Psychological Review
1985, Vol. 92, No. 2, 212-225

0033-295X/85/$00.75

Theoretical Interpretations of the Speed and Accuracy of Positive and Negative Responses

Roger Ratcliff
Northwestern University

The interpretation of the reaction time difference between positive and negative responses in two-choice matching tasks has been the subject of recent controversy. Proctor and his colleagues hold that the difference represents a difference in processing between same and different judgments, whereas Ratcliff and Hacker argued that the difference can be accounted for in terms of criteria settings. In this article, the way in which several models of choice reaction time and matching can account for this reaction time difference is examined and one particular model, the diffusion model of Ratcliff (1981), is fitted to the data from three published experiments. The results of these fits provide a clear interpretation of the reaction time difference in terms of criteria settings. It is concluded that interpretation of such positive–negative reaction time differences in the absence of a specific model is hazardous at best.

In two-choice reaction time tasks, subjects sometimes show faster positive responses than negative responses and sometimes faster negative responses than positive responses. These differences between positive and negative latencies have been of central concern in the development of recent models for a number of tasks (see Ratcliff & Hacker, 1981). In research on matching processes, interpretation of the difference between positive and negative latencies has been the subject of extended theoretical controversy (Bamber, 1969; Krueger, 1978; Proctor, 1981; Ratcliff & Hacker, 1981; Taylor, 1976). In essence, the issue is whether this difference demonstrates the existence of a separate process or difference in processing between *same* and *different* judgments or merely demonstrates the way criteria are set in decision processes. The point of the present article is to show that models incorporating decision processes powerful enough to accomodate the wide range of relationships among various aspects of data, including reaction time, accuracy, and reaction time distributions, are also capable of accounting for positive–negative differences in terms of criterion settings and do not require an extra process or processing stage.

In the first section of this article, models that attempt to explain both reaction time and accuracy data are noted, and their ability to deal with yes/no bias effects and performance on the matching task is considered. In the second section, several specific models developed for the matching task are considered with respect to their ability to relate reaction time and accuracy. In the third section, the model of Ratcliff (1981), which was designed specifically to account for reaction time and accuracy in the matching task, is applied to the data from three sets of experiments, including two from Proctor and his associates' most recent work (Proctor & Rao, 1983a; Proctor, Rao, & Hurst 1984, Experiment 1). Finally, generalizations from these applications of the model are discussed.

In a typical matching task, two strings of letters are presented sequentially to the subject, and the subject is asked to judge whether they are identical. The result usually found is that *same,* or *positive,* responses are faster than *different,* or *negative,* responses; this result has been labeled the fast *same* effect. Bamber (1969), Proctor (1981), and Taylor (1976) (among others) have argued that the fast *same* effect reflects a special process or a

This research was supported by Grants BNS-8203061 and HD-18812.

I would like to thank Gail McKoon, Reid Hastie, Lester Krueger, Duncan Luce, Ben Murdock, Robert Proctor, and an anonymous reviewer for their comments on this manuscript.

Requests for reprints should be sent to Roger Ratcliff, Psychology Department, Northwestern University, Evanston, Illinois 60201.

processing difference that gives an advantage to *same* responses. However, Ratcliff and Hacker (1981) rejected the notion of special processing, and claimed that the fast *same* effect is simply the result of the way subjects set criteria in decision processes. To support this claim, Ratcliff and Hacker showed that the fast *same* effect could be eliminated (and replaced by a fast *different* effect) by introducing difficult negative trials or by adjusting the bias of the subjects to be cautious on *same* trials. The elimination of the fast *same* effect would seem to argue against the existence of a special stage of processing. But, Proctor and Rao (1982) replied that a careful analysis of the data presented by Ratcliff and Hacker showed that even with the effects of bias there was still a fast *same* effect because the reaction time bias to *same* responses was greater than for *different* responses. Ratcliff and Hacker (1982) responded that error rates must be taken into account, and when this is done, it appears that fast *same* responses are associated with more false *same* responses, suggesting an explanation based on shifting criteria.

The critical issue in this argument concerns the relationship between *accuracy* and reaction time for *same* and *different* responses. The typical (but not universal) result is that the fast *same* effect is accompanied by a greater proportion of false *different* responses (responding *different* in error; Krueger, 1978). Most recently, Proctor and Rao (1983a) and Proctor et al. (1984) manipulated the relative probability of *same* and *different* trials and showed that, when error rates for *same* and *different* responses are equal (interpolating between probability conditions), there is still a fast *same* effect. It has been claimed by Krueger (1978), Proctor et al. (1984), and Proctor and Rao (1982, 1983a) that the effect of bias on accuracy and reaction time is simple: In order for subjects to be fast on *same* responses, they have to sacrifice accuracy by making more false *same* responses (responding *same* in error). So, for a response bias account to be sufficient, in any set of data the fast *same* effect would have to be accompanied by more false *same* responses than false *different* responses. They then argue that the observed pattern of more false *different* responses than false *same* responses rules out models that try to account for the fast *same* effect in terms of criteria settings.

The most important problem with this argument is that there are few, if any, models in the reaction time literature capable of dealing with reaction time, reaction time distributions, and the relationship between accuracy and reaction time that would require such a simple interpretation of the relationship between positive and negative reaction times and accuracies. In fact, Ratcliff and Hacker (1982) noted specifically that the model of Ratcliff (1981) was capable of producing the fast *same* effect at the same time as *more* false *different* responses than false *same* responses. Because this model, as well as other classes of models, have not been addressed in the recent work by Proctor and his colleagues noted above, these issues are taken up here.

Models of Reaction Time and Accuracy

In order to examine the relationship between positive and negative responses in reaction time, it is necessary to examine both reaction time and accuracy and to consider models that account for both reaction time and accuracy. The models are described here briefly and many of their properties are not mentioned, but their ability to deal with positive–negative reaction time differences is considered. The review that is most pertinent is the one by Audley (1973). He evaluated models of choice reaction time with respect to empirical findings dealing with sequential effects, probability manipulations, and reaction time–accuracy relationships, all of which relate to bias between positive and negative responses. The conclusion he reached was that there are two main classes of models able to provide a coherent account of the data: the mixture models and the sequential sampling models. In the time since Audley's review, sequential sampling models of some types have been elaborated and there has been some development of mixture models, but there are no substantially new classes of models.

In the class of mixture models, there are two main exemplars: the preparation model (Falmagne, 1965; Falmagne, Cohen, & Dwivedi, 1975) and the fast-guess model (Ollman, 1966; Yellott, 1967, 1971). Falmagne's (1965) model was designed to account for sequential effects and bias effects by assuming that a response could either be in a prepared state or an unprepared state. Reaction time for a prepared response is assumed to be faster than for an unprepared response. Predictions for sequential effects in both reaction time and accuracy are made by assuming a probability mixture of prepared and unprepared responses over trials. Thus, if the previous response is the same as the present response, it is more likely to be in a prepared state and so have a faster response (e.g., Falmagne et al., 1975).

The fast-guess model is a two-state model that assumes a stimulus-controlled response state and a fast-guess state. The fast-guess response is assumed to have shorter reaction time than does the stimulus-controlled response, and this allows speed–accuracy trade-off to be modeled; in conditions that stress speed, subjects produce more fast guesses, leading to faster but less accurate responses. This model leads to some precise and elegant tests involving linear relationships between terms involving accuracy and reaction time. However, these linear relatinships are not always satisfied (see Green & Luce, 1973).

Both of the initial versions of these models were found to be inadequate for complementary reasons. The fast-guess model could not account for sequential effects, and the preparation model could not account for speed–accuracy trade-offs. Revisions of these models produced a common model with both preparation and fast-guess states (Falmagne et al., 1975, Model II; Yellott, 1971, FG2).

There are two main problems in applying mixture models to results from matching tasks. First, there is little evidence for fast guesses in the sense of Swensson (1972), that is, reaction times around 200 ms, and chance accuracy. The fast-guess model would have to assume fast guesses as slow as 500 ms (Ratcliff, 1981, Figure 4) with high accuracy, which does not seem to be the kind of data for which the model was originally designed. Second, there is no mechanism to account for stimulus-controlled processes. In the mixture models, it is assumed that stimulus-controlled processes have constant reaction time and accuracy, but clearly for the matching task, this would not be sufficient to account for similarity effects, serial position effects, order effects, and so on. In order to advance mixture models, additional mechanisms would be needed.

The second class of models is the sequential sampling models; random walk, counter, and runs models. In these models, information is assumed to be accumulated gradually, and reaction time and accuracy are related through a noisy evidence-accumulation process. The notion of bias was important in early versions of these sequential sampling models; for example, Stone (1960), in the first application of random walk models to reaction time, presented expressions that relate reaction time means and variances, error rates, and frequency of alternatives. Since the initial work of Stone, there have been a number of other random walk models presented and applied in the reaction time literature by, for example, Laming (1968; 1979a, 1979b), Link (1975), and Link and Heath (1975). These models have been well developed to the point of providing some exact quantitative and distribution-free tests. However, they are not without their problems (Laming, 1979a; Link & Heath, 1975; Luce, in press). In addition to these random walk models, there are a number of other stochastic models: the counter model (Pike, 1973; Pike, Dalgliesh, & Wright, 1977), the runs model (Audley, 1960), and the accumulator model (Vickers, 1979).

Among the sequential sampling models, the random walk (diffusion) model of Ratcliff (1981) was specifically designed to account for data from the sequential letter-matching task. This model was chosen to be applied to data from the matching task, but it should be noted that modified versions of the other sequential sample models could account for these in similar ways.

The diffusion model (Ratcliff, 1981) uses a continuous version of the random walk process—the diffusion process—to represent the process of comparing a test item to items in memory. More specifically, the diffusion model assumes that evidence toward a *same* or *different* response is accumulated gradually. When either an upper criterion (*same* response) or a lower criterion (*different* response) is reached, a response is produced.

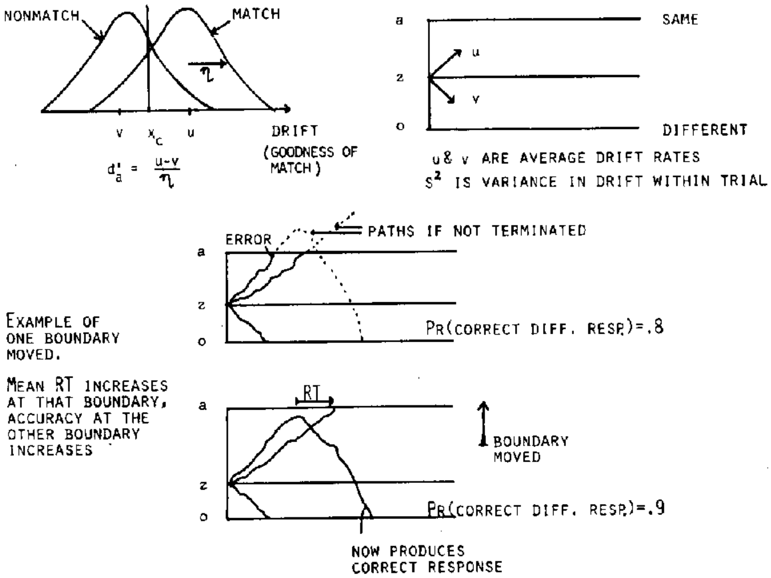

Figure 1. Illustration of the diffusion model. (The distributions represent goodness of match for matching and nonmatching comparisons over trials. Goodness-of-match translates into drift rate in the diffusion comparison process, and u and v represent the means of the match and nonmatch drift rate distributions over trials. The bottom two panels provide an example of the change in reaction time and accuracy as one boundary is moved to illustrate speed–accuracy relationships.)

The rate at which the process proceeds toward either the *same* boundary or the *different* boundary is determined by the goodness-of-match: The more the test stimulus matches the studied stimulus, the faster the process approaches the *same* boundary, whereas the less the two strings match, the faster the process approaches the *different* boundary. In the model presented by Ratcliff (1981), the goodness-of-match for a particular condition is derived from a memory model that assumes that encoded items are distributed over position. The amount of overlap between the study and test strings is used to determine the goodness-of-match.

There are three criteria that can be adjusted by the subject (see Figure 1). The first two are the boundaries of the random walk: Increasing the position of the upper boundary means that more evidence is required for a *same* response so that reaction times will be slower and accuracy of *different* responses will increase. Similarly, if the *different* boundary position is increased, more evidence will be required for a *different* response and reaction time will increase and accuracy for *same* responses will increase (see the bottom two panels in Figure 1).

The third criterion is the zero point of the rate of accumulation of evidence. It is assumed that across trials there is variability in the goodness-of-match (signal and noise distributions in signal detection theory correspond to goodness-of-match values for match and nonmatch comparisons, respectively). The model assumes that goodness-of-match is mapped into drift in the diffusion process by setting a zero point of drift (corresponding to the criterion in signal detection theory) to correspond to some point between the match and nonmatch distributions. Values of goodness-of-match that are greater than the criterion correspond to positive average drift, and values less than the criterion correspond to negative average drift (see the top left panel of Figure 1).

It is important here to differentiate between the goodness-of-match criterion that is free to be set and adjusted by the subject and the parameters of the model that are allowed to vary in fitting the model to the data. In the model, for a fixed d', only the criterion is free to be adjusted by the subject. However, in fitting the model to data, u and v, the mean drift rate for matching and nonmatching comparisons respectively, are allowed to vary (u for *same* responses and v for *different* responses). This is equivalent to allowing both d' and the criterion to vary, in signal detection terms. The relationship between the usual signal detection parameters, d' and x_c (the criterion position) is $d' = (u - v)/\eta$, and the criterion position, scaled from the nonmatch distribution, is $x_c = -v/\eta$, where v is usually a negative number. Thus, in experiments in which d' is assumed to be constant across instruction or probability conditions, $u - v$ should be approximately constant (i.e., constant $d' = (u - v)/\eta$).

It is the three criteria (the two diffusion process boundaries and the goodness-of-match criterion) that can be adjusted by the subject, and as will be shown, changes in these criteria are capable for accounting for bias effects (as well as sequential and speed–accuracy effects) in the matching task.

The two classes of models, sequential sampling models and mixture models, have the major advantage that they were devised initially to deal with the relationship between reaction time and accuracy in choice reaction time. There have been a number of other models specifically designed to account for data in the matching task (see Nickerson, 1972, for a review of the earlier models). In general, these earlier models do not do a good job of relating accuracy and reaction time. The models were discussed in Ratcliff and Hacker (1981) with respect to their ability to deal with bias effects, and a brief discussion is presented here for completeness.

Models for Matching

There are four models of the matching process (besides Ratcliff's 1981 model) that will be considered here. The models proposed by Bamber (1969) and Taylor (1976) are exemplars of one kind of model, the two-process model. These models account for *same* responses by assuming a holistic matching process and account for *different* responses by assuming analytic processing (item-by-item comparison, serial for Bamber, and parallel for Taylor). The holistic process is assumed to be faster than the comparison process, to give the fast *same* effect. The model proposed by Proctor (1981) provides a similar explanation. The comparison process is the same for positive and negative items and the fast *same* effect results from an encoding process that is faster for identicial strings. It is difficult to evaluate Proctor's model because it has not been presented in enough detail to allow quantitative comparisons between the model and reaction time, and accuracy data.

The model proposed by Krueger (1978, 1979) is a close relative of the class of sequential sampling models and so is related to the model proposed by Ratcliff (1981). The main difference is that Krueger's model is discrete in processing, whereas Ratcliff's model is continuous. Krueger's model assumes that a letter is composed of a fixed number of features. Comparison of study and test strings of letters is carried out by a feature matching process in discrete passes (200 ms duration per pass) in which all features are compared on every pass (see Krueger, 1978; Krueger & Shapiro, 1981; and Proctor & Rao, 1983b, for further discussion). These assumptions (and some others) have been criticized by Proctor and Rao (1983b), but with modification the sequential sampling nature of the comparison process is still defensible.

In contrast to Ratcliff's model, in Krueger's model, the pattern of data in which there are both fast *same* responses and a greater proportion of false *different* responses than false *same* responses, is not accounted for by the manner in which subjects set criteria in the comparison process. Instead, this pattern comes about from positive skew in the distributions of feature matches; this forces more passes for *different* responses. However, it should be noted that in the fits in Krueger (1978) for multielement strings, there are some systematic deviations between the model and the data. Alternatively (or in combination), it is possible that Krueger's model could be better fit to the data if the two criteria for positive and negative responses were made nonsymmetric (as in Krueger, 1979); then the fast *same* effect could be partly accounted for in terms of criteria settings, as in Ratcliff's (1981) model. Despite the differences, Krueger's model and Ratcliff's model are similar in aim and structure, particularly in the attempt to deal with the relationship between reaction time, accuracy, and reaction time distributions.

In addition to these models, Link (1978) and Vickers (1979) have applied their sequential sampling models to some aspects of *same/different* judgments, and these applications have much in common with the approaches of Krueger and Ratcliff. Link (1978) provided some relationships between accuracy and reaction time and applied these to data from a deadline experiment involving matching lines of different lengths. Vickers (1979) argued that the *same–different* disparity is simply a matter of bias, and whether the model makes similar predictions to those of the diffusion model is simply a matter of fitting the model to data.

All of these models (except Ratcliff's) have the limitation that there is no mechanism for the interaction of letters in transposition *different* conditions (e.g., test strings that have adjacent test letters interchanged). The models as they stand would predict that performance would be the same as for conditions in which the transposed letters were replaced by new letters; in fact, these conditions are slower and less accurate. The only model that has been specifically developed to deal with these transposition effects is Ratcliff's (1981), and fits to the transposition data will be noted along with the new fits presented later.

To summarize with respect to bias effects, models have been proposed by Bamber, Taylor, and Proctor that are either mute on the subject of bias effects or claim that, whatever the bias, there should always be a fast *same* effect reflecting a special processing stage or difference between *same* and *different* processing. In opposition, the models proposed by Ratcliff and Krueger have no special processing mechanism; instead the fast *same* effect and bias effects in general arise from the settings of criteria in a single-decision process or the shape of the feature match distributions. In the next section, it is shown that Ratcliff's model can account for *same–different* differences in reaction time and accuracy for several experiments in terms of criteria settings. It should be stressed that the way this model accounts for the experimental effects would be similar to several of the other sequential sampling models if they were modified to apply to data from the matching task.

Fitting the Diffusion Model to Bias Data

Ratcliff and Hacker (1982) noted that the diffusion model was capable of accounting for fast *same* responses at the same time that there are more false *different* errors than false *same* errors. The model can do this because there are three independent criteria to be set by the subject, the two boundary positions in the diffusion process and the zero point of drift. This point was not addressed by Proctor and Rao (1982, 1983a) and Proctor et al. (1984), so in this article the point is demonstrated by fitting the diffusion model to data. There are three sets of data to be fitted, those of Ratcliff and Hacker (1981, Experiment 1), Proctor et al. (1984), and Proctor and Rao (1983a); in addition, the results of fits of the model to other data in Ratcliff (1981) will be considered.

The diffusion model can be considered an integration of signal detection theory and the random walk (or diffusion) process. As would be done in signal detection theory, it is assumed that the goodness-of-match between the study and test strings for a particular condition varies over trials, leading to distributions of goodness-of-match for each of the various conditions. Values from these distributions are mapped into drift rate in the diffusion model. The parameters of the model are the diffusion process boundary positions—z the starting point and a the match boundary position (zero being the nonmatch boundary position, see Figure 1)—and the mean match drift rate, u, and the mean nonmatch drift rate, v. There is also an encoding and response time parameter, T_{er}, and two variance parameters—variance in the relatedness distributions, η^2, and variance in the drift of the diffusion process, s^2. The variance parameters are fixed at their values in Ratcliff (1981), and the encoding and response parameter is fixed in each experiment. In applying the model to data, mean reaction time and accuracy scores are fitted by adjusting the parameters representing the boundary positions and the drift rates for the various conditions.

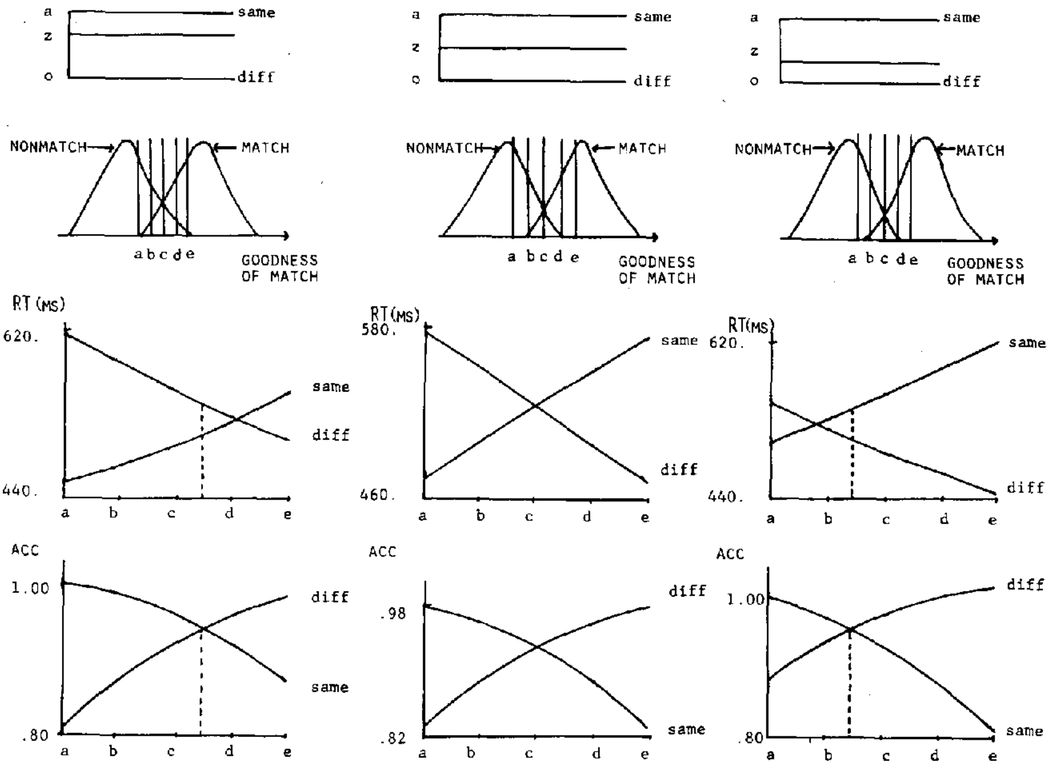

Figure 2. Reaction time and accuracy as a function of boundary position in the diffusion process (columns) and as a function of criterion placement in the signal detection analysis (*a–e*). (In these examples, T_{er} was set at 300 ms and a at 0.12; z varied from 0.07, to 0.06 to 0.05 across columns, v from −0.4 to −0.2, and u from 0.2 to 0.4, both v and u in 0.05 steps. The other parameters were set at the same values as in fits to the other paradigms.)

For further details of fitting, see the appendix in Ratcliff (1981) but note that only mean reaction times are fitted here (by integrating $tf(t)$, derived from the appendix in Ratcliff, 1978) rather than reaction time distributions.

Predictions of the Diffusion Model

Illustrations in Figure 2 demonstrate the ways in which changes in the diffusion-model parameters produce changes in reaction time and accuracy. The left column of Figure 2 shows the case in which the starting point of the diffusion process is nearer the *same* boundary than the *different* boundary. The five data points (*a* through *e*) represent cases in which the criterion in the goodness-of-match analysis varies from nearer the non-match distribution to nearer the match distribution. When the goodness-of-match criterion is just to the left of *d*, *same* responses are faster than *different* responses and there are more false *same* errors. The result that is of most note here is that as the goodness-of-match criterion moves toward the match distribution, the error rates become equal before the fast *same* advantage is erased. Thus it is possible to have equal error rates and fast *same* responses or even more false *different* responses at the same time as fast *same* responses. The greater the asymmetry between the diffusion process boundaries, the greater is this effect. It should also be noted that a fast *different* effect can occur as a mirror image of the fast *same* when the starting point of the diffusion process is nearer the *different* boundary than the *same* boundary; this is shown in the right column.

One specific conclusion to draw from these examples is that, if the diffusion process

Table 1
Parameter Values and Fits of the Diffusion Model to the Data From Ratcliff and Hacker (1981)

Trial	Data: Mean correct RT (ms)	Data: Acc	Fit: Mean correct RT (ms)	Fit: Acc	Model parameter: d'	u	v	a	z
Cautious same								0.125	0.056
Same	573	.891	577	.893		0.27			
Different									
1	605	.797	601	.791	2.33		−0.15		
2	517	.955	506	.950	3.28		−0.32		
3	480	.976	476	.974	3.61		−0.38		
4	461	.971	463	.981	3.78		−0.41		
Cautious different								0.125	0.069
Same	472	.967	476	.974		0.38			
Different									
1	695	.646	689	.641	2.58		−0.085		
2	584	.894	580	.889	3.56		−0.26		
3	531	.937	532	.946	4.00		−0.34		
4	518	.969	516	.960	4.17		−0.37		

Note. Other parameter values fixed across fits are T_{er} = 300 ms, s = 0.08, and η = 0.18. RT = reaction time; Acc = accuracy. From "Speed and Accuracy of Same and Different Responses in Perceptual Matching" by R. Ratcliff and M. J. Hacker, 1981, *Perception and Psychophysics, 30,* p. 305. Adapted by permission.

boundaries are asymmetric so that the starting point is nearer the *same* boundary, and if the goodness-of-match criterion is approximately centered between the distributions, then fast *same* responses along with a preponderance of false *different* responses are produced. The more general conclusion concerns the range of patterns of *same* and *different* response times and accuracy scores that can occur with bias manipulations. A manipulation can affect either the goodness-of-match criterion (*a* through *e* in Figure 2) or the position of the boundaries (different columns in Figure 2), or both. However, this model freedom applies only to the relative speed and accuracy of positive and negative responses and not to all kinds of data; Ratcliff (1978, 1981) discussed the kinds of data that constrain this model. In the following fits of the model to the data from the matching task, there are systematic changes in model parameters as a function of experimental manipulations (e.g., goodness-of-match criteria or diffusion-process boundaries) that allow clear theoretical interpretation of the effects of those experimental manipulations.

Ratcliff and Hacker (1981, Experiment 1)

In this experiment, subjects were presented with a four-letter study string followed by a four-letter test string. There were equal numbers of *same* and *different* trials, and the *different* trials were of four types: Either one, two, three, or four of the study letters were replaced by new letters. Reaction time and accuracy were the dependent measures. The main manipulation in this experiment was a manipulation of bias: In one group of sessions, subjects were instructed to respond *same* only when sure and *different* otherwise, and in another group of sessions, subjects had to respond *different* only when sure and *same* otherwise. The bias manipulation had a large effect on performance on both reaction time and accuracy, as can be seen in Table 1.

Also shown in Table 1 are fits of the diffusion model to the data. There are several things to note that about these fits. First, there is a very strong test of the model and that is the prediction of reaction time and accuracy for the four *different* conditions (e.g., one of the sets of data presented in Ratcliff, 1981, Figure 3, has the fits just missing the data). Once the boundary positions are fixed and the drift-rate parameters u and v (for one *different* condition) are fixed, then there is only one parameter free to vary over the other *different* conditions, and that is drift rate v. (In Figure 2, this would correspond to fixed boundaries and a fixed goodness-of-match criterion with several different nonmatch distributions.) Thus, reaction time and accuracy must fall on the function drawn by varying the parameter v. These fits are shown in Figure 3, and the model does a good job of accounting for the data.

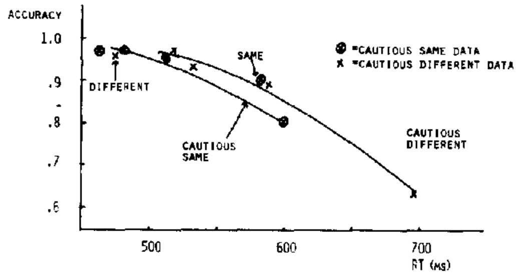

Figure 3. Fits of the diffusion model to the response time and accuracy data of Ratcliff and Hacker (1981), Experiment 1. (From "Speed and Accuracy of Same and Different Responses in Perceptual Matching" by R. Ratcliff and M. J. Hacker, 1981, *Perception and Psychphysics, 30,* p. 305. Adapted by permission.)

The second point to note is the effect of the bias manipulation on the boundary positions of the diffusion process for the two instruction conditions. The effect is symmetrical: The distance between the two boundaries is the same ($a = 0.125$) for the two conditions. For the cautious *same*, the position of the starting point is 0.056 above the *different* boundary, and for the cautious *different* the starting point is at 0.069 (=0.125 − 0.056).

This point is relevant to Proctor and Rao's (1982) criticism of Ratcliff and Hacker's (1981) conclusions from Experiment 1. Proctor and Rao subtracted reaction times for *same* and *different* responses for the two bias conditions and concluded that there was, overall, still a fast *same* effect. But Ratcliff and Hacker (1982) replied that, along with this fast *same* effect, were more false *same* errors. In other words, there was a trade-off: In responding more quickly on *same* trials, subjects also responded *same* incorrectly more often. This trade-off argued against Proctor and Rao's claim that the fast *same* effect represents a special stage of processing, because it is not clear how this special stage could produce the increase in false *same* errors. The above fits of the model to the data back up Ratcliff and Hacker's (1982) claim because the reaction time and error differences between the cautious *same* and cautious *different* conditions are produced by symmetric shifts in boundary positions. The reason that the difference in reaction times is not symmetric is because the goodness-of-match criterion settings are not symmetric (e.g., in Figure 2, a in the right column vs. d in the left column).

The third point is that performance in the cautious *different* condition is systematically better than in the cautious *same* condition. This shows up in the data, and as a result, in the fits of the model to the data. In the data in Table 1, reaction time and accuracy for *same* responses in the cautious *different* condition are about the same as reaction time and accuracy for the 3-*different* condition in cautious *same,* whereas reaction time and accuracy for *same* responses in the cautious *same* condition are about the same as those for 2-*different* in the cautious *different* condition. However, we shall see that in the data and fits to the data of Proctor and his colleagues, there are no systematic changes in d' as a function of probability condition. One possible difference between the Ratcliff and Hacker (1981) experiment and those of Proctor and colleagues is that the former used instruction condition as a between-session variable, whereas the latter varied probability within a session, leading to less possibility of altering performance.

As noted above, we would then expect that the model fits would mirror this difference because performance is symmetric for *same* and *different* processing. In the model fits, $d'_a = (u - v)/\eta$ is larger in the cautious *different* condition than in the cautious *same* condition (but only by a small but systemmatic 10%). This result suggests that bias manipulations produced changes not only in criteria but also in other aspects of performance, namely discriminability. To summarize, the most important results from these fits to Ratcliff and Hacker's (1981) Experiment 1 are the way the model fits the *different* conditions (a strong test of the model), and the way the caution instructions affect the bias parameters, both diffusion process boundaries and the goodness-of-match criterion.

It should also be noted that results for Experiment 2 in Ratcliff and Hacker (1981; in which difficult negatives were included) were replicated in Ratcliff (1981) and the model was fitted to these replication data. This experiment is mentioned here because adding difficult negatives alters relative bias settings. The result that is of main interest for present purposes is the finding that the easy *different* responses (strings with new letters) were much faster than *same* responses when there were difficult *different* trials (strings with one pair of letters transposed) in the experiment. This is a complete reversal of the usual fast *same* effect that is obtained when there are no difficult *different* trials in the experiment. The fits of the model to this data produced parameter estimates in line with those obtained above in the cautious *same* condition. The starting point of the diffusion process was nearer to the *different* boundary than the *same* boundary, and the size of the mean drift rate for easy *different* responses was larger than for *same* responses (i.e., $|v| > |u|$ so that the goodness-of-match criterion was nearer the match distribution than the easy nonmatch distribution). The pattern of results corresponds to criterion position *d* in the right panel of Figure 2. Further details of the fits can be found in Ratcliff (1981).

Proctor et al. (1984, Experiment 1)

Proctor et al. performed an experiment related to Ratcliff and Hacker's (1981) Experiment 1, using four-letter study and test strings. They used two kinds of presentation, sequential and simultaneous, but for the moment only the sequential condition will be considered. In Proctor et al.'s (1984) experiment, the *different* trials had only one letter replaced by a new letter (corresponding to 1-*different* in Ratcliff and Hacker, Experiment 1). The manipulation of most importance was a bias manipulation in which the probability of *same* and *different* trials was varied. There were four probability conditions, 20–80%, 40–60%, 60–40%, and 80–20% *same*–*different* trials, respectively. The data are shown in Table 2 along with the fits of the diffusion model. In both this and the next section on an experiment using single letters, the argument made by Proctor et al. was that there is still a fast *same* effect even when error rates are taken into account. When error rates for *same* and *different* responses are equal, there is still a reaction time advantage to *same* responses (as illustrated by the vertical dotted lines in the left column of Figure 2). This effect can be seen by interpolation in Table 2. When there are more false *different* errors (20S/80D condition), then reaction times are about equal, and if error rates were balanced (between the 20S/80D and 40S/60D conditions), there would still be a fast *same* effect. However, this does not demonstrate that the fast *same* effect cannot be produced by criterion changes in decision process. In fact the data presented in Table 2 are fitted well by the diffusion model, and the changes as a function of probability are well explained by changes in goodness-of-match criterion as in the left-hand panel in Figure 2.

The first thing to note about the fits of the model to the data shown in Table 2 is that except in the 80S–20D condition, boundary positions are approximately constant with the starting point nearer the *same* boundary. Thus the bias manipulation in this case was not strong enough to produce systematic changes in boundary positions. However, in the 80S–20D condition, the *same* boundary is 10% nearer the starting point (this corresponds to the decrease in reaction time for that condition).

The second point to note is that d'_a varies by less than 7% across the probability conditions. This suggests that discriminability is approximately constant across conditions (in contrast to the above fits to Ratcliff & Hacker's, 1981, Experiment 1). What does vary systematically with probability is the position of the criterion in the drift distributions. The parameter *u* decreases across conditions while *v* increases (which is equivalent to a change in the criterion). Thus, as probability conditions change from 80S–20D to 20S–80D, subjects move the criterion in the goodness-

Table 2
Parameter Values and Fits of the Diffusion Model to the Data of Proctor, Rao, and Hurst (1984)

	80s/20d		60s/40d		40s/60d		20s/80d	
Item	Mean correct RT (ms)	Acc	Mean correct RT (ms)	Acc	Mean correct RT (ms)	Acc	Mean correct RT (ms)	Acc
Data								
Same	559	.929	634	.890	664	.873	697	.834
Different	674	.796	693	.811	724	.817	690	.877
Fit								
Same	543	.921	629	.888	663	.879	695	.831
Different	673	.785	704	.811	729	.821	691	.877
Parameters								
a	0.135		0.15		0.16		0.16	
z	0.08		0.08		0.085		0.085	
u	0.26		0.23		0.22		0.18	
v	−0.17		−0.18		−0.185		−0.23	
d'	2.39		2.28		2.25		2.28	

Note. Other parameters fixed across conditions are T_{er} = 300 ms, s = 0.08, and η = 0.18. RT = reaction time; Acc = accuracy. From "An Examination of Response Bias in Multiletter Matching" by R. W. Proctor, K. V. Rao, and P. W. Hurst, 1984, *Perception and Psychophysics, 35*, p. 467. Copyright 1984 by the Psychonomic Society. Adapted by permission.

of-match distribution from nearer the non-match distribution to nearer the match distribution. This exact pattern is seen in the left column of Figure 2.

In Proctor and Rao (1983a), a case is made for the hypothesis that the same processes—except for a priming effect responsible for the fast *same* effect—underlie performance in the sequential matching task and the simultaneous matching task (in which the study and test strings are presented simultaneously and are left on display until a response has been made). However, use of the same processes would seem unlikely because, in the simultaneous case, subjects can engage in multiple comparisons of the stimuli on display. Also, in the multiletter comparison task, some aspects of the data suggest that the same processes are not used. In Proctor et al. (1984), the simultaneous matching condition with four-letter strings produced response latencies that were 400–500 ms longer than in the sequential matching condition. This is surprising because one would expect the simultaneous matching condition to be easier than the sequential matching condition because there is no memory load. However, error rates were lower in the simultaneous condition, so without an explicit model of the relationship between reaction time and error rate, the difficulty of the two conditions cannot be compared. For the purposes of this note, the possibility that subjects can engage in repeated comparisons (with perfect information) along with the large reaction time difference in the multiletter case makes the application of the diffusion model (using a single comparison process) questionable, and so fits will not be performed.

Proctor and Rao (1983a)

In another experiment, closely related to the preceding one, Proctor and Rao (1983a) varied probability in a single-letter matching task. The procedure was the same as above, and the probability manipulation was the same. The data and fits of the diffusion model are shown in Table 3. The fits of the diffusion model parallel those presented in the previous experiment. The boundary positions and d'_a are all roughly constant (to within 8% or 9%) across conditions. The parameter that does vary consistently across conditions is the position of the goodness-of-match criterion. As in the prior experiment, the starting point of the diffusion process was set nearer the *same* boundary leading to faster *same* responses

Table 3
Parameter Values and Fits of the Diffusion Model to the Data of Proctor and Rao (1983a)

	80s/20d		60s/40d		40s/60d		20s/80d	
Item	Mean correct RT (ms)	Acc	Mean correct RT (ms)	Acc	Mean correct RT (ms)	Acc	Mean correct RT (ms)	Acc
Data								
Same	412	.977	452	957	478	.939	482	.929
Different	493	.954	494	.966	493	.970	450	.983
Fit								
Same	413	.977	456	.961	481	.946	489	.934
Different	495	.953	491	.966	491	.967	445	.980
Parameters								
a	0.15		0.159		0.16		0.154	
z	0.085		0.09		0.09		0.084	
u	0.375		0.33		0.30		0.286	
v	−0.34		−0.37		−0.37		−0.41	
d'	3.97		3.89		3.72		3.87	

Note. Parameter values fixed across fits are T_{er} = 200 ms, s = 0.08, and η = 0.18. RT = reaction time; Acc = accuracy. From "Evidence That the *Same–Different* Disparity Is Not Attributable to Response Bias" by R. W. Proctor and K. V. Rao, 1983a, *Perception and Psychophysics, 34,* p. 74. Copyright 1983 by the Psychonomic Society. Adapted by permission.

when error rates are equal (interpolating in Table 3). As probability varies from 80S–20D to 20S–80D, the criterion varies from *a* to *d* in the left panel of Figure 2. Once again we see that the data and fits behave exactly as those shown in the left panel of Figure 2.

Generalizations About Criteria Setting in the Diffusion Model

In all four experiments described above, the diffusion model was successful in describing the data. Thus it is not necessary to invoke a separate, model-independent process to account for the fast *same* effect. As noted by Ratcliff and Hacker (1982), the diffusion model is capable of accounting for the fast *same* effect in terms of criteria effects.

From the fits presented above and the fits in Ratcliff (1981), it is possible to draw some general conclusions about the way in which bias manipulations affect the parameters of the model. First, probability manipulations were, on the whole, not sufficient to produce large changes in the values of the boundary positions. Instead, they affected the goodness-of-match criterion placement, which is the relative rate of accumulation of evidence between *same* and *different* comparisons. On the other hand, manipulation of caution instructions by Ratcliff and Hacker (1981) was sufficient to produce a change in the boundary positions along with a change in the goodness-of-match criterion. Another manipulation capable of producing a change in the boundary positions was the addition of difficult negatives (such as letter transpositions) to the *same* and replacement *different* comparisons (Ratcliff, 1981; Ratcliff & Hacker, 1981). In the fits in Ratcliff (1981), the starting point was placed nearer the *different* boundary (e.g., the right column of Figure 2), in contrast to the fits to Proctor and Rao (1983a) and Proctor et al. (1984) data (easy negative conditions). Whether these effects of probability manipulations, instructions, and the inclusion of difficult trials generalize to other paradigms in which the relative speed of *positive* and *negative* responses varies is a matter for further data collection and modeling.

A consistent pattern emerges in the behavior of the criteria in the model as it fit to the data from these experiments. The critical factor seems to be the relative difficulty of a judgment. If there are very difficult *different* trials along with easy *different* trials, then getting the *different* judgment correct is difficult and performance is best with the *same* boundary further away from the starting point than the *different* boundary (Figure 2, right column). This is because an error on a trial in which the stimuli are actually *different* (a false *same* error) is made when the *same* boundary is crossed so that the distance from the starting point to the *same* boundary controls the proportion of false *same* errors (see Figure 1). (This explanation is related to the assumption of heterogeneity of difference proposed by Krueger, 1979; see also Ratcliff, 1981, Figure 5.) On the other hand, if the *different* trials are all well discriminated from the *same* trials (e.g., Proctor et al., 1984), then the *same* trials may be perceived to be more difficult (it is easier to change from *same* to *different* by interference from other trials than from *different* to *same;* e.g., Krueger, 1978; Nickerson, 1972). Then the *different* boundary should be placed further away from the starting point than the *same* boundary (Figure 2, left column). The other criterion that can be changed is the goodness-of-match criterion. This has the same effect as in signal detection theory on accuracy (e.g., approximately as in the accuracy panel, the middle column of Figure 2) and also produces long reaction times in conjunction with high accuracy (Figure 2, changes in criterion from *a* to *e*). The pattern of behavior of these bias parameters as a function of experimental condition is reasonable and consistent, and the model provides a coherent account of these data.

Discussion

The fact that the diffusion model fits the data well raises the question of research strategy in examining effects such as the fast *same* effect. Once there are competing models that provide different accounts of the effect, then it is necessary to alter strategy from purely empirical investigations to including a strong component that involves testing and discriminating between models.

Discriminating among models requires consideration of what aspects of performance the various models are designed to explain. The diffusion model was designed to account for accuracy, reaction time, the shape of the reaction time distribution, and the growth of accuracy as a function of time. Ratcliff (1981) also provided a memory model that accounted for the relative difficulty of the *different* conditions (replacements and transpositions). Krueger's model (1978) has several features of the diffusion model (Ratcliff, 1978, 1981) at a global level. If some of the more detailed assumptions were relaxed and the process made more continuous, then many of the criticisms would be avoided and the model would take on the character of a sequential sampling model.

In contrast, the models proposed by Bamber (1969), Proctor (1981), and Taylor (1976) are limited in the way they can deal with bias data and speed–accuracy data in general. The mechanism for production of errors is not tied to the mechanism for producing reaction time differences, so that there is no natural way to model speed–accuracy relationships and changes in these relationships. In addition, there is no easy way for these models to deal with the transposition data of Ratcliff and Hacker (1981, Experiment 2).

It should be noted that other models of the sequential sampling class that have been applied to *same–different* judgments (e.g., Krueger, 1978, 1979; Link, 1978; Vickers, 1979) could probably be fitted to the data from the experiments considered here and provided a similar explanation. However, exactly how well the models might do and whether the same bias explanation for the fast *same* effect would be obtained can only be answered by detailed model fitting.

In the fits of the diffusion model to four experiments, some experimental manipulations (caution instructions, and addition of difficult negative trials) were sufficient to change the position of the diffusion-process boundaries and the position of the goodness-of-match criterion, whereas other manipulations (such as probability manipulations) were only sufficient to change the goodness-of-match criterion placement. The model explains the typical result of fast *same* responses along with a preponderance of false *different* errors by assuming that the *same* response boundary is placed nearer the starting point than is the *different* response boundary. One important conclusion that must be drawn from these fits and the model predictions in Figure 2 is that it is not sufficient to assume that an error bias in one direction must be accompanied by a reaction time difference in another direction for the results to be interpretable in terms of bias. This means that these fits call into question any conclusions based simply on raw data reaction times and error rates. In sum, what is needed is not so much verbal theories about how factors are related, but rather quantitative models that allow the relationships to be tested quantitatively.

References

Audley, R. J. (1960). A stochastic model for individual choice behavior. *Psychological Review, 67,* 1–15.

Audley, R. J. (1973). Some observations on theories of choice reaction time: Tutorial review. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 509-545). New York & London: Academic Press.

Bamber, D. (1969). Reaction times and error rates for "same"—"different" judgments of multidimensional stimuli. *Perception and Psychophysics, 6,* 169–174.

Falmagne, J. C. (1965). Stochastic models for choice-reaction time with applications to experimental results. *Journal of Mathematical Psychology, 2,* 77–124.

Falmagne, J. C., Cohen, S. P., & Dwivedi, A. (1975). Two-choice reactions as an ordered memory scanning process. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 296–344). New York: Academic Press.

Green, D. M., & Luce, R. D. (1973). Speed–accuracy tradeoff in auditory detection. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 567–569). New York: Academic Press.

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 and Performance, 5,* 277–288.

Krueger, L. E., & Shapiro, R. G. (1981). A reformulation of Proctor's unified theory for matching-task phenomena. *Psychological Review, 88,* 573–581.

Laming, D. (1968). *Information theory of choice reaction time.* New York: Wiley.

Laming, D. (1979a). A critical comparison of two random-walk models for two-choice reaction time. *Acta Psychologica, 43,* 431–453.

Laming, D. (1979b). Choice reaction performance following an error. *Acta Psychologica, 43,* 199–224.

Link, S. W. (1975). The relative judgment theory of choice reaction time. *Journal of Mathematical Psychology, 12,* 114–135.

Link, S. W. (1978). The relative judgment theory of the psychometric function. In J. Requin (Ed.), *Attention and performance VII* (pp. 619-630). Hillsdale, NJ: Erlbaum.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40,* 77–105.

Luce, R. D. (in press). *Response times and mental structure: Models of simple and choice reactions.* Oxford, England: Oxford University Press.

Nickerson, R. S. (1972). Binary-classification reaction time: A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements, 4,* (17, Whole No. 65).

Ollman, R. T. (1966). Fast guesses in choice reaction time. *Psychonomic Science, 6,* 155–156.

Pike, A. R. (1973). Response latency models for signal detection. *Psychological Review, 80,* 53–68.

Pike, R., Dalgleish, L., & Wright, J. (1977). A multiple-observations model for response latency and the latencies of correct and incorrect responses in recognition memory. *Memory & Cognition, 5,* 580–589.

Proctor, R. W. (1981). A unified theory for matching task phenomena. *Psychological Review, 88,* 291–326.

Proctor, R. W. & Rao, K. V. (1982). On the "misguided" use of reaction-time differences: A discussion of Ratcliff and Hacker (1981). *Perception and Psychophysics, 31,* 601–602.

Proctor, R. W., & Rao, K. V. (1983a). Evidence that the *same–different* disparity is not attributable to response bias. *Perception and Psychophysics, 34,* 72–76.

Proctor, R. W., & Rao, K. V. (1983b). Reinstating the original principles of Proctor's unified theory for matching-task phenomena: An evaluation of Kreuger and Shapiro's reformulation. *Psychological Review, 90,* 21–37.

Proctor, R. W., Rao, K. V., & Hurst, P. W. (1984). An examination of response bias in multiletter matching. *Perception and 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., & Hacker, M. J. (1981). Speed and accuracy of same and different responses in perceptual matching. *Perception and Psychophysics, 30,* 303–307.

Ratcliff, R., & Hacker, M. J. (1982). On the misguided use of reaction-time differences: A reply to Proctor and Rao (1982). *Perception and Psychophysics, 31,* 603–604.

Stone, M. (1960). Models for reaction time. *Psychometrika, 25,* 251–260.

Swensson, R. G. (1972). The elusive tradeoff: Speed vs. accuracy in visual discrimination tasks. *Perception and Psychophysics, 12,* 16–32.

Taylor, D. (1976). Effect of identity in the multiletter matching task. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 417–428.

Vickers, D. (1979). *Decision processes in visual perception.* New York: Academic Press.

Yellott, J. I. (1967). Correction for guessing in choice reaction time. *Psychonomic Science, 8,* 321–322.

Yellott, J. I. (1971). Correction for guessing and the speed–accuracy trade-off in choice reaction time. *Journal of Mathematical Psychology, 8,* 159–199.

Received December 9, 1983
Revision received July 22, 1984 ■