the goodness-of-match criterion for the remaining three  $d'$  values that were computed for each bias condition. These remaining values were based on responses to the *different* pairs that mismatched at 2, 3, and 4 positions, respectively, by varying only the distance between the criterion and the means of the respective *different* distributions.

Although this procedure of treating the *different* pairs as four distinct normal distributions may seem to eliminate the unequal-variance artifact from the  $d'$  measures, it does not. The reason why the artifact is not eliminated is that the subject's task is to discriminate the *same* pairs from the entire set of *different* pairs, not from a select subset. Thus, the  $d'$  calculations inappropriately compare responses for the entire set of *same* pairs to responses for only a subset of the *different* pairs.

The procedure has important consequences for the relative  $d'$  values computed when the bias is to respond "same" as opposed to when the bias is to respond "different." The complete decrease in performance for responses to *same* pairs as bias shifts from "same" to "different" is reflected in the relative  $d'$  values. For example, the complete change in the proportion of incorrect responses to *same* pairs from .033 to .109 in Ratcliff and Hacker's (1981) experiment is incorporated into each pair of  $d'$  values. However, for the corresponding increase in performance that occurs when responding to *different* pairs, only the portion of the overall change attributable to the particular subset is reflected in the  $d'$  values. For example, for the subset of *different* pairs that mismatched at



four positions in Ratcliff and Hacker's experiment, the  $d'$  values for the "same" bias and "different" bias conditions were based on the error rates of .031 and .029, respectively, rather than on the overall false "same" error rates of .138 and .075.

Because the complete decrease in performance for *same* pairs as bias shifts from "same" to "different" is offset by only a portion of the increase for *different* pairs, the  $d'$  measure is constrained artifactually to be less when the bias is to respond "different." Moreover, since the effect of bias on performance is less for the easy *different* pairs, the artifactual disparity in the  $d'$  values increases as the number of mismatching positions increases. This pattern corresponds to the  $d'$  values generated by Ratcliff (1985) for Ratcliff and Hacker's (1981) data. Thus, the effect of instructions on sensitivity is less likely to be "in the data" (Ratcliff, 1985, p. 220; 1987, p. 278) than in the procedures used to fit the data.

### What Parameter Reflects the Reaction-Time Asymmetry?

Instead of the RT asymmetry's being reflected primarily in the  $d'$  values, the asymmetry seems to be reflected in the goodness-of-match criterion of Ratcliff's (1985) model. As shown in Table 4, for the experiments fit by Ratcliff, the goodness-of-match criterion was set with a more extreme bias when subjects were biased to respond "same" than when they were biased to respond "different." Consistent with the RT data, the asymmetry in criterion settings results in an overall bias toward "same" when the settings are averaged across bias conditions.

In Ratcliff's model, then, the RT asymmetry occurs because subjects adopt asymmetric settings of the goodness-of-match criterion. However, no rationale is provided for why these particular settings of the criterion would be adopted (see Proctor, 1986). Asymmetric settings favoring "different" could be adopted just as easily that would

produce a greater tradeoff for "different" RTs than for "same" RTs. The fact that "same" RTs consistently show a greater tradeoff than "different" RTs suggests a more basic cause than arbitrary settings of a criterion.

### Some Alternative Accounts

The larger tradeoff for "same" than for "different" RTs can be explained in ways other than through the settings of a goodness-of-match criterion. For example, the asymmetric tradeoff could be due to information about sameness building up sooner and accumulating more gradually than information about difference. If so, subjects would have considerable latitude in the criterion that could be adopted for concluding that a pair is "same," but not in the criterion for concluding that a pair is "different." A more gradual buildup of information indicating sameness is a feature of two models.

According to Eriksen and Schultz's (1979) continuous flow model, information builds up gradually in the perceptual system when a stimulus is presented, with coarser features apparent before details. Eriksen, O'Hara, and Eriksen (1982) proposed that, as a consequence, the similarities between two stimuli that are being compared become evident early in processing, whereas differences become apparent only relatively late. Thus, the "same" response receives continuous, gradual priming, and the "different" response is activated more discretely. According to this sameness-primacy interpretation, the asymmetric tradeoff for RTs reflects a fundamental characteristic of the perceptual system that allows sameness to predominate early in information processing.

The second model attributes the asymmetry to the conjunctive decision rule used in the present experiments; that is, detection of a mismatch at only one position is sufficient for a pair to be classified with certainty as "different." However, detection of a match at one position does not indicate that the correct response is "same." Rather, four matches are necessary for a "same" response to be made with certainty. Thus, the conjunctive decision rule apparently constrains the range of states of information indicating sameness to be greater than the range indicating difference. As a consequence, bias manipulations could influence the amount of partial information that a person requires for responding "same" more than they could the amount of partial information required for responding "different." A decision-rule explanation of this type was proposed initially by Downing (1971) to explain the asymmetric tradeoff for pairs of multidimensional forms. It also has been suggested for multiletter matching by Taylor (1976) and by Proctor et al. (1984).

Whether the sameness-primacy account or the decision-rule account ultimately is correct is secondary to the point that the accounts share the following characteristic. Both accounts attribute the effects of bias manipulations, which yield the asymmetric tradeoff for RTs, to settings of response criteria. Thus, even though the asymmetry shows up in the goodness-of-match criterion of Ratcliff's (1985)

Table 4  
Distances Between the Goodness-of-Match Criterion and the Means of the Match (U) and Nonmatch (V) Distributions in Ratcliff's (1985) Fits

Bias	Model Parameter		Difference
	U	V	U-V
Ratcliff and Hacker (1981)			
"Same"	.380	.264	.116
"Different"	.270	.315	-.045
<i>M</i>	.325	.290	.035
Proctor, Rao, and Hurst (1984)			
"Same"	.260	.170	.090
"Different"	.180	.230	-.050
<i>M</i>	.220	.200	.020

Note—The difference score is an indicator of bias in the setting of the goodness-of-match criterion, with a positive value indicating a "same" bias and a negative value indicating a "different" bias. The bias conditions from Proctor, Rao, and Hurst (1984) are their two most extreme conditions.

model, the primary effects of both the instructional and probability manipulations of bias may be on response criteria.

### GENERAL DISCUSSION

Manipulations of bias by instructions (or payoffs) and probabilities typically have been considered to be equivalent procedures (the equivalency assumption) for influencing response criteria (the response-criteria assumption). However, Townsend and Ashby (1983) noted that little evidence exists for speeded choice tasks that is pertinent to these assumptions. They suggested that whereas instructions may affect response criteria, probabilities may affect the rate at which information builds toward the response criteria.

Recent fits by Ratcliff (1985) of his diffusion model to data from multiletter-matching experiments provided evidence consistent with Townsend and Ashby's (1983) general point that the effects of instructional and probability manipulations of bias may differ. In fits of his model to instructional data obtained by Ratcliff and Hacker (1981) and to probability data obtained by Proctor et al. (1984), Ratcliff found that both probabilities and instructions primarily affected a goodness-of-match criterion and that instructions also affected sensitivity. Because these deviations from the response-criteria and equivalency assumptions are potentially important, the experimental procedures of the previous studies and the model-fitting procedures used by Ratcliff were evaluated systematically in the present study.

Three experiments were conducted, one of which used an instructional manipulation of bias similar to that of Ratcliff and Hacker (1981) and two of which used probability manipulations similar to that of Proctor et al. (1984). All experiments varied the amount of heterogeneity of difference (i.e., whether the number of mismatching positions for *different* pairs was blocked or random). Neither heterogeneity of difference nor type of bias manipulation had much effect on the patterns of results that were obtained. The most salient feature of the results—that “same” RTs were traded off more than were “different” RTs—was apparent in all situations. Thus, the experimental results are consistent with the equivalency assumption that instructions and probabilities are similar procedures for manipulating bias.

The evaluation of Ratcliff's (1985) model-fitting procedures suggested that heterogeneity of difference for the *different* pairs in Ratcliff and Hacker's (1981) study was treated inadequately. The apparent effect of instructions on sensitivity likely is an artifact of unequal variances for the *same* and *different* distributions. Moreover, the asymmetric tradeoff for “same” and “different” RTs, which is reflected in the goodness-of-match criterion of Ratcliff's model, can be explained by alternative models for which the only criteria are response criteria. Thus, it remains

likely that the primary effects of the bias manipulations are on response criteria.

In summary, the evaluations of experimental and model-fitting procedures performed in the present study provide little reason to reject either the equivalency assumption or the response-criteria assumption for multiletter matching. Instructional and probability manipulations of bias have qualitatively similar effects that can be explained in terms of the relative settings of “same” and “different” response criteria.

### REFERENCES

- ASHBY, F. G. (1983). A biased random walk model for two choice reaction times. *Journal of Mathematical Psychology*, *27*, 277-297.
- BAIRD, J. C., & NOMA, E. (1978). *Fundamentals of scaling and psychophysics*. New York: Wiley.
- BAMBER, D. (1969). Reaction times and error rates for “same-different” judgments of multidimensional stimuli. *Perception & Psychophysics*, *6*, 169-174.
- COLTHEART, M., & CURTHOYS, I. (1968). Short-term recognition memory for pitch: Effect of a priori probability on response times and error rates. *Perception & Psychophysics*, *4*, 85-89.
- DOWNING, B. D. (1971). Response probabilities and “same”-“different” reaction times. *Perception & Psychophysics*, *9*, 213-215.
- ERIKSEN, C. W., O'HARA, W. P., & ERIKSEN, B. A. (1982). Response competition effects in *same-different* judgments. *Perception & Psychophysics*, *32*, 261-270.
- ERIKSEN, C. W., & SCHULTZ, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, *25*, 249-263.
- GESCHIEDER, G. A. (1985). *Psychophysics: Method, theory, and application* (2nd ed.). Hillsdale, NJ: Erlbaum.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- HYMAN, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *45*, 188-196.
- KRUEGER, L. E. (1978). A theory of perceptual matching. *Psychological Review*, *85*, 278-304.
- KRUEGER, L. E. (1979). A model of unidimensional perceptual matching. *Journal of Experimental Psychology: Human Perception & Performance*, *5*, 277-288.
- KRUEGER, L. E. (1986). Positive effect of heterogeneity of difference on the *same-different* disparity in letter matching. *Perception & Psychophysics*, *39*, 117-122.
- KRUEGER, L. E., & SHAPIRO, R. G. (1981). Intertrial effects of *same-different* judgments. *Quarterly Journal of Experimental Psychology*, *33A*, 241-265.
- LAMING, D. R. J. (1968). *Information theory of choice-reaction times*. New York: Academic Press.
- LINK, S. W. (1975). The relative judgment theory of two choice response time. *Journal of Mathematical Psychology*, *12*, 114-135.
- McNICOL, D. (1972). *A primer of signal detection theory*. London: Allen & Unwin.
- PROCTOR, R. W. (1986). Response bias, criteria settings, and the fast-same phenomenon: A reply to Ratcliff. *Psychological Review*, *93*, 473-477.
- PROCTOR, R. W., & HEALY, A. F. (1987). Task-specific serial position effects in multiletter matching. *Perception & Psychophysics*, *42*, 180-194.
- PROCTOR, R. W., & RAO, K. V. (1982). On the “misguided” use of reaction-time differences: A discussion of Ratcliff and Hacker (1981). *Perception & Psychophysics*, *31*, 601-602.
- PROCTOR, R. W., & RAO, K. V. (1983). Evidence that the *same-different* disparity in letter matching is not attributable to response bias. *Perception & Psychophysics*, *34*, 72-76.

- PROCTOR, R. W., RAO, K. V., & HURST, P. W. (1984). An examination of response bias in multiletter matching. *Perception & Psychophysics*, *35*, 464-476.
- RATCLIFF, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.
- RATCLIFF, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*, 552-572.
- RATCLIFF, R. (1985). Theoretical interpretations of speed and accuracy of positive and negative responses. *Psychological Review*, *92*, 212-225.
- RATCLIFF, R. (1987). More on the speed and accuracy of positive and negative responses. *Psychological Review*, *94*, 277-280.
- RATCLIFF, R., & HACKER, M. J. (1981). Speed and accuracy of same and different responses in perceptual matching. *Perception & Psychophysics*, *30*, 303-307.
- SNODGRASS, J. G., LUCE, R. D., & GALANTER, E. (1967). Some experiments on simple and choice reaction time. *Journal of Experimental Psychology*, *75*, 1-17.
- STONE, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*, 251-260.
- SWETS, J. A. (1964). *Signal detection and recognition by human observers*. New York: Wiley.
- TAYLOR, D. A. (1976). Effect of identity in the multiletter matching

task. *Journal of Experimental Psychology: Human Perception & Performance*, *2*, 417-428.

TOWNSEND, J. T., & ASHBY, F. G. (1983). *Stochastic modeling of elementary psychological processes*. New York: Cambridge University Press.

#### NOTES

1. Except for *F* ratios that are designated explicitly as being from Experiment 3, the reported degrees of freedom for the error term are correct only for Experiments 1 and 2.

2. It should be noted that in neither Experiment 1 nor Experiment 2 was the method  $\times$  pair type interaction significant. Moreover, a comparison of "same"-"different" disparity scores for errors on the combined data of Experiments 1 and 2 showed no significant difference between blocked and random presentation [ $t(19) = 1.05$ ]. Thus, the error data are, at best, weakly consistent with predictions of the external noise principle.

(Manuscript received September 15, 1987;  
revision accepted for publication July 26, 1988.)