Journal of Experimental Psychology:
Human Perception and Performance
2000, Vol. 26, No. 1, 127–140

Copyright 2000 by the American Psychological Association, Inc.
0096-1523/00/$5.00 DOI: 10.1037//0096-1523.26.1.127

A Diffusion Model Account of Masking in Two-Choice Letter Identification

Roger Ratcliff and Jeffrey N. Rouder
Northwestern University

The diffusion model developed by R. Ratcliff (1978, 1981, 1985, 1988) for 2-choice decisions was applied to data from 2 letter identification experiments. In the experiments, stimulus letters were displayed and then masked, and the stimulus onset asynchrony between letter and mask was manipulated to vary accuracy from near chance to near ceiling. A standard reaction time procedure was used in one experiment and a deadline procedure in the other. Two hypotheses about the effect of masking on the information provided to the decision process were contrasted: (a) The output of perception to the decision process varies with time, so that the information used by the decision process rises and falls, reflecting the stimulus onset and mask onset. (b) The output of perception to the decision is constant over time, reflecting information integrated over the time between the stimulus and mask onsets. The data were well fit by the diffusion model only with the assumption of constant information over time.

Sequential sampling models have been used extensively to describe rapid two-choice decisions about simple perceptual and cognitive stimuli. For example, when subjects are asked to decide, as quickly as possible, which of two tones, two lights, two line lengths, or two letters was presented, sequential sampling models are generally successful in fitting most aspects of response time and accuracy data (Audley & Pike, 1965; Heath, 1981; Laming, 1968; Link & Heath, 1975; Smith & Vickers, 1988; Vickers, 1979). In this article, we extend one sequential sampling model, Ratcliff's (1978) diffusion model, to investigate data from a two-choice letter-matching paradigm in which visually presented stimuli were masked shortly after being displayed.

The question we address is how a mask affects the stimulus information that enters the decision process. One hypothesis is that the information coming from early perceptual processes rises and then falls, reflecting the earlier onset of the stimulus and subsequent obscuring of the stimulus by the mask. A second hypothesis is that the information entering the decision process is constant over time, representing the total amount of information encoded from the stimulus during the time from its onset to the onset of the mask. Figure 1 illustrates these two hypotheses in terms of the diffusion model (Ratcliff, 1978). In this model, information from a stimulus is accumulated from a starting point (z) toward one or the other of two criteria (a and 0), with each criterion representing one of the two possible responses. A decision is made when the amount of information reaches one of the criteria. *Drift rate* is the rate at which information is accumulated. The top of Figure 1 shows "nonstationary" drift; the drift rate starts out with a positive value (v) and then drops to zero, reflecting the rise of stimulus information and then its termination because of the mask. With a longer stimulus duration (2 vs. 1 in the figure), the accumulation of information proceeds farther before the drift rate drops to zero. The bottom of Figure 1 shows the "stationary" drift rate; drift rate is a constant value determined by stimulus duration. The value of v is greater as stimulus duration increases. Note that we did not examine all possible nonstationary models, just the ones that assume that information at the decision processes follows availability of stimulus information at input.

The two experiments described in this article were designed to distinguish between the stationary and nonstationary drift hypotheses. For both experiments, on each trial a letter was briefly displayed and then masked, and subjects were asked to decide which of two letters the stimulus matched. The mask was a square filled with random contours that was slightly larger than the letters. As we discuss, the stationary and nonstationary hypotheses make different predictions about the relation between response speed and accuracy. Empirical examination of the predictions requires a full range of accuracy values, from near-perfect performance to near chance. To accomplish this, in Experiment 1 we used a range of stimulus onset asynchronies (SOAs) between stimulus onset and mask onset and we varied the difficulty of the discrimination between the candidate letters. In Experiment 2, we used a deadline procedure to trace the time course of the decision process from early in processing (near-chance performance) to later, asymptotic performance.

Roger Ratcliff and Jeffrey N. Rouder, Department of Psychology, Northwestern University.

Jeffrey N. Rouder is now at the Department of Psychology, University of Missouri.

This research was supported by National Institute of Mental Health Grant HD MH 44640 and National Institute for Deafness and Other Communication Disorders Grant R01-DC01240.

We thank Gail McKoon for extensive comments on this article.

Correspondence concerning this article should be addressed to Roger Ratcliff, Department of Psychology, Northwestern University, Evanston, Illinois 60208. Electronic mail may be sent to r-ratcliff@nwu.edu.

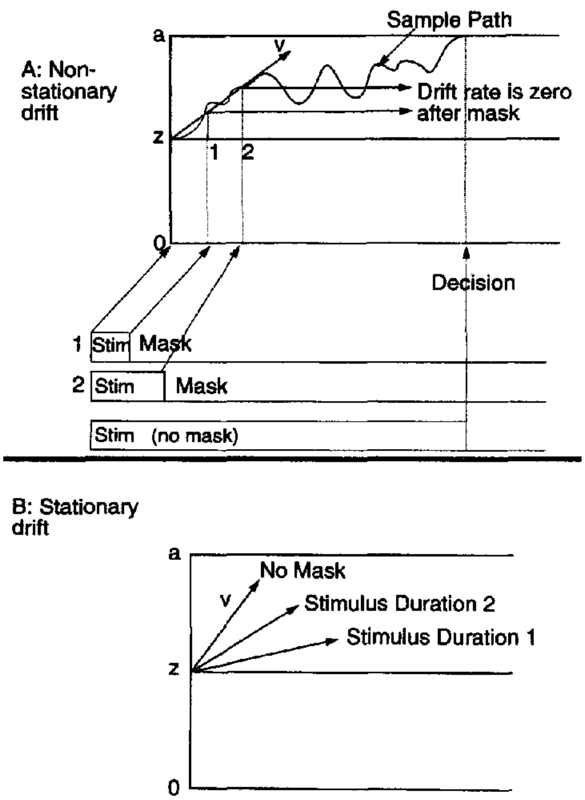

Figure 1. An illustration of the two assumptions about how stimulus mask (Stim) stimulus onset asynchrony relates to drift rate. A: Drift rate set to zero when the mask is presented. B: Constant drift rate. Note that the drift rates shown are means; the actual paths vary around the means.

Background

The two hypotheses, stationary versus nonstationary drift, are hypotheses about the information that enters the decision process. They are not hypotheses about the effects of a mask on perception, and, for this reason, they are not directly tied to theories of masking. That is, either stationary or nonstationary drift would be consistent with any of the current ideas about how masking affects perception. Here, we briefly review several theories about the effects of masking to make this point.

In 1968, Kahneman reviewed the then-current literature on masking and identified "integration" and "interruption" theories (see also Turvey, 1973). Integration theories assume that perceptual processes sum the stimulus and the mask so that they form a composite, making the stimulus less intelligible. Both theories describe early perceptual processing. They do not describe how the information involved in this processing might enter the later decision process. The information entering the decision process might be nonstationary, reflecting early visual information, information that rises and then falls before being integrated or interrupted. Alternatively, the information entering the decision process might be constant (stationary), reflecting the amount of information derived from the stimulus before the mask interrupted perceptual processing or reflecting the amount of information available from the integrated stimulus and mask.

Like the earlier integration and interruption hypotheses, newer ideas about the effects of masking on perception also do not provide any means of deciding in what form information enters the decision process. Much research has been concerned with "metacontrast" masking in which performance on the stimulus is adversely affected even though the mask is displayed at a different spatial location than the stimulus. Explanations of this effect (e.g., Breitmeyer, 1984; Breitmeyer & Ganz, 1976; Dixon & Di Lollo, 1994) involve mechanisms of lateral suppression and temporal integration of visual information as well as how those mechanisms determine whether a stimulus can be perceived in an identification task, but they do not specify at what point or what kind of information is read out to decision processes.

There have also been several theories of letter and word identification that model the effect of stimulus presentation time on the acquisition of information in perception. Loftus and Ruthruff (1994) and Busey and Loftus (1994) provided evidence that performance in a digit identification task depends on the integral of information encoded by the perceptual system. Their model assumes that perceptual information is convolved with an impulse response function to produce a sensory response function. This sensory response function is integrated over time to produce the total sensory response, which serves to determine the probability with which a stimulus is identified (probability equals $1 - \exp$ [sensory response]). Like the other models reviewed here, this model is designed to explain under what conditions a stimulus can be identified, not to specify what kind of information from perception might enter a later decision process. However, if the model were extended to explicitly describe the time course of the integration mechanism, then it might be possible to enter the integrated information over time into the decision process of the diffusion model so that as the size of the integral grows rapidly, information driving the decision process would grow rapidly and then remain stationary until a decision is made.

Another model that deals with the time course of acquisition has been proposed by Bundesen (1990). His model is designed to explain visual attention in paradigms such as letter identification. The model assumes that stimuli are represented as features and that features are extracted with exponential finishing times. A mask terminates processing of the stimulus, and the response is based on the features extracted to that point. For visual search, the model uses a deadline mechanism to predict mean reaction time. The model could be made consistent with our constant drift assumption if the categorical information extracted from the stimulus before the mask was used as the quantity determining a constant drift rate.

The one model of the effects of masking on perception that explicitly models the time course of information accumulation and the output of that information to response mechanisms is the *interactive activation model* (McClelland & Rumelhart, 1981; see also Grainger & Jacobs, 1996). Its prediction about the output of information falls between the stationary and nonstationary drift hypotheses. In the model, activation is passed among three layers of nodes in a localist connectionist network. The input layer consists of nodes for features of letters that are connected to letter nodes, which in turn are connected to word nodes. There is a different network of features and letters for each letter position in a word. Within a layer, the nodes inhibit each other. When a stimulus is input to the system, activation flows from the feature nodes up to the letter nodes and from there to the word nodes. Activation also flows downward, so that activated words send activation down to letters consistent with them. The model assumes that at output, a weighted average of activation, activation integrated over time, is used to compute a response strength for each node, word nodes if the task is word identification or letter nodes if the task is letter identification. Which response is selected is determined by Luce's choice rule.

A mask is input to the system by turning off the activation in all the feature nodes, which rapidly suppresses activation in the whole system. However, response strength decreases less abruptly at output than does activation at the word and letter nodes because it is a weighted average of activation (see McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). Response strength decays as a function of time at about the same rate as the rise in response strength as a function of SOA (see McClelland & Rumelhart, 1981, Figure 8; see also Loftus & Ruthruff, 1994, for a similar response function to a square pulse being transmitted through filters). Thus, in a masking paradigm, the function for the rise and fall of response strength is approximately an inverted V. If this strength is input as drift rate to the decision process in the diffusion model, then the drift rate should be nonstationary but with an inverted-V shape rather than the abrupt change to zero drift for the nonstationary hypothesis. We used the data from the experiments to test this inverted-V hypothesis as well as the two hypotheses already described.

The Diffusion Model

The diffusion model is one of a class of sequential-sampling, random walk–diffusion models that have proved successful in accounting for a range of data across a range of experimental paradigms such as choice reaction time (Heath, 1981, 1992; Laming, 1968; Link, 1975; Link & Heath, 1975; Stone, 1960), simple reaction time (Smith, 1995), memory retrieval (Ratcliff, 1978, 1980, 1988), letter matching (Ratcliff, 1981, 1985), numerosity judgments (Ratcliff, Van Zandt, & McKoon, 1999), various perceptual judgments (Ratcliff & Rouder, 1998), visual scanning (Strayer & Kramer, 1994), and decision making (Busemeyer & Townsend, 1993). In the contexts of these models, both stationary and nonstationary processes for the accumulation of evidence have been proposed. Stationarity is the default assumption in most of the models. Nonstationarity has been used in models of simple reaction time (Smith, 1995) and models of decision making (Busemeyer & Townsend, 1993). In the perception and masking literature, Heath's (1981, 1992; see also Ratcliff, 1980) tanden random walk model proposes a nonstationary drift rate, and Audley and Pike (1965) alluded to a nonstationary hypothesis for a random walk model: The probability of a step toward one response criterion versus the other is greater than chance when the stimulus is displayed but returns to chance at mask onset. However, none of these suggestions was fully implemented and explicitly fit to the kind of comprehensive data we present here.

We chose the diffusion model as the main vehicle with which to investigate the stationary versus nonstationary hypotheses because, in the domains to which it has been applied, it is capable of accounting for the behavior of all the dependent variables: mean response times for correct responses and error responses, probabilities of correct and error responses, and the shapes of the distributions of response times. It also explains the relative speeds of correct versus error responses, something other models have not been found to do (except that of Smith & Vickers, 1988, in one experimental procedure).

The diffusion process is a general decision mechanism that accumulates information over time toward one or the other of two possible response criteria. It is designed to describe single-stage decision processes with mean response times of not much more than about 1–1.5 s, not decision processes that require multiple stages. The components of the model are the starting point for the accumulation of evidence (z), variability across trials in the starting point (s_z), the response criteria (boundary positions 0 and a) toward which evidence is accumulated, the rate of accumulation (drift rate v), the variability in the drift rate within a trial (s), the variability in drift rate across trials (η; assuming a normal distribution for drift rates), and the nondecision components of reaction time (T_{er}).

In the top panel of Figure 2, the mean drift rates for two letters are labeled v_+ and v_- and both distributions have standard deviation η. Eta is the across-trial variability in drift rate; it reflects variability in the value of the quality of information about a stimulus on different trials. In memory paradigms, for example, a word to be remembered might be more strongly encoded on one trial than another or for one subject than another. In perception paradigms, exactly the same stimulus might be better encoded on one trial than another.

In Experiment 2, the deadline procedure allows measurement of an asymptotic value of d', that is, d'_a (see the top panel of Figure 2). In the diffusion model, this is computed by subtracting the mean drift rates for the two letters (one is usually positive and the other negative) and dividing by the standard deviation.

There is also within-trial variability in drift (see the sample path in the top panel of Figure 1). Because of this variability, the amount of accumulated evidence can reach the wrong boundary even with a large positive or negative value of drift. Movement of the boundary positions allows the model to account for speed–accuracy trade-offs. Boundaries close to the starting point mean fast response times but high error rates. Boundaries farther from the starting point

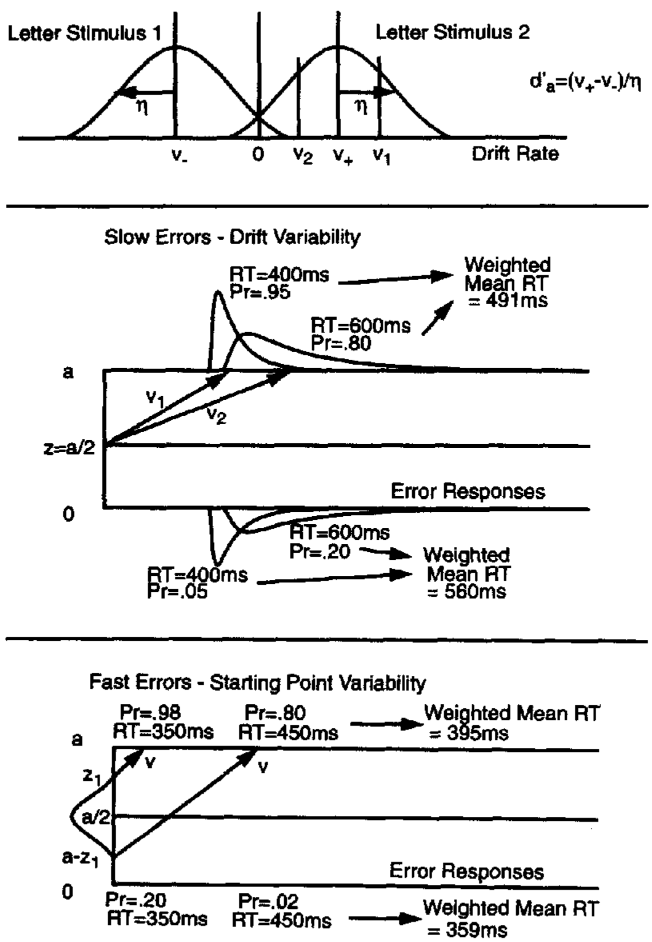

Figure 2. Illustration of how variability in drift across trials leads to slow errors and how variability in starting point across trials leads to fast errors. The top panel shows distributions of drift rates across trials for the two response choices, the middle panel shows averaging two drift rates, and the bottom panel shows the averaging of two starting points. RT = reaction time; Pr = probability.

mean slower response times but a greater chance for the accumulation of evidence to reach the correct boundary.

Empirical response time distributions are typically positively skewed. The diffusion model naturally predicts this shape by simple geometry: Equal size decreases in drift rate (e.g., from $3v$ to $2v$ to v) do not lead to equal size increases in response time but instead to increasingly larger increases in response time.

A recent discovery about the diffusion model is that the combination of across-trial variability in drift rate and across-trial variability in starting point allows the model to account for the relative speeds of error versus correct response times, something no other model has been able to do (Ratcliff & Rouder, 1998; Ratcliff et al., 1999; see also Van Zandt & Ratcliff, 1995). The middle panel of Figure 2 illustrates the effect of variability in drift rate using two values of drift rate (v_1 and v_2 in the top panel) rather than the whole distribution of drift rates that would be used in a real implementation of the model. Because the larger drift rate (v_1) produces fast error reaction times but fewer of them than the smaller drift rate (v_2), the weighted average reaction time for errors is longer than the weighted average for correct responses. The bottom panel of Figure 2 shows the effect of variability in starting point, again using two values (z_1 and $a–z_1$) for illustrative purposes instead of a whole distribution (variability in the boundary positions would produce the same result as variability in starting point). When the starting point is near the error boundary, it hits quickly and with high probability, whereas when it is nearer the correct boundary, errors occur with low probability and they are slow. The weighted average leads to faster error responses than correct responses.

The combination of variability in these two parameters leads to one of three patterns of results: errors faster than correct responses if starting point variability is large; errors slower than correct responses if drift variability is large; and a crossover such that errors at intermediate levels of accuracy (e.g., 0.5–0.9) are slower than correct responses and errors at extreme levels of accuracy (e.g., above .95) are faster than correct responses if both kinds of variability have moderate-to-large values. The data from Experiment 1 below show the crossover pattern.

Predictions for Latency Probability Functions

One of the strengths of the diffusion model is that it jointly and simultaneously predicts speed and accuracy. The relation between the two measures can be displayed with latency probability functions (Audley & Pike, 1965; Vickers, 1979; Vickers, Caudrey, & Willson, 1971). Values of response time are plotted against values of the probability of a response to form a parametric plot, where drift rate is the parameter that traces out the function. Latency probability functions have not often been used, probably because they immediately force evaluation of error response times that few models can fit with any accuracy.

Figure 3 shows sample latency probability functions. The vertical axis is response time, and the horizontal axis is the probability with which one of the two possible responses is made (the probability values for the other response would be a mirror image). When the probability of the response is high, generally above .5, it is the correct response; when the probability is low, it is not the correct response. In other words, points on the right-hand side of the functions are generally correct responses, and points on the left-hand side are generally errors. Figure 3 shows the latency probability functions that the diffusion model produces for two values of boundary position (with the two boundaries set equidistant from the starting point so that $z = a/2$) and for stationary and nonstationary values of drift. For stationary drift (the black filled symbols), points on the latency probability function correspond to drift rates of 0.3, 0.2, 0.1, and 0.05. The points for correct responses for these drift rates are on the right-hand side of the function, and the points for error responses are on the left-hand side. The higher the value of drift rate, the more likely are correct responses (farther right points) and the less likely are errors (farther left points). The

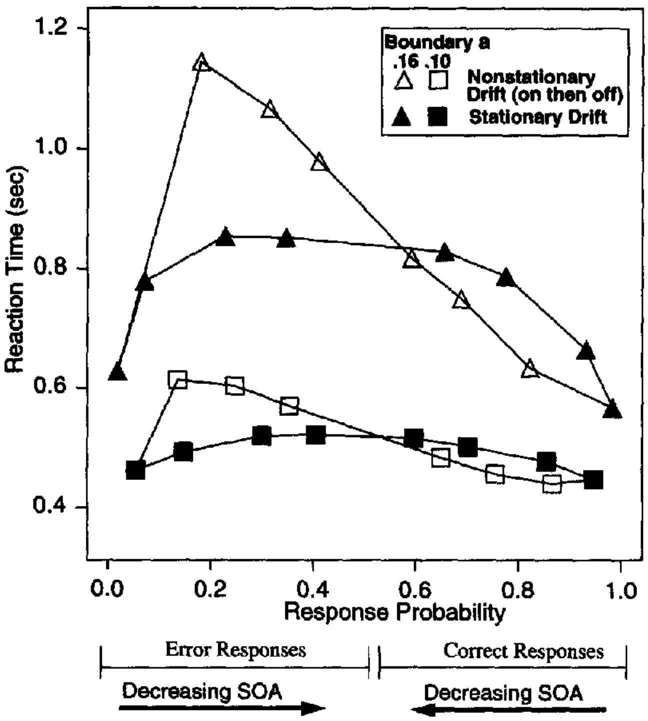

Figure 3. Latency probability functions for stationary drift and nonstationary drift for two boundary positions in the diffusion model. SOA = stimulus onset asynchrony.

four values of drift rate represent varying qualities of stimulus information, as would result, for example, from varying the SOA between letter onset and mask onset. For nonstationary drift (the open symbols), the points on the function correspond to varying the time at which drift rate goes from 0.3 to 0; the times are 0.05, 0.1, or 0.2 s or never. If the drift rate of 0.3 never drops to 0, then the predicted response times and probabilities are the same as for the stationary process with drift rate 0.3, so that the farthest right and farthest left points are the same for the stationary and nonstationary functions. So the response times correspond numerically to what would be expected from real data, we added a nondecisional constant time, T_{er} = 0.3 s, to each reaction time value. For both the stationary and nonstationary functions, the other parameters of the model were as follows: η = .08 and s = .1; these values were close to those used to fit the data from Experiments 1 and 2. For nonstationary drift, exact solutions are not available, so we simulated the diffusion process with a random walk with a small step size (cf. Feller, 1968); the resulting predictions are within a few percentage points of the predictions that would be generated by the continuous diffusion model.

The asymmetry of latency probability functions provides a display of the relative speeds of correct versus error responses. The response time for any particular value of correct response probability can be compared with the response time for the corresponding value of error probability. For example, response times for correct responses with a probability of .6 would be compared to with response times for error responses with a probability of .4. Points on the left-hand side of the function are errors, and if they are higher than their corresponding points on the right side, then the errors are slower than the correct responses. The latency–probability function also provides a visual presentation of the rate of change of reaction time as a function of accuracy, another way it highlights the joint behavior of the two dependent variables.

The latency probability functions show clearly different predictions for stationary versus nonstationary drift rates. For stationary drift, the latency probability functions are nearly symmetrical with errors only modestly slower than correct responses, a pattern similar to what has been observed before in situations where drift rate was assumed to be stationary (Ratcliff & Rouder, 1998; Ratcliff et al., 1999). For nonstationary drift, the latency probability function is highly asymmetrical. Errors are substantially slower than correct responses, except at extreme response probabilities for which most responses would have terminated before the mask affected the decision process. The reason for the very slow errors is not intuitively apparent. They come about because the positive drift rate before the mask drives processes away from the starting point toward the correct boundary; when drift rate goes to zero, the process acts like a process with zero drift rate and starting point close to one boundary. In this situation, error responses are very slow (see Ratcliff, 1988, for a description of the distribution of nonterminated processes left in the decision process as a function of time).

Experiment 1

Experiment 1 was designed to map out latency probability functions to test the stationary versus nonstationary drift rate hypotheses. To obtain a wide range of accuracy values, we varied the SOA between letter onset and mask onset from 12 to 84 ms, and there were two levels of difficulty for the discrimination between the letter choices.

Method

Subjects. Eighteen Northwestern University undergraduates participated to fulfill a course requirement.

Apparatus. Letters were displayed on video graphics array computer monitors in text mode (80 columns, Courier font). The video card driving each monitor was reprogrammed so that it would execute a refresh every 4 ms (250 Hz; e.g., Von Brisinski, 1994). The mask was fixed across trials and consisted of a square outline with random horizontal, vertical, and diagonal lines inside it.

Procedure. Five levels of SOAs were used: 12, 20, 32, 52, and 84 ms. The more difficult to discriminate letter pairs were E versus F, P versus R, and C versus G; the easier pairs were E versus C, P versus G, and C versus F.

Each subject participated in one session that consisted of 18 blocks of 50 trials each, resulting in a total of 900 trials (90 trials per SOA per easy vs. difficult letter pair). At the beginning of each trial, the display consisted of a plus sign fixation point and the two letter alternatives. The plus sign was located in the center of the display. One response alternative was located above and to the left of the plus sign and the other above and to the right of the plus sign. This initial display lasted 600 ms, and then the stimulus letter was

Table 1
Accuracy, Correct and Error Response Times (RTs), and Standard Errors in RTs

Stimulus onset asynchrony (ms)	Probability correct	Correct RT (ms)	Standard error in RT (ms)	Error RT (ms)	Standard error in RT (ms)
		Easy discrimination			
12	.568	569	7	583	8
20	.716	538	5	581	12
32	.861	497	4	513	13
52	.914	474	4	495	24
84	.934	476	3	400	18
		Hard discrimination			
12	.494	573	8	579	8
20	.576	545	7	566	10
32	.723	522	5	546	13
52	.846	510	5	524	17
84	.904	500	4	419	14

displayed and then masked. The mask remained on until the subject made a response. Subjects were instructed to press the *z* key on the computer keyboard if the stimulus letter matched the left alternative and the "/" key if it matched the right alternative.

The same two letters were used as response alternatives for all the trials of a block. The blocks were grouped in 3s such that the three blocks of a group each used one of the three easier to discriminate letter pairs or the three blocks each used one of the three more difficult letter pairs. The easy versus more difficult groups of blocks alternated with their order counterbalanced across subjects. The SOA for each trial was chosen randomly from among the five possible values. The first 20 trials of the session and the first trial of each block were considered warmup trials and were not included in the data analyses. Subjects received no feedback about accuracy or RT. They were told that accuracy and reaction time were being measured and to go with their first impression of the stimulus. They were given short breaks between blocks of trials.

Results and Fits of the Diffusion Model

Table 1 shows mean reaction times and accuracy rates as a function of the two levels of discrimination difficulty and five SOA values. The data are plotted in latency probability functions in Figures 4 and 5 for easy and difficult discriminations, respectively. In the experimental data, errors were faster than correct responses for extreme values of accuracy and were slower than correct responses for intermediate levels of accuracy (for other examples of this crossover in correct vs. error responses, see Ratcliff & Rouder, 1998; Ratcliff et al., 1999; Smith & Vickers, 1988). The functions have the shape predicted from the stationary drift hypothesis; they are not consistent with the nonstationary drift hypothesis.

Representative cumulative reaction time distributions are shown in Figure 6. The means of five quantiles are plotted. Only five were used because there were few error observations per subject at the highest values of accuracy. To compute the mean quantiles for a subject, we divided the response times into five equal-sized groups ranging from fastest to slowest and computed the mean of each group. Figure 6 shows the means of these means across subjects. The distributions show a skew to the right, typical of reaction time distributions.

For fitting the diffusion model with the constant drift assumption, we used a general fitting program called Simplex (Nelder & Mead, 1965). The program minimized the sum of squares between theoretical and observed values of accuracy and theoretical and observed means of the five quantiles of the correct and error response time distributions for each SOA and difficulty condition. The theoretical values were computed from exact numerical solutions involving numerical integration for variability in drift and variability in starting point (see Ratcliff et al., 1999). In minimizing the sum of squares, accuracy values were weighted twice as much as the RT quantiles, and the first and fifth reaction time quantiles were weighted half as much as the other quantiles. This weighting scheme was used because standard errors in accuracy were smaller than those in the reaction time

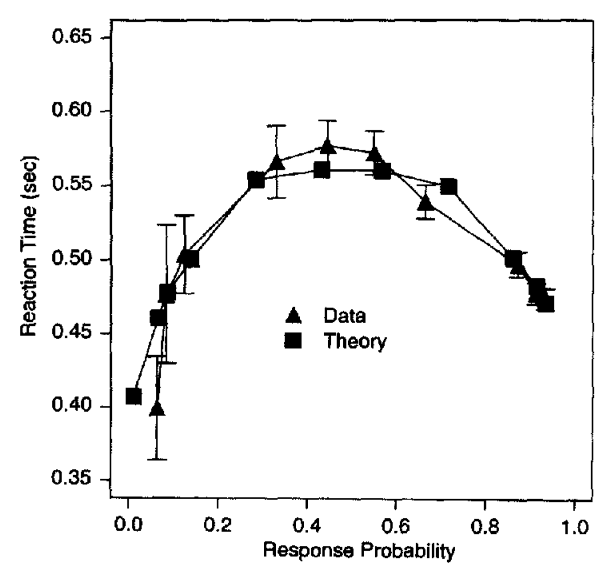

Figure 4. Latency probability functions for the data for the easy discrimination condition in Experiment 1 and fits of the diffusion model with stationary drift.

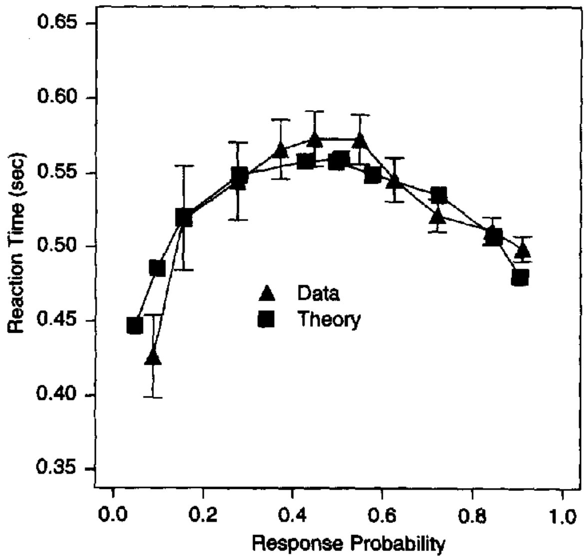

Figure 5. Latency probability functions for the data for the hard discrimination condition in Experiment 1 and fits of the diffusion model with stationary drift.

quantiles, and the first and fifth reaction time quantiles were more likely to have outliers associated with them. The program adjusted the values of the boundary parameter (a), the nondecisional component of reaction time (T_{er}), the variability in drift across trials (η), and the variability in starting point (s_z), and there was a different value of drift (v) for each SOA and difficulty condition.

The fits of the model to the latency–probability functions are shown in Figures 4 and 5, and fits to the sample cumulative response time distributions are shown in Figure 6. In general, the model fit the data well: With only drift rate varying among conditions, the shapes of the latency probability functions and the response time distributions are well described. Furthermore, the theoretical functions fall within or close to 2 *SE*s of the experimental data, except for the errors when accuracy is more than 95% and the data are based on relatively few observations. Table 2 and Table 3 show the parameter values for the fits. Although the boundary position parameters (a, $z = a/2$), the variability in the starting point parameter (s_z), and the encoding and response parameter (T_{er}) were allowed to vary freely between the easy and difficult conditions, the resulting values were within a few percentage points of each other, and setting them equal to each other would not alter the quality of the fits. Therefore, according to the model, the only components of processing that varied appreciably across conditions were the across-trial variability in drift rate and the drift rates for the different conditions.

The three parameters other than drift rate (a, η, and s_z) specify the form of the latency probability function. The latency probability function can be viewed as a parametric plot, with the parameter of the plot being drift rate. In other words, the predicted line is based on three parameters that describe the allowable values of correct and error response

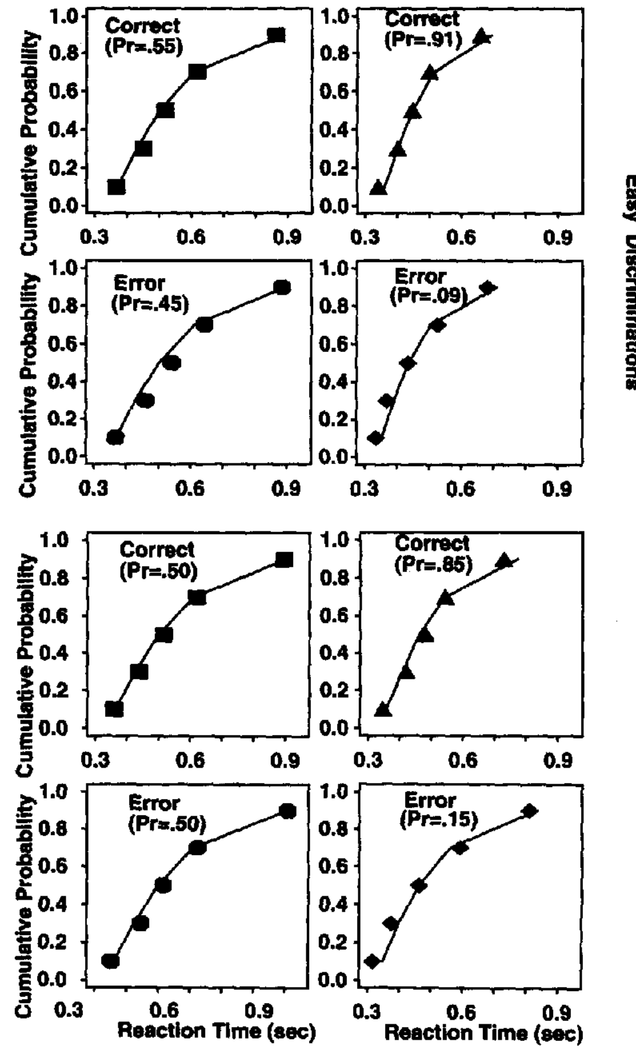

Figure 6. Cumulative response time distributions for sample data from Experiment 1. The top half shows easy discriminations and the bottom half hard discriminations. For each of these conditions, one high-accuracy and one low-accuracy condition is shown (one figure for correct responses and one for error responses). Pr = probability.

times and response probabilities. (The parameter T_{er} adjusts the vertical position of the function, not the shape.) The three parameters also predict the possible range of shapes of response time distributions. The different values of drift rate

Table 2
Parameters of the Fits of the Diffusion Model to Experiments 1 and 2

Parameter	Experiment 1: Hard condition	Experiment 1: Easy condition	Experiment 2
a	.108	.104	
z	.054	.052	
s_z	.018	.014	
T_{er}(s)	.316	.319	.303
s	.1	.1	.1
η	.111	.071	.170

Table 3
Drift Rates and d' Values for the Different Conditions From the Fits of the Diffusion Model to Experiments 1 and 2

SOA (ms)	Drift rates: Experiment 1: Hard condition	Drift rates: Experiment 1: Easy condition	d'_a values: Experiment 1: Hard condition	d'_a values: Experiment 1: Easy condition	d'_a values: Experiment 2
12	0.026	0.024	0.47	0.68	
20	0.067	0.080	1.21	2.25	1.76[a]
32	0.124	0.224	2.19	6.28	3.07[a]
52	0.216	0.274	3.89	7.71	
84	0.298	0.311	5.37	8.76	
None					6.63

Note. d'_a values are computed from 2(drift rate)/η, that is, $(v_+ - v_-)/\eta$, where v_- is $-v_+$. SOA = stimulus onset asynchrony.
[a]Actual average mask durations were 22 and 31 ms.

specify particular points on the latency probability function for the particular SOAs used in the experiment.

For the nonstationary drift model with drift rate going to zero at mask onset and the nonstationary inverted-V from the interactive activation model, the fits are shown in Figures 7 and 8; both figures show data and fits to the easy discrimination condition. The program used to fit these models simulated the diffusion process by a random walk with a small step size (0.1 ms per step) using approximations derived from Busemeyer and Townsend (1992). Because this was a simulation, many simulated trials (10,000 here) were required to produce accurate predictions for reaction time and accuracy. A simulation of the diffusion process was required because there are no general explicit solutions for cases in which the drift rate or other parameters of the model are changed during the time course of retrieval. Various numerical methods are available (Smith, 1995, Appendix B), but the simulation method is much easier to use as long as there is enough computer power available. The simulation of the diffusion process was embedded within the simplex minimization routine that adjusted parameters of the model to minimize squared differences between theoretical and experimental values of accuracy and mean reaction time for correct and error responses. Figure 7 shows the fits of the model with the drift rate constant for the duration of stimulus presentation and then zero after that. Figure 8 shows the fits of the model with drift rate increasing linearly for the duration of the stimulus and decreasing to zero at the same rate (the inverted-V function). The best fitting functions miss the data by several standard errors in some places, and the shapes of the latency probability functions are not like those shown by the data. The best

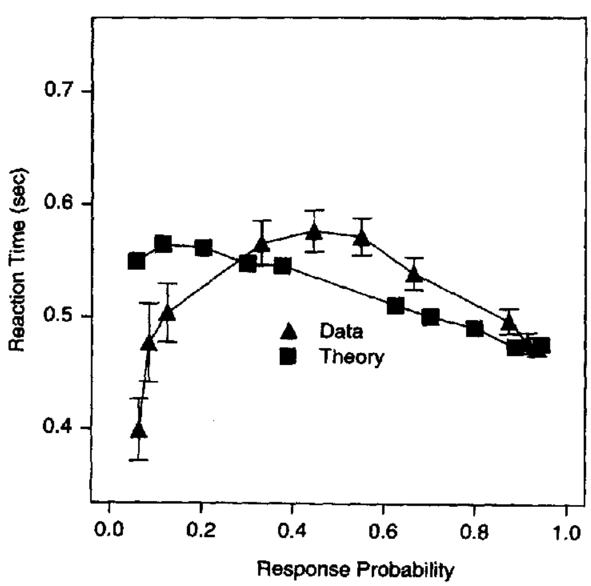

Figure 7. Latency probability functions for the data for the easy discrimination condition in Experiment 1 and fits of the diffusion model with the drift constant up to the mask onset and then zero until the response.

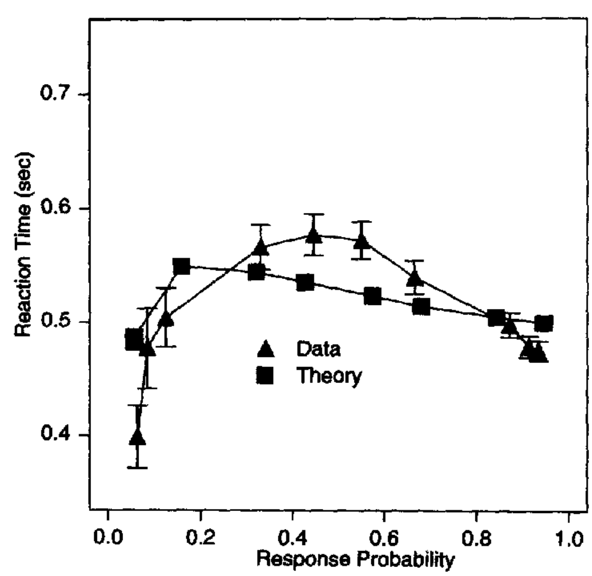

Figure 8. Latency probability functions for the data for the easy discrimination condition in Experiment 1 and fits of the diffusion model with drift linearly increasing up to the time of mask and then linearly decreasing to zero at the same rate.

fitting parameters of the nonstationary models are shown in Table 4.

One feature of the fits is the inability of the two nonstationary models to fit the fast extreme errors while fitting slow errors when accuracy is low. The model with constant drift up to mask presentation is not able to produce fast errors; the predicted error reaction times are slower than correct responses. The inverted-V drift model produces fast extreme errors relative to correct responses, but not fast enough to match the data. This is because the drift rate increases over the stimulus duration and then falls and does not reach zero until it is twice the SOA. In the extreme accuracy condition (84-ms SOA), the drift rate does not fall to zero until 168 ms after stimulus onset.

Although neither of the two nonstationary models we fit to the data gave a satisfactory account, some small nonstationarity in drift superimposed on constant drift (e.g., a rapid rise or a small and slow decrease in drift rate) might be able to fit the data, but the nonstationarity would have to be small relative to the constant drift component.

The nonstationary drift models have fewer parameters than the stationary drift model: The stationary drift model has one parameter for each drift rate, and the nonstationary model has only the point at which drift rate changes. Thus, one might conclude that the stationary drift model has more flexibility than the nonstationary models. As noted earlier, however, the shape of the latency probability function is determined by parameters other than drift rate (a parametric plot with the drift rate being the parameter of the function), so the comparisons using the latency probability function are made on the basis of the same number of parameters: boundary position, nondecision reaction time, and variability in drift and starting point of the process.

We used standard errors in the experimental data as our criteria for judging the quality of the fits. The diffusion model with constant drift was close to the experimental data, but the two nonstationary drift models missed substantially. Sophisticated methods of model evaluation will not be needed until alternative models are developed that fit the data within standard error criteria (i.e., as well as the stationary drift diffusion model).

Table 4
Parameters of the Fits of the Nonstationary Diffusion Models to the Data From the Easy Condition in Experiment 1

Parameter	Constant drift up to mask and then zero drift	Drift linearly increasing to mask then linearly decreasing
a	0.0650	0.0551
z	0.0325	0.0276
s_z	0.0080	0.0030
T_{er} (s)	0.437	0.461
s	0.1	0.1
η	0.096	0.102
Drift rate	0.697	2.455

Note. The drift rate for the linearly increasing drift is the maximum drift rate attained when the stimulus onset asynchrony is 84 ms and the drift rate then falls linearly to zero at the same rate.

From these fits to experimental data, we can conclude that the diffusion model with stationary drift is capable of fitting the data to (nearly) within standard error criteria. The two nonstationary models are not capable of fitting the data. Thus, the perceptual system is passing information at a constant rate to the decision process to produce a constant drift rate over the time course of the decision, and the size of the drift rate is determined by duration of the stimulus.

Experiment 2

In Experiment 2 we tested the stationary versus nonstationary hypotheses in a different way by using a deadline procedure. The task was letter matching with masking, like Experiment 1, but subjects were required to respond in advance of experimenter-determined deadline times.

Figure 9 shows the diffusion model predictions for the stationary and nonstationary hypotheses for growth of accuracy as a function of time. The assumptions are that subjects stop the decision process before the deadline time and that if the process is above the starting point, one response is produced, and if it is below the starting point, the other response is produced. Figure 9 shows predictions for growth of accuracy as a function of the time at which processing is terminated for three different SOA conditions. For stationary drift (the open symbols), the different SOAs are represented by three different values of constant drift: 0.05, 0.1, and 0.3. Accuracy rises gradually and monotonically with the rate of rise increasing with drift rate. For nonstationary drift (the black filled symbols), the different SOAs are represented by drift changing from 0.3 to 0 at 0.05, 0.1, and 0.2 s after the start of the decision process. Accuracy rises and then, at the time at which drift rate changes to 0, it begins to fall. Accuracy falls because when drift rate goes to 0, the diffusion process randomly drifts with the result that the accumulation of evidence sometimes ends up below the starting point at the wrong boundary (see Ratcliff, 1988, for a description of nonterminated processes as a function of time).

Method

Subjects. Six Northwestern University undergraduates participated in the experiment. Each was compensated $6 for each of six 35-min sessions.

Procedure. Stimuli were displayed in the same way as for Experiment 1. Two pairs of letters were used: E and Q, and L and P.

There were two variables in the experiment: SOA and deadline. The SOA between stimulus and mask was either long or short, set individually for each subject, or there was no mask. The deadlines were 300, 400, 525, or 2,000 ms.

The first two sessions for each subject were used for practice, and they were also used to set the SOA values for each subject. Two SOA values were chosen such that accuracy for the short one was between 0.6 and 0.7 and the long one was between 0.7 and 0.85. Calibration was done ad hoc using two SOAs for the first practice session, guessing appropriate SOAs for the second practice session from the data from the first practice session, and then using the results from both practice sessions for the final calibration. Across subjects, the short SOA ranged from 16 ms to 32 ms (M = 22 ms), and the long SOA ranged from 24 ms to 44 ms (M = 31 ms).

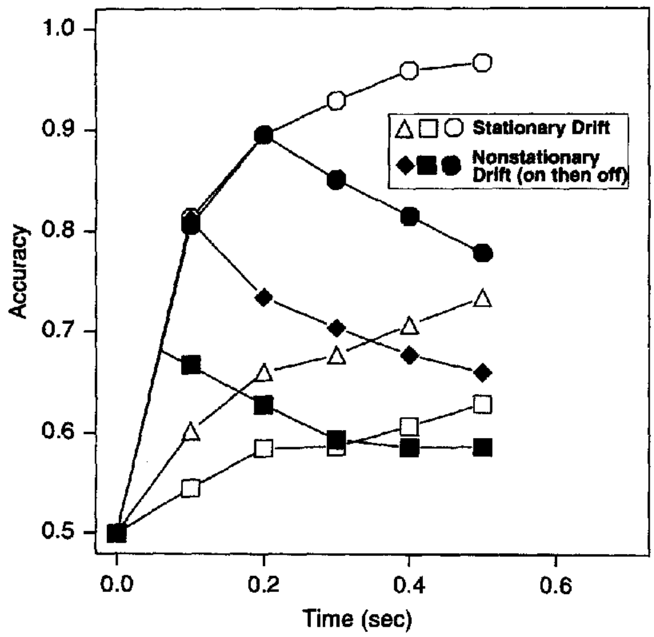

Figure 9. Predictions for the growth of accuracy in a deadline experimental procedure. The nonstationary drift rate is assumed to be constant at 0.3 up until the mask and then set to zero. Points at which drift rate is set to zero for nonstationary drift are 0.05, 0.1, and 0.2 s for the filled squares, diamonds, and octagons, respectively. For stationary drift, the drift rates are 0.3, 0.1, and 0.05 for open octagons, triangles, and squares, respectively. The standard deviation in drift rate between-trials η is 0.08, the standard deviation of within-trials variability s is 0.1, and nondecisional processing time T_{er} is 0.

The deadlines were imposed using feedback. At the beginning of each trial, the display consisted of a plus sign fixation point, the two letter alternatives, and the deadline (displayed in milliseconds). The plus sign was located in the center of the display, one response alternative was displayed above and to the left of the plus sign, and the other was displayed above and to the right. The deadline was displayed above the right alternative. This display lasted 500 ms, and the stimulus letter was then presented. If the letter was to be masked, the mask appeared after the appropriate SOA and remained on the screen until the subject made a response. If the letter was not masked, the letter remained on the screen until the response. Subjects pressed the *z* key if the stimulus was the left alternative and the "/" key if it was the right alternative. After the response, feedback was given. If the response was correct and occurred before the deadline, the message "correct" was displayed along with the response time. If the response occurred after the deadline, the message "too slow" was displayed.

Each session was made up of eight blocks of 100 trials; each block had equal numbers of the five SOA conditions in random order. The same deadline was used for all the trials of two consecutive blocks. Across all the blocks of the experiment, there were equal numbers of blocks for each deadline and the order of the pairs of blocks was randomly assigned except that the deadline was always switched from one of the two short deadlines (300 or 400 ms) to one of the two long deadlines (525 or 2,000 ms) or vice versa after every second block. All the trials of a block used the same pair of letters, and the pair was switched after every block. Subjects were given a short break after every block.

Results

Response times greater than 1,000 ms and less than 100 ms (less than 0.2% of the data) were eliminated from the analyses. Figure 10 shows accuracy plotted against response time. As the figure shows, the mean response times in each deadline condition were generally faster than the deadline, considerably so for the 525- and 2,000-ms conditions. Only for the 300-ms deadline was the mean response time slower than the deadline. The figure also shows that across the SOA conditions, mean response times at each deadline were about the same except at the 2,000-ms deadline, where responses in the no-mask condition were faster than in the mask conditions.

The main result is that the functions had exactly the shape predicted by the diffusion model with stationary drift. There was no hint of the sharp bend predicted by nonstationarity.

Figure 11 shows cumulative response time distributions based on mean quantiles (see Ratcliff, 1979). The distributions for the first three deadlines were collapsed over mask conditions (because there were no differences between them), as were the two mask conditions for the 2,000-ms deadline. But the 2,000-ms deadline no-mask condition differed from the two mask conditions and so is plotted separately. Mean quantiles were computed by ordering the response times, dividing them into 20 groups, and then taking the mean response time for each group. These means were then averaged over subjects. For comparison, the figure also shows a normal distribution with about the same standard deviation (53 ms) as that of the first three distributions plotted. The cumulative distributions are parallel to each other except for the mask conditions at the 2,000-ms deadline. The cumulative distributions are almost normally distributed, except that the slowest quantile is skewed a little. These distributions indicate that the deadlines induced

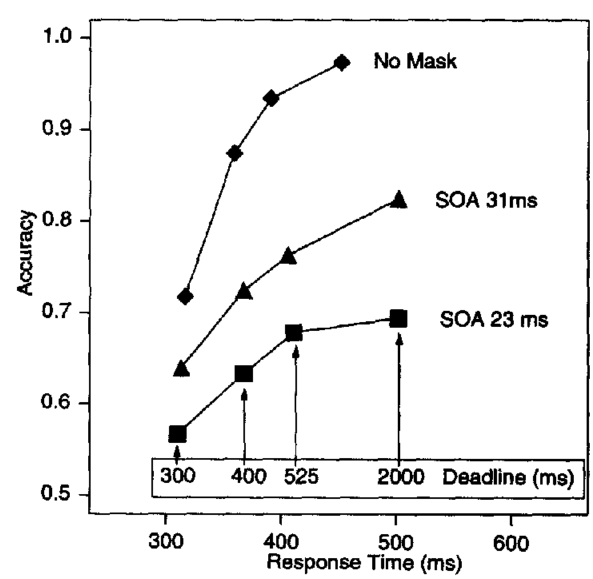

Figure 10. Accuracy as a function of deadline and stimulus onset asynchrony (SOA) from Experiment 2.

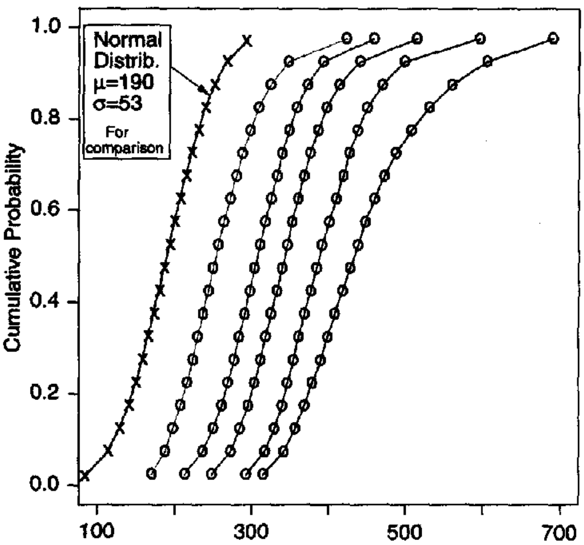

Figure 11. Cumulative response time (RT) distributions (Distrib.) for the data from Experiment 2. Circles represent quantiles for RT distributions for deadline conditions. The order from left to right is 300-, 400-, and 525-ms deadlines, followed by the no-mask and mask conditions for the 2,000-ms deadline.

subjects to stop processing with a stopping time that was nearly normally distributed.

To fit the data with the stationary drift diffusion model, we used the function derived by Ratcliff (1978) for the growth of d' as a function of time:

$$d'(t) = \frac{d'_a}{\sqrt{1 + \dfrac{s^2}{\eta^2(t - T_{er})}}}$$

This function is based on the assumption that the diffusion process evolves over time without boundaries. When the deadline is presented, a process above the starting point produces one response and a process below the starting point produces the other response. This gives a reasonable fit to experimental data from response signal methods (see Ratcliff, 1978; for a discussion of alternative functions, assumptions with models with response boundaries, and issues of mimicking, see Ratcliff, 1988).

We converted the accuracy values (probabilities) shown in Figure 10 to d' values using the accuracy value as the hit rate and one minus the hit rate as the false-alarm rate (because a hit for one target was one minus the false-alarm rate for the other target because response time data and accuracy for the two letter targets were symmetrical). We then fit Equation 1 to the d' values using least squares minimization with three asymptotic d'_a parameters (the difference in positive and negative drift rates divided by standard deviation in drift, η; see the top panel of Figure 2), one for each mask duration and another for the no-mask condition, one rate parameter (s^2/η^2), and the nondecisional time parameter T_{er}. The best fitting functions are shown in Figure 12, and the data and the parameter values are shown in Tables 2 and 3. The best fitting functions fall close to the empirical functions.

The parameter values obtained by fitting Equation 1 were consistent with those from Experiment 1; T_{er} was about the same and the asymptotic d'_a values fell between those for the hard and easy conditions in Experiment 1. η was larger than the two η values from Experiment 1, but it was not out of line with the ranges found in other experiments (see Ratcliff & Rouder, 1998; Ratcliff et al., 1999). Thus, the model shows reasonable parameter consistency across experimental procedures.

General Discussion

In this research we used the diffusion model framework to contrast two hypotheses about the information that enters the decision process in a two-choice masking task. The first hypothesis was that drift rate would increase and then decrease, reflecting the earlier onset of the stimulus and then onset of the mask. The second hypothesis was that drift rate would be constant over the course of decision making, with its value being a function of stimulus duration before mask onset. Differential predictions were generated from these hypotheses for the standard reaction time procedure used in Experiment 1 and for the deadline procedure used in Experiment 2. The nonstationary drift rate models predict slow error reaction times in the standard reaction task, with an approximately linear increase in reaction time from correct through to error responses and a sudden downturn for extreme errors. They also predict that accuracy should rise and then fall in the deadline task. In contrast, for the standard reaction time task, the stationary drift rate model

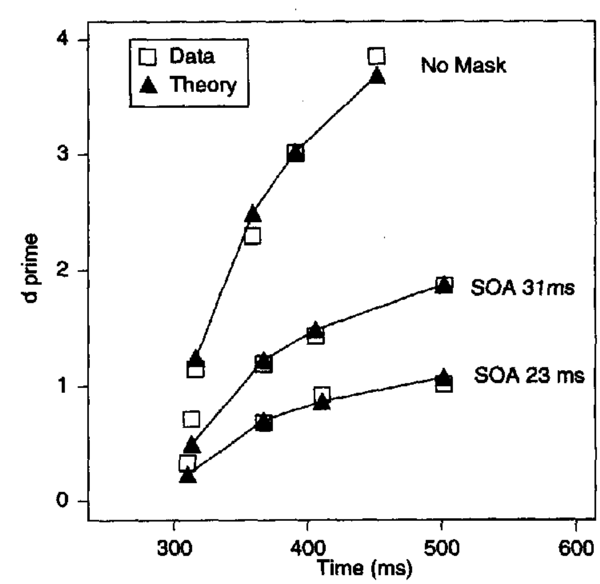

Figure 12. Fits of the stationary drift diffusion model (without response boundaries) to the data from Experiment 2. SOA = stimulus onset asynchrony.

predicts errors only a little slower than correct responses at intermediate levels of accuracy and errors faster than correct responses at extreme levels of accuracy. For the deadline task, stationary drift predicts a monotonic growth of accuracy. The patterns of data obtained in the experiments matched those predicted by stationary drift; the data showed neither the pattern of slow errors nor the nonmonotonic growth of accuracy predicted by nonstationary drift. Whether the finding of stationary drift rate will extend to other domains is an open question. For example, it might extend to word identification paradigms because words, like letters, can be represented categorically. However, it might not extend to perceptual stimuli (cf. luminous squares on random dot patterns; Smith, 1995) for which a cognitive representation may not be available to serve as the output of perceptual processing.

With the stationary drift rate assumption, the diffusion model provided a good quantitative fit to the data. The masking manipulation allowed accuracy to be varied from near ceiling to near floor, which in turn allowed mapping of full latency–response probability functions and provided a range of reaction time distributions (see Figure 10). The crossovers of error versus correct response times, in conjunction with the latency probability functions, were stringent tests of the model.

Not every nonstationary model is ruled out by the experimental results. For example, suppose that the drift rate started at zero and rapidly rose after stimulus onset until mask onset and then remained constant for short stimulus presentation times. Such a model would not produce predictions that are discriminable from the stationary model presented here. (In fact, variability in the starting point can be considered to be the result of processing before stimulus onset; see Laming, 1968, for a premature sampling followed by a step function in drift rate.)

The diffusion model provides a good quantitative fit to the sets of data from the individual experimental conditions with few free parameters compared with the number of degrees of freedom in the data (which included accuracy values, correct and error reaction times, and the shapes of reaction time distributions). It also provides consistent estimates of parameter values across conditions and experiments. Although the data from the hard and easy discrimination conditions of Experiment 1 were fitted separately, the estimates of the parameters that would not be expected to be different as a function of the stimulus difficulty (boundary separation and the encoding and response parameter, T_{er}) were within a few percentage points of each other. The value of variability in starting point (s_z) was small for the easy and difficult conditions in Experiment 1, so small that the difference (0.018 vs. 0.014) would not produce significant changes in the quality of the fits. The d'_a values from Experiment 2 fell in the same range as those for the hard and easy conditions of Experiment 1. Furthermore, for all the parameters, the values from Experiments 1 and 2 are in the range of the values that have been found in related paradigms (Ratcliff & Rouder, 1998; Ratcliff et al., 1999). These parameter invariances provide strong support for the model.

The good fits to the data from Experiments 1 and 2 also support the diffusion model by extending the model to deal with a kind of limitation on performance that had not previously been addressed by the model, namely variability resulting from impoverished stimulus encoding. Varying mask onset time varied the quality of the information from the stimulus (i.e., it reduced performance so that variability from perceptual processes operating on a brief stimulus determined the noise in processing, such that noise was internal to the stimulus). In previous research, the diffusion model has been tested with noise external to the stimulus. Ratcliff and Rouder (1998) and Ratcliff et al. (1999; see also Ashby & Gott, 1988; Espinoza-Varas & Watson, 1994; Lee & Janke, 1964) varied external noise by varying the probability with which the correct response to a stimulus was one alternative versus another. For example, in the Ratcliff et al. experiments, subjects were asked to decide whether the number of asterisks in a display was "high" or "low." The number of asterisks was drawn from one of two overlapping distributions: the low distribution with mean 38 and the high with mean 56 (SDs = 14). After the response to each stimulus, subjects were given feedback about from which distribution the stimulus had been drawn. Subjects could not be perfectly correct in their responses because any given stimulus might have been drawn from either distribution. A display of 20 asterisks, for example, would usually have been drawn from the low distribution, but it might also have been drawn from the high distribution. In this situation, the spread of accuracy from floor to ceiling is accomplished by external variability; it is the variability in the probability of which is the correct response to the stimulus not the quality of the stimulus itself that leads to errors. Numbers of asterisks near 50 have low accuracy because it is about equally likely that they came from the high versus low distributions; numbers of asterisks near 0 or 100 have high accuracy because they are much more likely to have come from one of the distributions than the other. The diffusion model does a good job of quantitatively fitting the data from this paradigm as well as similar paradigms using external variability, such as red–green discrimination, auditory discrimination, and brightness same–different judgments. With the new experiments presented here, the diffusion model is extended to account for data from conditions in which perceptual information is impoverished and the limitation in processing is internal variability.

In quantitatively fitting empirical data, the diffusion model provides a measure of d'_a for the information that enters the decision process in each experimental condition (see the top panel of Figure 2), a measure of d'_a that is different in an important way from the standard d' measure. The diffusion model's d'_a measure, which is drift rate divided by the standard deviation in drift rate, combines data from reaction time and accuracy into a single measure (by applying the diffusion model to the data), unlike the standard measure of d', which is based on accuracy alone. The standard measure is flawed because it does not take into account the possibility of speed–accuracy interactions. In most tasks, if subjects are instructed to be as accurate as possible, then their accuracy will increase relative to an uninstructed baseline, leading to a larger value of d'. On the other hand, if they are instructed to respond as quickly as possible, their accuracy will decrease, leading to a smaller value of d'. In other words, d' computed from accuracy alone is not invariant across speed–accuracy criteria settings. This is not reasonable; d' should measure the information entering the decision process, and this information should be invariant across criteria settings. The diffusion model provides a method for extracting such a measure (see Ratcliff & Rouder, 1998, Experiment 1).

The research presented in this article focused on the decision process and did not deal directly with perceptual processes beyond hypotheses about the time course of availability of perceptual information to the decision process. Most of the research on perceptual processing and masking deals with processing at input and encoding. We emphasize that none of the earlier research has produced a model that can fully account for the data that the paradigms produce. Any response made in these paradigms has a response time associated with it, and when responses can be errors, there are also error response times. Thus, we argue, the field should begin to work with models of the time course of how decisions are made and link them to the models of perceptual processing. Our research provides one example of this strategy (for other examples, see Rouder, 1995; Smith, 1995; Smith & Vickers, 1988) and shows that within the framework of the diffusion model, we can say that the information provided to the decision process is constant (or nearly constant) over time.

Because the diffusion model fits the experimental data well, it is natural to ask whether the model is falsifiable. In fact, it is remarkably well constrained and easily falsified. One simple way of seeing this is that the diffusion model with nonstationary drift was easy to falsify. Furthermore, the model is highly constrained by the shapes and locations of the reaction time distributions. For example, a change in drift rate alone changes the leading edge of the reaction time distribution by little relative to the spread in the tail. Therefore, any factor that alters the leading edge significantly without also increasing the spread falsifies a model in which drift rate alone accounts for the effect of the factor.

The success of the diffusion model sets standards for competing models of the decision process. They should account for accuracy, reaction times for correct and error responses, and the shape of reaction time distributions as well as the growth of accuracy as a function of time in procedures using response signals or deadlines. The usefulness of such modeling is exemplified by the findings from the experiments presented here: We conclude that the processing of a letter stimulus when the letter is masked after brief presentation involves integrating the stimulus information to provide a constant value (categorical representation?) of perceptual output.

References

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 33–53.

Audley, R. J., & Pike, A. R. (1965). Some alternative stochastic models of choice. *British Journal of Mathematical and Statistical Psychology, 18,* 207–225.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach.* New York: Oxford University Press.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review, 83,* 1–36.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547.

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences, 23,* 255–282.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459.

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review, 101,* 446–469.

Dixon, P., & Di Lollo, V. (1994). Beyond visible persistence: An alternative account of temporal integration and segregation in visual processing. *Cognitive Psychology, 26,* 33–63.

Espinoza-Varas, B., & Watson, C. (1994). Effects of decision criterion on latencies of binary decisions. *Perception & Psychophysics, 55,* 190–203.

Feller, W. (1968). *An introduction to probability theory and its applications.* New York: Wiley.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Heath, R. A. (1981). A tandem random walk model for psychological discrimination. *British Journal of Mathematical and Statistical Psychology, 34,* 76–92.

Heath, R. A. (1992). A general nonstationary diffusion model for two-choice decision-making. *Mathematical Social Sciences, 23,* 283–309.

Kahneman, D. (1968). Methods, findings, and theory in studies of visual masking. *Psychological Bulletin, 70,* 404–425.

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

Lee, W., & Janke, M. (1964). Categorizing externally distributed by stimulus samples for three continua. *Journal of Experimental Psychology, 68,* 376–382.

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

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

Loftus, G. R., & Ruthruff, E. (1994). A theory of visual information acquisition and visual memory with special application to intensity-duration trade-offs. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 33–49.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308–313.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461.

Ratcliff, R. (1980). A note on modelling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology, 21,* 178–184.

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. (1988). Continuous versus discrete information processing: Modeling the accumulation of partial information. *Psychological Review, 95,* 238–255.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9,* 347–356.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review, 102,* 261–300.

Rouder, J. N. (1995). *Assessing the roles of change discrimination and accumulation: Evidence for a hybrid race model of perceptual decision making in luminance discrimination.* Unpublished doctoral dissertation, University of California, Irvine.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review, 102,* 567–591.

Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology, 32,* 135–168.

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

Strayer, D. L., & Kramer, A. F. (1994). Strategies and automaticity: 1. Basic findings and conceptual framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 318–341.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 80,* 1–52.

Van Zandt, T., & Ratcliff, R. (1995). Statistical mimicking of reaction time distributions: Mixtures and parameter variability. *Psychonomic Bulletin and Review, 2,* 20–54.

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

Vickers, D., Caudrey, D., & Willson, R. J. (1971). Discriminating between the frequency of occurrence of two alternative events. *Acta Psychologica, 35,* 151–172.

Von Brisinski, I. S. (1994). Ultrafast display buildup with standard VGA on MS-DOS computers. *Behavior Research Methods Instruments and Computers, 26,* 335–336.

Received May 18, 1998
Revision received January 7, 1999
Accepted February 17, 1999 ■