

## Temporal Properties of Human Information Processing: Tests of Discrete versus Continuous Models

DAVID E. MEYER, STEVEN YANTIS, ALLEN M. OSMAN,  
AND J. E. KEITH SMITH

*University of Michigan*

Cognitive psychologists have characterized the temporal properties of human information processing in terms of discrete and continuous models. Discrete models postulate that component mental processes transmit a finite number of intermittent outputs (quanta) of information over time, whereas continuous models postulate that information is transmitted in a gradual fashion. These postulates may be tested by using an adaptive response-priming procedure and analysis of reaction-time mixture distributions. Three experiments based on this procedure and analysis are reported. The experiments involved varying the temporal interval between the onsets of a prime stimulus and a subsequent test stimulus to which a response had to be made. Reaction time was measured as a function of the duration of the priming interval and the type of prime stimulus. Discrete models predict that manipulations of the priming interval should yield a family of reaction-time mixture distributions formed from a finite number of underlying basis distributions, corresponding to distinct preparatory states. Continuous models make a different prediction. Goodness-of-fit tests between these predictions and the data supported either the discrete or the continuous models, depending on the nature of the stimuli and responses being used. When there were only two alternative responses and the stimulus-response mapping was a compatible one, discrete models with two or three states of preparation fit the results best. For larger response sets with an incompatible stimulus-response mapping, a continuous model fit some of the data better. These results are relevant to the interpretation of reaction-time data in a variety of contexts and to the analysis of speed-accuracy trade-offs in mental processes. © 1985 Academic Press, Inc.

---

Preparation of this article was supported by grants to D. E. Meyer from the National Institute of Mental Health (R01 MH 38845) and the Horace Rackham School of Graduate Studies. The laboratory used in the project was also supported by a Biomedical Research Grant from the National Institutes of Health to the Vice President for Research at the University of Michigan. S. Yantis participated under a National Science Foundation graduate fellowship. Portions of the research were first reported at the Franco-American Conference on Preparatory States and Processes and the meetings of the Psychonomic Society and Midwestern Psychological Association (Meyer, Yantis, Osman, & Smith, 1984; Meyer, Osman, & Yantis, 1982; Yantis, Osman, & Meyer, 1982). We thank D. Frey, T. Recker and D. Tang for technical assistance. Suggestions by E. Hunt, J. Johnston, K. Koh, S. Kornblum, J. Miller, S. Sternberg, J. Townsend, and an anonymous reviewer are also gratefully acknowledged. Requests for reprints should be sent to David E. Meyer, Human Performance Center, Department of Psychology, University of Michigan, 330 Packard Rd., Ann Arbor, MI 48104.

In formulating theories of human information processing, an important goal is to characterize the temporal properties of mental processes such as stimulus encoding, memory retrieval, decision making, and response preparation. The characterization should address a number of related questions: What is the duration of each component process? How are the individual durations combined to determine the overall reaction time? Do processes start at the same moment or at different moments? Must one process finish before another can be initiated and/or completed? Does each process produce a single discrete packet of information as its output, or is the flow of information continuous?

Efforts to answer such questions began with the birth of experimental psychology (e.g., Donders, 1868–1869) and have persisted ever since (e.g., Smith, 1968; Sternberg, 1969). One popular approach has been to measure reaction time, that is, the time taken by a subject to react overtly to a presented stimulus. It is commonly accepted that reaction times may reveal whether mental processes proceed sequentially or concurrently and whether they produce discrete or continuous outputs (Pachella, 1974). As a result, much has been learned regarding perception, memory, and cognition, but considerable uncertainty still remains about the temporal properties of human information processing (McClelland, 1979; Ollman, 1977; Sternberg, 1969; Townsend, 1974; Townsend & Ashby, 1983; Wickelgren, 1977).

The uncertainty stems partly from the fact that reaction times provide only indirect reflections of the durations of component mental processes. Because these durations are not directly observable, patterns of factor effects on reaction times are open to multiple theoretical interpretations. For example, Sternberg (1969) has argued that if two experimental factors affect reaction time additively, then one may infer that there are at least two temporally separate, discrete stages of processing. This logic formed the basis of his well-known additive-factor method, which has been applied in a variety of psychological domains. However, McClelland (1979) has shown that under some plausible circumstances, two or more factors can affect reaction time additively even though they influence simultaneously active processes whose outputs are continuous. He has also shown that two factors can have interactive effects on reaction time even though they do not influence the same process. The latter demonstration conflicts with a corollary of the additive-factor method, which asserts that if two factors affect reaction time interactively, then they influence at least one processing stage in common. Given such conflicts, it is difficult to reach firm conclusions about the temporal properties of information processing.<sup>1</sup>

<sup>1</sup> The difficulty is illustrated by the work of Meyer, Schvaneveldt, and Ruddy (1975), who reported that lexical (word–nonword) decisions were faster for pairs of semantically

The difficulty is aggravated by possible speed-accuracy trade-offs (Pachella, 1974). Subjects may reduce their reaction time at the expense of committing more errors, or they may reduce their error rates at the expense of taking more time to react. Without strong and perhaps unjustified assumptions about the exact quantitative form of the trade-off, it becomes more complicated to evaluate how much the duration of a particular process is affected by some factor and to determine the temporal relations among various processes. Increasing doubts have arisen about whether there even exists a "true" reaction time that corresponds to a moment when all information processing is finished after the onset of a test stimulus. This has led some investigators to suggest that conventional reaction-time procedures should be abandoned and that full speed-accuracy trade-off curves should be obtained instead (Ollman, 1977; Wickelgren, 1977; cf. Schmitt & Scheirer, 1977).<sup>2</sup>

Nevertheless, there are still efforts being made to study the temporal properties of mental processes by measuring reaction times. For example, we have combined preemptive response signals with a conventional reaction-time procedure to reach inferences about the time course of stimulus encoding and memory retrieval (Meyer & Irwin, 1981). Our results indicate that these processes may function as discrete stages during word recognition, supporting some previous conclusions (e.g., Meyer, Schvaneveldt, & Ruddy, 1975) made about the recognition process. Moreover, Miller (1982, 1983) has manipulated stimulus-response compatibility to reach conclusions about the time course of response preparation. He found evidence that the preparation process can start before other processes have completely finished, but that this process receives quantal packets of partial information as inputs and proceeds through a finite sequence of intermediate states. The outcome of such efforts suggests that reaction times may continue to yield significant insights about information-processing dynamics.

---

related words (e.g., BREAD-BUTTER) than for pairs of unrelated words (e.g., NURSE-BUTTER). More facilitation occurred when the second word of a pair was visually degraded than when it was displayed intact, producing a legibility-relatedness interaction. Following Sternberg's (1969) additive-factor method, Meyer et al. (1975) inferred that stimulus legibility and semantic relatedness both influence an early discrete perceptual-encoding stage. McClelland (1979), on the other hand, argued that the legibility-relatedness interaction could be explained better in terms of separate factor effects on two functionally different, but simultaneously active, processes such as encoding and memory retrieval.

<sup>2</sup> For example, Wickelgren (1977, pp. 83-84) claimed that "In those cases where the focus is on testing theories regarding information processing dynamics, speed-accuracy tradeoff [SAT] experiments are far superior to reaction time experiments. . . . The increased power of [the SAT] method for deciding between theories of information processing dynamics should be obvious . . . the prospects of achieving cumulative scientific progress in quantitative studies of cognitive dynamics seems so much greater when one obtains the entire SAT function that such data should often be well worth the extra time and expense."

The present article deals further with the temporal properties of mental processes that underlie response preparation. As in past research, measures of reaction time constitute a principal source of evidence here. However, some useful extensions are introduced in collecting and evaluating the data. Our approach incorporates an adaptive response-priming procedure and a quantitative analysis of entire reaction-time distributions (Meyer, Yantis, Osman, & Smith, 1984). With this procedure and analysis, we specifically address the questions of whether component mental processes produce discrete or continuous outputs of information over time and whether response preparation involves a finite or infinite number of intermediate preparatory states.<sup>3</sup>

In what follows, we first survey some relevant alternative models of information processing. Next we outline the basic steps of our response-priming procedure and reaction-time analysis. Then we report some illustrative experiments and results. Our aim is to demonstrate the potential of this approach and to help stimulate future research on information-processing dynamics. The present results do not constitute definitive evidence about the temporal properties of human information processing under all circumstances, but they do place some additional constraints on viable models. For example, it appears that the form of the outputs by the mental processes leading to response preparation and the number of intermediate preparatory states associated with them depend significantly on the number of alternative responses, the difficulty of each response, and the compatibility of assigned stimulus-response mappings. A simple mapping with relatively few response alternatives may tend to induce discrete processing and limit how many intermediate preparatory states there are, whereas a complex mapping with numerous response alternatives may tend to induce continuous processing and more intermediate states.

### MODELS OF INFORMATION PROCESSING

Two broad classes of information-processing models are most relevant here. One class consists of discrete finite-state models, and the other consists of continuous models. Each class incorporates a number of functionally distinct mental processes (e.g., encoding, retrieval, and decision), which accept inputs, transform information in various ways, and produce

<sup>3</sup> Our work also bears at least indirectly on the question of whether two or more component processes may be simultaneously active in a system. In particular, to the extent that information appears to be output continuously, this would suggest that both the source and the recipient of the information can function as simultaneous, although perhaps contingent, processes (McClelland, 1979). By contrast, if a set of component processes produces only a small number of discrete outputs and yields few or no intermediate preparatory states, then one might plausibly infer that the processes do not overlap in time (cf. Miller, 1982).

outputs to be used by other parts of the system (Smith, 1968). The primary differences between the two classes concern their assumptions about how various processes are related temporally and what forms the outputs have.

### *Discrete Models*

Discrete models assume that mental processes produce a finite number of distinct information outputs (quanta) over time. With this form of output, a processing system may enter and exit one or more intermediate preparatory states as information passes through the system, but the number of different states is limited. Here we discuss two representative cases in detail: the discrete stage model (Sternberg, 1969), and the asynchronous discrete-coding model (Miller, 1982, 1983). For other relevant examples, see Broadbent (1958, 1971), Donders (1969/1868–1869), and Schweickert (1978, 1983).<sup>4</sup>

*Discrete stage model.* According to the discrete stage model, there is no temporal overlap among various mental processes. Rather, the processes occur in a strict sequence such that each process (stage) starts only after the immediately prior one has finished. It is further assumed that a stage of processing produces a single discrete quantum of information whose quality does not depend on how long the process takes to finish and that, correspondingly, there is only a finite number of intermediate preparatory states before the response to a stimulus occurs.<sup>5</sup>

The discrete stage model has important implications for the interpretation of reaction-time data. If the model is correct, then the overall reaction time observed on a trial equals the sum of the durations of component mental processes (Donders, 1969/1868–1869; Sternberg, 1969), and the effect of a factor on reaction time equals the total amount by which the factor changes those durations. It was these considerations

<sup>4</sup> It should be noted that the above informal dichotomy glosses over a third logically possible class of models, namely, ones with a countably infinite number of distinct outputs and intermediate preparatory states. While these additional models may be of some theoretical interest, and are discrete in a technical sense, we cannot distinguish them empirically from members of the continuous class. Thus, in subsequent exposition, the term “discrete” is restricted to those cases with a relatively small number of distinct outputs and intermediate states.

<sup>5</sup> There may be a positive correlation between the number of preparatory states and the number of stages in a system. For example, stage models of word recognition that include a process of tentative stimulus identification and a subsequent process of stimulus verification could yield at least two higher states of preparation following an initial unprepared state (Meyer & Irwin, 1981). The number of higher states would not necessarily equal the number of stages, however, because the outputs of some stages may not be sufficient by themselves to allow any useful changes in preparation for later events. Consequently, in some discrete stage models, as few as two distinct preparatory states are possible.

that motivated Sternberg's (1969) development of the additive-factor method. Reasoning backward from them, he proposed that the existence of discrete, temporally separate, processing stages may be inferred whenever two or more factors have additive effects on reaction time (Sternberg, 1969, p. 282). This logic admittedly has loopholes, because conclusions based on affirming the consequent of an implication (i.e., if  $p$ , then  $q$ ;  $q$ ; therefore  $p$ ) are not necessarily valid. However, it is plausible in that additive effects seem most likely to occur when two or more factors influence different discrete stages.

The discrete stage model is also relevant for assessing and controlling speed-accuracy trade-offs. If subjects process information through a series of temporally separate stages, each of which produces a single discrete output, then they would have relatively few options available to trade accuracy and speed. Decision making and response preparation could not be readily based on partial outputs from early stages to generate better-than-chance responses before information processing is complete. This limits the form of the trade-off function to one involving a random fast guess or a deadline strategy (Ollman, 1966, 1977; Ollman & Billington, 1972; Yellott, 1967, 1971). In particular, the trade-off might be either linear (Ollman, 1966; Yellott, 1967, 1971) or isomorphic to the cumulative distribution function of reaction times from trials on which guessing does not occur (Meyer & Irwin, 1981; Schmitt & Scheirer, 1977). The resulting constraint would facilitate inferences about the duration of completed processing, leading to better measures of "true" reaction time (cf. Pachella, 1974). Thus, it is especially important to determine when, and under what circumstances, discrete stage models may hold in preference to other alternatives that impose weaker constraints.

*Asynchronous discrete-coding model.* Another member of the discrete class is Miller's (1982, 1983) asynchronous discrete-coding (ADC) model. Like the discrete stage model, it assumes that mental processes produce intermittent quanta of information, yielding a finite number of intermediate preparatory states. However, the ADC model relaxes some of the basic assumptions of the discrete stage model.

According to the ADC model, a component process may produce more than one output as a stimulus is evaluated over time, and two or more processes may be simultaneously active, even though one's output provides the other's input. "Partial information about a stimulus can be transmitted . . . when the information is complete with respect to an internal perceptual code" (Miller, 1982, p. 292). In particular, Miller (1982) proposed that the process of response preparation receives multiple inputs from processes that start before it but that do not finish until after some preparation has already occurred. This proposal implies that the overall reaction time would not be the sum of the durations of the

underlying component processes, contrary to the discrete stage model. Nor would the observed amount by which a factor affects reaction time necessarily reflect how much that factor increases or decreases the total duration of those processes, contrary to Sternberg's (1969) additive-factor method.

Miller's (1982) ADC model also imposes fewer constraints than the discrete stage model does on the possible form of the speed-accuracy trade-off. With multiple outputs from intermediate processes, a subject might trade accuracy for speed not only by making random fast guesses but also by making more-sophisticated guesses based on available partial information. The increased flexibility would diminish the prospects of measuring "true" reaction times associated with complete processing of a stimulus, although not as much as continuous models do.

### *Continuous Models*

Continuous models assume that information is transmitted between component mental processes in a gradual fashion.<sup>6</sup> The transmission involves more than just a finite or countably infinite sequence of intermittent discrete packets. Response preparation may therefore vary smoothly over time as other processes that provide inputs to it progress from start to finish.

Consequently, prototypical members of the continuous class differ from the discrete stage model to an even greater extent than does the ADC model. There are no implications that reaction time should equal the sum of the durations of underlying component processes or that overall factor effects on reaction time should reflect the total amount by which a factor changes the duration of those processes. Under a continuous model, subjects may have virtually unlimited flexibility in trading speed and accuracy as a function of personal preferences and/or task demands (Ollman, 1977; Wickelgren, 1977). Any level of intermediate output could be used to generate sophisticated guesses, depending on the relative importance of fast reaction times versus low error rates. Strong empirical evidence of a continuous information flow would constitute a severe blow to experimenters who seek a measure of "true" reaction time (Pachella, 1974).

*Cascade model.* An illustrative member of the continuous class is the cascade model (McClelland, 1979). According to this model, a sequence of functionally distinct but simultaneously active processes leads from

<sup>6</sup> Specifically, the change must have the formal properties of a continuous mathematical function that relates the information output (dependent variable) to the length of time since processing started (independent variable). It is not necessary that the change involve a strict monotonic increase. Periods of time could occur during which the output remains constant. In at least some well-known cases, however, there is strict monotonicity (e.g., the cascade model; McClelland, 1979).

stimulus input to response output. The flow of information between these processes is represented quantitatively by a negatively accelerated activation curve with an exponential rate parameter and a positive asymptote. The current output of a process determines the asymptote that can be reached by its immediate recipient. When a stimulus is presented, activation levels of all processes increase gradually until a threshold is crossed in a final response unit, evoking an overt response.

The cascade model provided a concrete theoretical framework for McClelland's (1979) criticisms of the additive-factor method. He noted that if two factors influence the rate parameters of two different processes in the model, then they could have additive effects on observed reaction time, even though the processes overlap temporally. It would not be justified to infer that there are nonoverlapping, discrete stages. Also, if one factor influences the rate parameter of some process in cascade, but another factor influences the asymptote parameter of a different process, then the two factors could have interactive effects on reaction time, even though they still influence functionally distinct processes. It would not be justified to infer that the effects occur within the same process (stage).<sup>7</sup> Thus, before applying the additive-factor method or dismissing possible artifacts due to speed-accuracy trade-offs, an investigator should have some principled theoretical or empirical bases for assessing the cascade model, as we consider here.

*Other examples.* There are also other influential examples of the continuous class that merit attention. These include concurrent-contingent models (Eriksen & Schultz, 1979; Turvey, 1973), random-walk models (Edwards, 1965; Link, 1975; Ratcliff, 1978; Stone, 1960), and logogen models (Morton, 1969). Perhaps the most general case is the interactive-activation model (McClelland & Rumelhart, 1981). It allows information to be transmitted continuously in multiple directions through a processing system, so that functionally superordinate processes can send concurrent feedback to subordinate processes and receive feedforward from them. The amount of interaction in this model extends beyond what is possible even under the cascade model.

## EMPIRICAL APPROACH

Our empirical approach to testing discrete versus continuous models builds on previous studies of stimulus and response-priming effects. A

<sup>7</sup> In essence, this account violates the uniform-quality assumption of the discrete stage model and the additive-factor method, which presume that the quality of the output by a process does not depend on the process's duration. Thus, it is not surprising that the rationale of the additive-factor method fails to hold under the cascade model (Sternberg, 1969, p. 283).



number of investigators have demonstrated that prime stimuli can significantly affect the latency of responses to subsequent test stimuli. Facilitation or inhibition may occur because a prime stimulus alters temporal uncertainty about when a test stimulus will appear (Bertelson, 1967), spatial uncertainty about where the test stimulus will appear (Posner, Snyder, & Davidson, 1980), and conceptual uncertainty about what the identity of the test stimulus will be (LaBerge, Van Gelder, & Yellot, 1970; Meyer & Schvaneveldt, 1971, 1976; Neely, 1976, 1977; Posner & Snyder, 1975). The prime stimulus can also provide specific information to program a required response, influencing the selection of the effector, direction, and extent of a movement (Leonard, 1958; Miller, 1982, 1983; Rosenbaum, 1980; Rosenbaum & Kornblum, 1982). Typically, such effects approach an asymptote as the temporal interval between the onsets of the prime and test stimuli increases (Bertelson, 1967; Leonard, 1958; Neely, 1976; Posner & Snyder, 1975).

#### *An Adaptive Response-Priming Procedure*

Mindful of these findings, we have developed an adaptive response-priming procedure for determining if and when information processing is discrete or continuous (Meyer et al., 1984). This procedure incorporates a series of reaction-time trials on which a subject is presented with an informative prime stimulus followed by a test stimulus. Overt responses are required to the test stimulus but not to the prime stimulus. Reaction time is measured as a function of the type of prime stimulus, the type of test stimulus, and the length of the interval between the onsets of the prime and test stimuli. A staircase tracking algorithm adjusts the priming interval (stimulus onset asynchrony; SOA) to maximize the diagnosticity of the results. Hence, the procedure is termed "adaptive."

The experiments reported below included three basic priming conditions: unprimed, partially primed, and completely primed, which differed with respect to the duration of the priming interval and/or type of prime stimuli being used. In the completely primed condition, the prime stimuli provided valid predictive information about the subsequent test stimuli and responses, and the priming interval was long enough to allow full processing of the primes. In the partially primed condition, the prime stimuli were valid, but the priming interval had a medium duration, permitting only an intermediate amount of processing to be achieved before the onset of the test stimulus. In the unprimed condition, we presented the valid primes for a very short duration or replaced them with neutral primes, which provided no information about the test stimuli and responses. Across the three priming conditions, the subjects' degree of preparation thus varied systematically, ranging from highly prepared to

unprepared. Observed performance was compared against alternative quantitative predictions derived from discrete and continuous models, thereby assessing whether or not processing of the prime stimulus involves a small finite set of distinct preparatory states.

#### *Ancillary Assumptions*

This application of the adaptive response-priming procedure requires certain assumptions about subjects' strategies in processing the prime and test stimuli. We assume that any available partial information derived from a valid prime stimulus is used to prepare for the subsequent test stimulus. However, processing of the prime is assumed to stop at the onset of the test stimulus. It is also assumed that if no partial information becomes available from processing a prime stimulus, then at the onset of the test stimulus, the processing of the test stimulus starts from scratch, as if no prime had been presented at all. Our experiments were designed to help ensure the validity of these assumptions, and some support for them may be found in the obtained results.

### PREDICTIONS OF THE MODELS

The major predictions examined here concern mixture distributions of reaction times. Because of their simplicity and elegance, such distributions have already played a significant role in theoretical descriptions of visual masking, memory search, and binary choice-reaction time (Eriksen & Eriksen, 1972; Falmagne, 1965, 1968; Sternberg, 1973; Theios & Smith, 1972; Townsend & Ashby, 1983). Our analyses broaden this role to more general tests of discrete and continuous models.

#### *Mixture Distributions*

Mixture distributions are produced by sampling probabilistically from a set of underlying basis distributions (Everitt & Hand, 1981). For example, suppose that we have two normal distributions with means  $\mu_1$  and  $\mu_2$  and that independent observations are drawn with probability  $\pi_1$  from the first of these distributions and with probability  $\pi_2$  from the second, where  $\pi_1 + \pi_2 = 1$ . Then the overall set of observations will have a mixture distribution with the two original normal distributions as the bases. More generally, a mixture distribution may have a total of  $n$  basis distributions ( $n \geq 2$ ) with probabilistic sampling weights ( $\pi_i; i = 1, 2, \dots, n$ ) that are nonnegative and sum to one.

The relevance of mixture distributions for present purposes is straightforward. Discrete models predict that the unprimed, partially primed, and completely primed conditions of the adaptive response-priming procedure should yield reaction times that have mixture distributions with a finite number of bases. Under these models, each basis distribution corre-

sponds to a different state of response preparation. There may be an unprepared state, a fully prepared state, and other distinct intermediate preparatory states. Manipulating the length of the priming interval and the type of prime stimuli affects the probability of achieving a particular state. The state of preparation achieved on a trial then uniquely determines the distribution from which a subject's reaction time comes, thus constraining the data in a quantitatively specifiable fashion.

At least some continuous models make an alternative (nonmixture) prediction. According to them, manipulating the interval between a prime and test stimulus should produce gradual changes in a subject's degree of preparation over time, reflecting the extent to which information conveyed by the prime stimulus has been processed before the test stimulus appears. The reaction times observed in the unprimed, partially primed, and completely primed conditions would not come from a family of mixture distributions based on a finite number of preparatory states. Instead, the various priming conditions would yield reaction times whose distributions have a different relation to each other, as discussed below.

To illustrate these predictions, we focus closely on two extreme cases: a two-state version of the discrete stage model (Sternberg, 1969) and a prototypical version of the cascade model (McClelland, 1979). These models exemplify the marked differences in predictions by some members of the discrete and continuous classes. We also briefly consider predictions by other specific models. However, it will not be possible to deal with all individual cases at length, since each class contains many members. Rather, our intent is to treat them at the most general level possible.

### *Two-State Discrete Stage Model*

According to the two-state discrete stage model, the processing of a prime stimulus includes discrete stages like encoding, retrieval, and decision, which lead to a discriminative output about the identity of the prime. This output then serves as an input to a response-preparation stage that adjusts the information-processing system for a rapid response to a subsequent test stimulus. Here preparation functions as a unitary all-or-none process that either places the system in a fully prepared state or leaves the system entirely unprepared. Essentially, this model embodies the theoretical proposals that some investigators (e.g., Meyer et al., 1975; Sternberg, 1969) have made about the performance of word-recognition and memory-search tasks.<sup>8</sup>

<sup>8</sup> Of course, other discrete stage models might also be applicable under our procedure. One could instead postulate several substages of preparation, each with its own individual output, such as various perceptual adjustments for processing the stimuli and motor adjustments for selecting the effector, the direction, and the extent of movement (e.g., Ro-

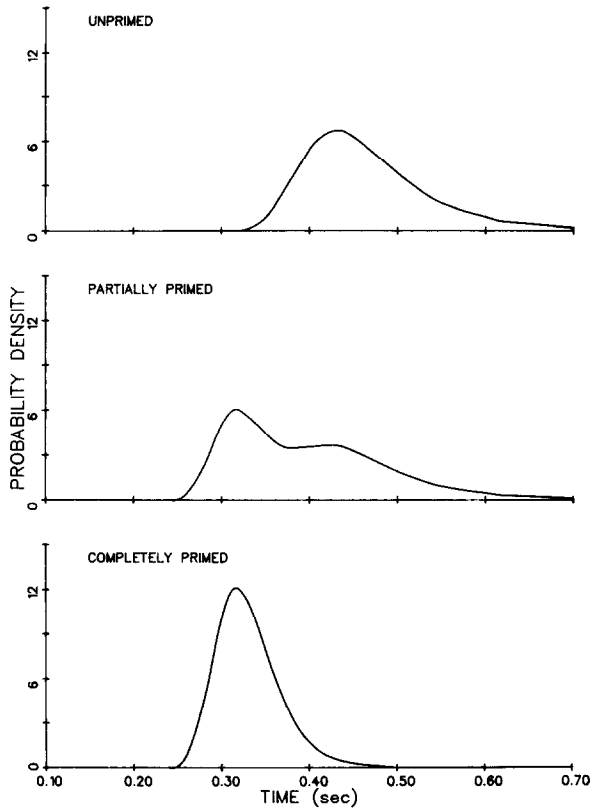


FIG. 1. Distributions of reaction times predicted by the two-state discrete stage model for the unprimed (top panel), partially primed (middle panel), and completely primed (bottom panel) conditions of the adaptive response-priming procedure. The distribution in the middle panel is a 50–50 mixture of those in the top and bottom panels.

Some predictions derived from the two-state discrete stage model are illustrated in Fig. 1. The figure shows theoretical distributions (probability-density functions) of reaction times expected in the unprimed, partially primed, and completely primed conditions of the adaptive response-priming procedure. Although the distributions for the unprimed and completely primed conditions have shapes similar to translated gamma distributions, this is not essential to the illustration. Instead, our concern is mainly with how the relative locations of these distributions and their shapes are related to the distribution for the partially primed condition.

---

senbaum, 1980). Nevertheless, the two-state discrete stage model is useful for illustrative purposes (cf. Footnote 5).

Given the assumptions discussed earlier, the two-state discrete stage model implies some systematic relationships among the reaction-time distributions for the unprimed, partially primed, and completely primed conditions. The unprimed condition places a subject in the unprepared state, yielding a distribution of relatively slow reaction times. The completely primed condition places the subject in the fully prepared state, yielding a distribution of relatively fast reaction times. It is assumed that the distributions associated with these two distinct states are unique to these states. However, on a trial of the partially primed condition, the subject may be in either one or the other of the two possible states when the test stimulus appears, depending on the speed with which the prime stimulus happens to be processed on that trial; the moment of transition between states is a random variable. This will yield a mixture distribution of slow and fast reaction times, the bases of which are the distributions from the unprepared and fully prepared states, respectively. The effect of manipulating the duration of the priming interval is to vary the amount of time available for fully processing the prime stimulus, thereby determining the exact mix of the two states in the partially primed condition.

The mixture prediction of the two-state discrete stage model can be expressed more precisely in terms of Eq. (1):

$$f_p(t|d) = \pi(d)f_c(t) + [1 - \pi(d)]f_u(t). \quad (1)$$

Here  $f_u(t)$  and  $f_c(t)$  denote the probability-density functions of the unprimed and completely primed reaction times, respectively, and  $f_p(t|d)$  denotes the probability-density function of the partially primed reaction times given a medium priming interval of duration  $d$ . The constant  $\pi(d)$  is a mixture parameter. It represents the probability that the subject finishes processing a valid prime stimulus and preparing the anticipated response before the test stimulus appears in the partially primed condition. The duration of the medium priming interval determines the value of  $\pi(d)$  by influencing the proportion of trials on which a transition occurs from the unprepared to the prepared state. As the medium interval increases, there will be more and more trials on which the subject is fully prepared, yielding an increase of the mixture parameter and a decrease of the mean reaction time.

From Eq. (1), several interesting properties of  $f_p(t|d)$  follow directly. The mean  $M_p(d)$  and variance  $V_p(d)$  of  $f_p(t|d)$  in the partially primed condition must satisfy the following equations:

$$M_p(d) = \pi(d)M_c + [1 - \pi(d)]M_u, \quad (2)$$

$$V_p(d) = \pi(d)V_c + [1 - \pi(d)]V_u + \pi(d)[1 - \pi(d)](M_u - M_c)^2, \quad (3)$$

where  $M_c$  and  $M_u$  are the means of  $f_c(t)$  and  $f_u(t)$ , respectively, and  $V_c$  and  $V_u$  are the corresponding variances. Thus,  $M_p(d)$  has to fall between  $M_c$  and  $M_u$ , because the latter two means are weighted by the fractions  $\pi(d)$  and  $1 - \pi(d)$ , respectively. In contrast,  $V_p(d)$  can be larger than both  $V_c$  and  $V_u$  because it is inflated by the factor  $\pi(d)[1 - \pi(d)](M_u - M_c)^2$ , which is relatively large when  $\pi(d)$  has a value close to one-half and  $M_c$  differs significantly from  $M_u$ . For a derivation of Eq. (3) and related expressions, see Appendix 1.

The two-state discrete stage model also makes some other interesting predictions. If  $f_u(t)$  and  $f_c(t)$  intersect at some value of  $t$ , then  $f_p(t|d)$  must intersect them at that value too, forming a confluence of the three probability-density functions. Falmagne (1968) has called this the "fixed-point property" of mixture distributions. Similarly, the tails of  $f_p(t|d)$  should overlap entirely with those of  $f_u(t)$  and  $f_c(t)$ , satisfying what Sternberg (1973) has called the "short-RT" and "long-RT" properties.<sup>9</sup> Furthermore, the partially primed distribution of reaction times may have two separate modes, even though the unprimed and completely primed distributions have only one.<sup>10</sup>

Of the preceding predictions, the basic mixture prediction [Eq. (1)] is strongest. All of the ancillary predictions stem directly from it. Our subsequent analyses therefore test the mixture prediction primarily and refer back to the others only in passing.

### *Higher-Order Discrete Models*

This application of mixture distributions can be extended to characterize other members of the discrete finite-state class, such as Miller's (1982, 1983) ADC model and higher-order stage models (i.e., ones with three or more preparatory states). For example, suppose that there are

<sup>9</sup> The short-RT property, which derives from Sternberg's (1973) analysis of self-terminating search processes, states that for all  $t$ ,  $F_p(t|d) \geq \pi(d)F_c(t)$ , where  $\pi(d)$  is the mixture parameter of the two-state discrete model. If a "short" reaction time is defined to be any time  $t$  such that  $t \leq T_s$ , and if the probability of a short time in the completely primed condition is  $\gamma_{cs}$  [i.e.,  $\gamma_{cs} = F_c(T_s)$ ], then according to the short-RT property, the probability  $\gamma_{ps}(d)$  of a short reaction time in the partially primed condition must be at least  $\pi(d)\gamma_{cs}$ . The long-RT property is a complement of this. It states that for all  $t$ ,  $F_p(t|d) \leq F_c(t)$ ; consequently, the probability of a reaction time less than  $T_s$  in the partially primed condition cannot exceed the corresponding probability in the completely primed condition [i.e.,  $\gamma_{ps}(d) \leq \gamma_{cs}$ ]. In fact, an even stronger property can be formulated for our situation. According to it,  $[1 - F_p(t|d)] \geq [1 - \pi(d)][1 - F_u(t)]$  for all  $t$ . The probability of a reaction time greater than  $T_s$  in the partially primed condition must therefore be at least  $1 - \pi(d)$  of the corresponding probability in the unprimed condition.

<sup>10</sup> For example, if  $f_u(t)$  and  $f_c(t)$  are normal distributions whose standard deviations both equal  $S$ , and if  $\pi(d)$  equals one-half, then  $f_p(t|d)$  will have two modes whenever  $|M_u - M_c| > 2S$  [Everitt & Hand, 1981, Eq. (2.4)].

$n$  distinct states of preparation ( $n \geq 3$ ), including an initial unprepared state, followed by  $n - 2$  intermediate preparatory states, and a final state of full preparation. Suppose also that  $\pi_i(d)$  ( $i = 1, 2, \dots, n$ ) denotes the probability of being in the  $i^{\text{th}}$  state after a medium priming interval, and that  $f_i(t)$  denotes the probability-density function of reaction times associated with this state. Then the probability-density function of the partially primed reaction times should satisfy the following equation:

$$f_p(t|d) = \sum_{i=1}^n \pi_i(d) f_i(t). \quad (4)$$

Here the function  $f_i(t)$  is analogous to the unprimed density function  $f_u(t)$  of Eq. (1), and the function  $f_n(t)$  is analogous to the completely primed density function  $f_c(t)$ . Moreover, as shown in Appendix 1, the mean,  $M_p(d)$ , and variance,  $V_p(d)$ , of  $f_p(t|d)$  would be

$$M_p(d) = \sum_{i=1}^n \pi_i(d) M_i, \quad (5)$$

$$V_p(d) = \sum_{i=1}^n \pi_i(d) V_i + \sum_{i=1}^n \pi_i(d) [M_i - M_p(d)]^2, \quad (6)$$

where  $M_i$  and  $V_i$  are the mean and variance, respectively, of the probability-density function associated with the  $i^{\text{th}}$  preparatory state.

Equations (4) through (6) generalize the mixture prediction of the two-state discrete stage model. Given this generalization, the mean of the partially primed reaction times [ $M_p(d)$ ] must again fall between the means of the unprimed and completely primed reaction times ( $M_1$  and  $M_n$ ). Also, the variance [ $V_p(d)$ ] of the partially primed reaction times can again exceed those of the unprimed and completely primed reaction times. When the mixture parameters  $\pi_1(d)$  and  $\pi_n(d)$  are both positive,  $f_p(t|d)$  will include nondegenerate contributions from the unprimed and completely primed density functions [ $f_1(t)$  and  $f_n(t)$ ], causing the partially primed density function to have upper and lower tails that overlap entirely with those of the other functions, as required by Sternberg's (1973) short-RT and long-RT properties (cf. Footnote 9). However, the partially primed reaction times do not have to constitute a perfect mixture of times obtained in the unprimed and completely primed conditions, because  $f_p(t|d)$  may include contributions from one or more of the basis distributions associated with the intermediate preparatory states [ $f_i(t)$ ;  $i = 2, 3, \dots, n - 1$ ].

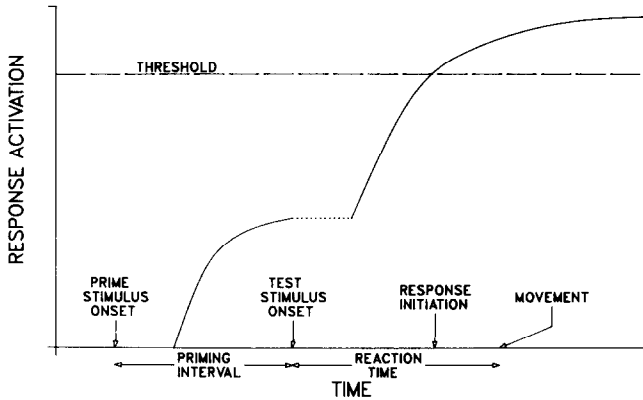


FIG. 2. A prototypical cascade model for deriving predictions about performance in the adaptive response-priming procedure.

### *Cascade Model*

Some alternative predictions by a prototypical cascade model may be derived in terms of Fig. 2. The figure illustrates the growth of activation over time for our adaptive response-priming procedure, assuming that both the prime and test stimuli activate a common response-preparation unit. Activation is depicted as starting to grow at some moment shortly after the onset of a prime stimulus and continuing throughout the priming interval (lower left solid curve) until the onset of the test stimulus. The latter event terminates the growth of activation induced by the prime, but then the growth resumes again in response to the test stimulus (upper right solid curve), progressing on from its previously attained level toward a higher asymptotic level. When a threshold (dashed line) is crossed, a response is initiated, leading eventually to an overt movement (e.g., key-press). Reaction time is measured as the duration of the interval between the onset of the test stimulus and the occurrence of the movement.<sup>11</sup> This yields a priming effect whose magnitude depends directly on how long the priming interval is. An increase of the priming interval from short to long would gradually reduce the measured reaction time because, in essence, the processing of the prime stimulus decreases the amount of time that it takes for the test stimulus to induce enough activation to cross the threshold.

<sup>11</sup> In McClelland's (1979) formulation of the cascade model, the time increment from the threshold crossing until the overt movement was treated as a constant (e.g., 100 ms). Thus, variability in reaction times, and the effects of factors on those times, would arise solely from the processes that take place before the threshold crossing.



The form of the activation-growth curves shown here follows previous statements of the cascade model (Ashby, 1982; McClelland, 1979). We assume that the activation,  $A_p(\tau)$ , associated with processing the prime stimulus has a time course that corresponds approximately to a delayed exponential approach to a limit:

$$A_p(\tau) = \begin{cases} b_p, & 0 \leq \tau < \Delta\tau_p \\ b_p + a_p\{1 - \exp[-r_p(\tau - \Delta\tau_p)]\}, & \Delta\tau_p \leq \tau \leq d \\ A_p(d), & \tau > d \end{cases} \quad (7)$$

where  $\tau$  is time measured from the onset of the prime stimulus,  $b_p$  is the initial base level of prime activation,  $\Delta\tau_p$  is the "dormant time" (delay) between the onset of the prime stimulus and the moment when activation starts to rise significantly above its base level,  $r_p$  is the rate of increase in the activation after it starts rising,  $a_p$  is the asymptotic activation level, and  $d$  is the length of the priming interval. As McClelland (1979, p. 296) demonstrated, the period of dormant time between the onset of a stimulus and the initial rise of activation would depend mainly on the rates of the relatively fast component processes (i.e., ones with steep transfer functions) in a cascaded system. If there are many such processes, then the dormant time could be substantial.<sup>12</sup> By contrast, once the activation starts to rise significantly, its rate of increase would depend mainly on the rate of the slowest process in the system, and its ultimate asymptote (upper bound) would depend jointly on the asymptotes of all the system's processes (McClelland, 1979). We assume similarly that the additional activation induced by a subsequent test stimulus involves a delayed exponential approach to a limit. The specific parameters (i.e., delay, rate, and asymptote) of the latter growth may, however, differ from those involving the prior prime stimulus, as expressed by Eq. (8):

$$A_s(t|d) = \begin{cases} A_p(d), & 0 \leq t < \Delta t_s \\ A_p(d) + a_s\{1 - \exp[-r_s(t - \Delta t_s)]\}, & t \geq \Delta t_s \end{cases} \quad (8)$$

where  $t$  is time measured from the onset of the test stimulus, and  $\Delta t_s$ ,  $r_s$ , and  $a_s$  are, respectively, the dormant time, activation rate, and asymptotic activation associated with processing the test stimulus.

Within this context, the cascade model has two distinct ways of generating reaction-time distributions for a test stimulus that follows a constant priming interval. One way is to treat the base level and/or asymptote parameters of activation as random variables. Stochastic fluctuations of

<sup>12</sup> In particular, suppose that there are  $n$  relatively fast processes whose rate parameters  $r_i$  ( $i = 1, 2, \dots, n$ ) each exceed the rate of the slowest process by a factor of 4 or more. Then, according to McClelland (1979, p. 296), the dormant time would approximately equal the sum of the reciprocals of these rate parameters (i.e.,  $\Delta\tau_p = \sum_{i=1}^n r_i^{-1}$ ).

these parameters would cause the moment at which the activation-growth curve crosses the response-threshold criterion to vary randomly over trials, yielding a distribution of reaction times. Such an approach has been pursued by Ashby (1982), who derived specific probability-density functions with the model, using straightforward analytical techniques. Another way of generating reaction-time distributions is to treat the rate parameters of the component processes in a cascaded system as random variables. Again this would cause the threshold crossings to vary randomly over trials, but unfortunately, it does not allow any easy mathematical solutions for the resultant distributions of reaction times; instead the problem becomes rather intractable, and is more readily attacked through stochastic simulation techniques (Ashby, 1982). The present article therefore focuses most heavily on the first (random base level and/or asymptote) approach, and considers the second (random rate) approach only as an ancillary possibility.

*Random base level and/or asymptote parameters.* Some representative distributions of reaction times predicted by a prototypical cascade model for our adaptive response-priming procedure appear in Fig. 3. We obtained these distributions by extending Ashby's (1982) mathematical analysis directly to the situation where response activation from a test stimulus grows on top of activation from a prior prime stimulus, and the asymptotes and/or base level of the growth curves may vary randomly over trials. The top, middle, and bottom panels of the figure illustrate individual reaction-time distributions corresponding to unprimed, partially primed, and completely primed conditions, respectively. For a formal derivation of the equations used to generate these densities, see Appendix 2.

Several facts should be noted about the predicted reaction-time distributions. As under the two-state discrete stage model, there is a substantial priming effect. The completely primed density function [ $f_c(t)$ ] includes faster times than does the unprimed density function [ $f_u(t)$ ]. However, the partially primed density function [ $f_p(t|d)$ ] is not a perfect mixture of these other functions. The variance of  $f_p(t|d)$  is intermediate to the variances of  $f_p(t)$  and  $f_u(t)$ , rather than being larger than both of them, and the tails of the partially primed function do not overlap much with those of the unprimed or completely primed functions. All three functions have similar shapes. Hence, this pattern provides a strong contrast to the predictions made by the two-state discrete stage model (cf. Fig. 1).<sup>13</sup>

<sup>13</sup> Some other members of the continuous class make predictions similar to those of the prototypical cascade model. For example, Ratcliff's (1978) random-walk diffusion model assumes that response strength moves gradually toward one or the other of two decision criteria, yielding an overt output when a criterion threshold is crossed. The information

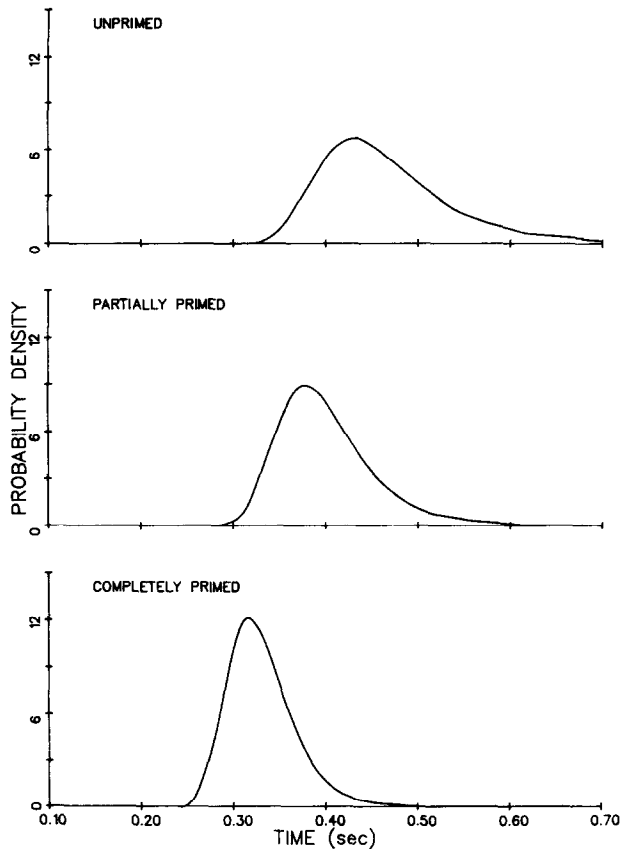


FIG. 3. Distributions of reaction times predicted by a prototypical cascade model for the unprimed (top panel), partially primed (middle panel), and completely primed (bottom panel) conditions of the adaptive response-priming procedure. See Appendix 2 for the parameters used to generate these densities. Also, note that the top and bottom panels in Fig. 1 are the same as those here.

More generally, one might conjecture that a prototypical cascade model yields reaction times whose distributions are not exactly consistent with any finite-state discrete model. Although we do not have a formal proof to support such a conjecture, it seems at least intuitively plausible. Re-

---

provided by a prime and test stimulus could be combined to determine the overall response strength. This would yield reaction-time distributions qualitatively like those in Fig. 3. Similar predictions are also made by related random-walk models (Edwards, 1965; Link, 1975; Stone, 1960), logogen models (Morton, 1969), contingent-concurrent models (Eriksen & Schultz, 1979; Turvey, 1973), and interactive-activation models (McClelland & Rumelhart, 1981).

sponse activation under the cascade model increases continuously during the priming interval. If the model is correct, then there must be a non-countable infinity of different preparatory states, each with its own unique intermediate activation level from which to proceed toward the final response-threshold criterion. This, in turn, should yield an infinity of distinct basis distributions of reaction times, which vary as a function of how long the priming interval is, whereas discrete models involve only a limited number of underlying preparatory states and basis distributions.

*Random rate parameters.* The situation appears more complex, however, when some more atypical versions of the cascade model are pursued further. For example, suppose that variability of reactions times is attributed to random fluctuations of the rate parameters of component processes in a cascaded system, not merely to random fluctuations of the base level and/or asymptote parameters (cf. Ashby, 1982). Suppose also that the activation rates of these processes are extremely high, and that there are many (e.g., dozens) of such fast processes in the system. Then the cascade model can, in principle, mimic the predictions of even a two-state discrete stage model to any desired degree of approximation. The approximation depends primarily on the number of processes in the system and on the rate of the slowest process. As more and more processes are added to the system, and as the rate of the slowest process increases, the resultant activation-growth curve will approach a step function with a relatively long dormant time (i.e., period of inactivity after stimulus onset) followed by a sharp transition from a low level of response preparation to a high level (cf. McClelland, 1979). This will yield essentially just two distinguishable preparatory states, as the discrete stage model does. Also, random fluctuations of the rate parameters will cause the moment of transition to vary from trial to trial, yielding a virtually all-or-none mechanism of preparation at the onset of a test stimulus in the partially primed condition, and producing quasi-mixture distributions of reaction times based on which preparatory state a subject happens to achieve.

### *Theoretical Caveats*

These considerations thus call for some theoretical caveats in testing discrete versus continuous models of information processing. Given enough parameters and sufficient flexibility, a continuous model may be difficult or impossible to distinguish from even a simple discrete model. The degree of separability hinges on what assumptions are made about the rates of continuous change from one level of preparation to another. A particular data set will never allow all conceivable continuous models to be rejected in favor of a discrete model, unless parsimony or some other criteria are taken into account (cf. Wickelgren, 1977).

Conversely, discrete models can, with selected ancillary assumptions, mimic the properties of continuous models to any desired level of approximation. The approximation will be more or less good, depending on the number of distinct processes placed in a system and on the "grain size" of the intermittent information packets produced by each process (Miller, 1982). As more and more processes are added, and as the granularity of the outputs is decreased, an apparently smooth transition may result among preparatory states indistinguishable from what a prototypical cascade model implies (e.g., Fig. 2). The state of affairs is analogous to ones sometimes encountered in physics and other natural sciences, where empirical phenomena do not always allow discrete processes to be distinguished easily from continuous processes, and either class of model may fit observed data quite closely.<sup>14</sup>

Nevertheless, there are still some reasonable goals for us to pursue. One of these goals involves developing informative methodological and analytical techniques that reveal how well a particular discrete or continuous model, having a specific set of parameter values, fits reaction-time distributions from subjects' performance of a given cognitive task. Another goal involves assessing how the relative goodness-of-fit achieved by alternative models varies across different tasks. This pursuit could provide investigators with ways of deciding, for their own purposes, which members of what theoretical classes are most suitable to adopt in interpreting observed magnitudes of reaction times, patterns of factor effects, and speed-accuracy trade-offs. It might also provide insights about how the temporal properties of mental processes vary systematically as a function of changing task demands.

### GOODNESS-OF-FIT TESTS

To test the goodness-of-fit for discrete versus continuous models, we have developed a new iterative maximum-likelihood statistical technique (Smith, Meyer, Yantis, & Osman, 1982).<sup>15</sup> The technique produces maximum-likelihood estimates of the mixture parameters,  $\pi_i(d)$ , and reaction-time frequency distributions associated with the bases [i.e.,  $f_i(t)$ ,  $i = 1, \dots, n$ ] of a general  $n$ -state discrete model ( $n \geq 2$ ). These estimates are

<sup>14</sup> Tests of discrete versus continuous models of information processing are also complicated in other respects. A theoretical distinction can be drawn between the activity within a stage of processing and the form of output produced by that stage (Miller, 1982). The stage itself might involve continuous growth of internal activation, but might hold its output until an asymptotic level is reached, transmitting the ultimate product discretely. Available experimental procedures, including the present one, are not powerful enough to evaluate such models separately from other discrete and continuous ones (cf. Footnote 4).

<sup>15</sup> Copies of the paper that describes this technique may be obtained from J. E. Keith Smith, Dept. of Psychology, University of Michigan, Ann Arbor, MI 48104.

constrained to satisfy the model's mixture prediction [Eq. (4)] exactly. In addition, the technique produces a  $\chi^2$  statistic that quantifies the extent to which an observed set of empirical frequency distributions of reaction times deviates from the mixture prediction.

Figure 4 illustrates how this analysis works. Here we assume, for purposes of simplicity, that there are empirical frequency distributions of reaction times from three conditions: unprimed, partially primed, and completely primed, as discussed previously. The distributions are denoted, respectively, as  $\hat{f}_u(t)$ ,  $\hat{f}_p(t|d)$ , and  $\hat{f}_c(t)$ . In this example, each distribution is divided into three frequency bins, which correspond to "fast" (e.g., less than 200 ms), "moderate" (e.g., between 200 and 400 ms), and "slow" (e.g., greater than 400 ms) reaction times. The axes of the three-dimensional space in the left panel of Fig. 4 represent the three bins, and the vectors represent the proportions of reaction times that each distribution has in each bin [e.g., .2, .3, and .5 for  $\hat{f}_u(t)$ ]. Because the proportions of fast, moderate, and slow times must sum to 1, these vectors have endpoints on an oblique two-dimensional plane in the three-dimensional space.

When a two-state discrete model is to be tested, we seek three collinear points that fall on the two-dimensional plane and that come closest in an overall maximum-likelihood sense to the vector endpoints of the empirical frequency distributions of reaction times. The results of such a search are illustrated by the dashed line and solid points in the right panel of Fig. 4. The individual collinear (solid) points closest to  $\hat{f}_u(t)$ ,  $\hat{f}_p(t|d)$ , and  $\hat{f}_c(t)$ ,

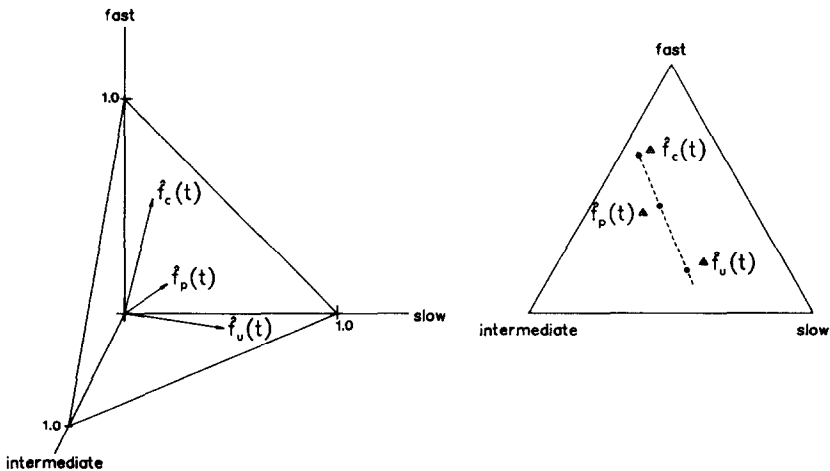


FIG. 4. An illustration in which the iterative maximum-likelihood statistical technique (Smith et al., 1982) is applied to test the goodness-of-fit of the two-state discrete stage model.

respectively, constitute the estimates of the unprimed, partially primed, and completely primed distributions of reaction times according to the model's mixture prediction. Furthermore, in this context, the estimated mixture parameter  $\pi(d)$  equals the ratio of two quantities: (i) the distance between the collinear points found for the unprimed and partially primed distributions, and (ii) the distance between the collinear points found for the unprimed and completely primed distributions. The two-state model, which involves a weighted linear combination of probability-density functions [Eq. (1)], implies that these collinear points should all coincide perfectly with the vector endpoints of the empirical distributions of reaction times, if the data were not noisy. When the data are noisy, but the model is valid, then the fit should still be relatively good, and our statistical technique would produce a test statistic that has an approximate  $\chi^2$  distribution with 1 degree of freedom in the present case. Values of this statistic beyond the usual Type I error cutoffs of a  $\chi^2$  distribution would constitute evidence against the model.

Tests of general  $n$ -state discrete models ( $n \geq 2$ ) can be obtained by extending the illustration in Fig. 4. The extension requires empirical reaction-time distributions from at least  $n + 1$  different priming conditions, including one unprimed condition,  $n - 1$  partially primed conditions, and one completely primed condition. Each distribution must be divided into at least  $n + 1$  separate bins, ranging from relatively fast to relatively slow reaction times. Our maximum-likelihood statistical technique yields a set of estimated basis frequency distributions that satisfy Eq. (4), a corresponding set of mixture parameters, and a test statistic whose value indicates the  $n$ -state model's goodness-of-fit. If the model holds, and if an experiment includes a total of  $P$  priming conditions ( $P \geq n + 1$ ) together with  $B$  bins per empirical distribution ( $B \geq n + 1$ ), then this statistic will have an approximate  $\chi^2$  distribution with  $(P - n)(B - n)$  degrees of freedom.

It is possible to specify the power of our tests of  $n$ -state discrete models ( $n \geq 2$ ) versus continuous models when the number of priming conditions ( $P$ ) and bins ( $B$ ) per distribution both exceed  $n$ . For example, when we tested the mixture prediction of a two-state discrete stage model [Eq. (1)] against the distributions of reaction times expected from a prototypical cascade model (Fig. 3), a highly significant  $\chi^2$  statistic emerged [ $\chi^2(8) = 31.0, p < .001$ ].<sup>16</sup> This result clearly violates the two-state model, helping to gauge how sensitive the experiments are.

<sup>16</sup> In this test of power, 10 bins were created to span the reaction-time distributions predicted by the prototypical cascade model for the unprimed, partially primed, and completely primed conditions. The number of bins was similar to those used in analyzing our actual data, which, given the present experimental sample sizes, would closely approximate

In general, the power of the goodness-of-fit tests depends on the parameters of the models at hand (cf. Theoretical Caveats), the magnitude of the overall priming effects, and the variances of the unprimed and completely primed distributions. This can be seen intuitively by looking at Eq. (3) for the variance of the partially primed reaction times [ $V_p(d)$ ]. As the overall priming effect ( $M_u - M_c$ ) increases and/or the variances of the unprimed and completely primed distributions ( $V_u$  and  $V_c$ ) decrease, the relative impact of the inflation factor  $\pi(d)[1 - \pi(d)](M_u - M_c)^2$  on performance increases, making it easier to discriminate the two-state discrete stage model from members of the continuous class. Likewise, the discriminability is increased by selecting a medium priming interval for which the distribution of partially primed reaction times has an intermediate mean. With an intermediate mean,  $M_p(d)$ , the term  $\pi(d)[1 - \pi(d)]$  of the inflation factor reaches a maximum, further increasing its relative impact under the two-state model. We will have more to say later about the power of our tests against various alternatives (see General Discussion).

### SELECTION OF PRIMING INTERVALS

The selection of durations for the short and long intervals between the prime and test stimuli is relatively easy. We typically use a short interval of 0 ms which corresponds to no visible presentation of a valid prime stimulus and allows no useful response preparation. The duration of the long interval can be based on results from tasks in which speeded responses are made directly to stimuli like the primes. In past lexical decision experiments, for example, our subjects have usually taken about 500 ms to discriminate words from nonwords (Meyer et al., 1975). Thus, with words and nonwords as prime stimuli, we use long priming intervals of about 750 ms for the present experiments, which should be enough for subjects to finish preparatory activities completely before a test stimulus occurs.

Selecting a medium priming interval is more problematic. The partially primed reaction-time distribution should fall about midway between those obtained with the short and long intervals, because this maximizes the power of our statistical tests. However, the medium interval that best yields partial priming could depend on many factors, including the type of prime and test stimuli, the particular subjects being tested, the amount of practice, the alertness of the subjects, and so forth. We have therefore

---

the distributions in Fig. 3, if the model were correct. Of course, the exact result of the test only applies to the particular parameter values that we chose in the illustration. More general information about the power of our tests with respect to a larger parameter space appears in the General Discussion.



relied on a staircase tracking algorithm to adjust this interval. The algorithm is designed to converge on an ideal medium interval in a relatively small number of trials, much like staircase tracking in standard psychophysical paradigms (Levitt, 1971).

#### *Staircase Tracking Algorithm*

In implementing our staircase tracking algorithm, samples of reaction times are first obtained from trials on which the valid prime stimuli precede the test stimuli by either short or long priming intervals. The initial samples let us estimate theoretical distributions of unprimed and completely primed reaction times. We denote the cumulative distribution functions of these times as  $F_u(T)$  and  $F_c(T)$ , respectively [i.e.,  $F_u(T) = P(t \leq T | \text{no priming})$ , and  $F_c(T) = P(t \leq T | \text{complete priming})$ , where  $t$  is reaction time, and  $T$  is any specified temporal value]. As before, the probability-density functions for these distributions are  $f_u(t)$  and  $f_c(t)$ , which correspond to the derivatives of the cumulative distribution functions.

Next we estimate a temporal cutpoint,  $T_x$ , midway between  $f_c(t)$  and  $f_u(t)$  in the sense that the estimated proportion of  $f_u(t)$  above  $T_x$  equals the estimated proportion of  $f_c(t)$  below  $T_x$ . The cutpoint is intended to satisfy the following equation:

$$F_c(T_x) = 1 - F_u(T_x). \quad (9)$$

When  $F_c(T)$  and  $F_u(T)$  are continuous strictly increasing functions whose probability densities overlap each other, there will be a unique  $T_x$ . The staircase tracking algorithm then seeks an ideal medium priming interval,  $d_m^*$ , for which the distribution of partially primed reaction times [ $f_p(t|d)$ ] has a median equal to  $T_x$ . For more details about how this works, see Appendix 3.

#### *Psychometric Priming Function*

An instructive way of representing the objective of the staircase tracking algorithm appears in Fig. 5. This figure illustrates what we call a "psychometric priming function," denoted  $\Psi(d)$ , which is analogous to psychometric functions from sensory psychophysics. The psychometric priming function has qualitative properties like a plot of mean reaction times and/or mean priming effects versus the duration of a prime stimulus (e.g., Meyer et al., 1972; Neely, 1976; Posner & Snyder, 1975; Sabol & DeRosa, 1976; Yellott & Hildreth, 1969), but its quantitative properties are somewhat different and more useful for our purposes.

Given each possible duration,  $d$ , of the interval between the prime and test stimuli,  $\Psi(d)$  specifies the probability that a reaction time obtained with that interval will fall below the temporal cutpoint,  $T_x$ . The lower bound of  $\Psi(d)$  equals the area under the unprimed reaction-time distri-

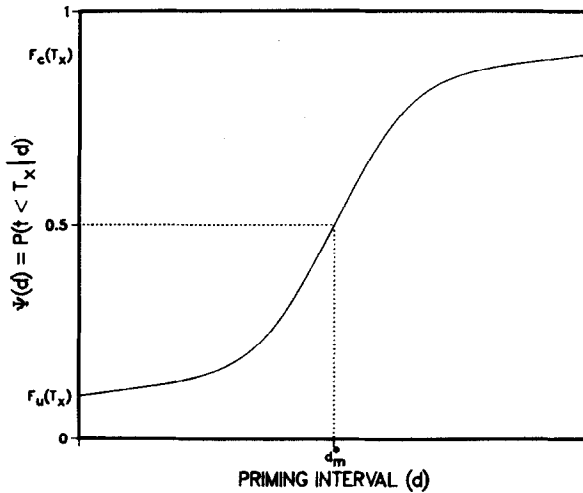


FIG. 5. A psychometric priming function that relates the duration,  $d$ , of a priming interval to the probability,  $\Psi(d)$ , that a reaction time obtained with this interval will fall below the temporal cutpoint  $T_x$ . The parameter  $d_m^*$  is the ideal medium priming interval for which the psychometric priming function equals 0.5 and the median of the partially-primed density function,  $f_p(t|d_m^*)$ , equals  $T_x$ .

bution to the left of the cutpoint [i.e.,  $F_u(T_x)$ ]. This bound will be low to the extent that the unprimed condition produces few relatively fast reaction times. The upper bound of  $\Psi(d)$  equals the area under the completely primed distribution of reaction times to the left of the cutpoint [i.e.,  $F_c(T_x)$ ]. This bound will be high to the extent that the completely primed condition produces many relatively fast reaction times. The tracking algorithm converges on an ideal medium priming interval  $d_m^*$  that corresponds to the 50% point on the psychometric priming function. As we indicate more fully later (see General Discussion),  $\Psi(d)$  may also bear on other important theoretical considerations concerning the two-state discrete stage model.

## OVERVIEW OF EXPERIMENTS

Three experiments are reported here to demonstrate our approach in testing discrete versus continuous models. For each experiment, the test stimuli were arrows, and the responses were keypresses by designated fingers. We used words and nonwords as the valid prime stimuli. Neutral primes (uninformative rows of X's) were also included. Whenever the prime stimulus was a word, it accurately cued the subject that the subsequent test stimulus would require a response with a particular hand and/or finger, and whenever the prime stimulus was a nonword, it accurately cued the subject that a response with another hand and/or finger would be required instead. The subjects practiced extensively on the

experimental task. The number of alternative test stimuli and responses, and the compatibility of the mapping between them, changed from experiment to experiment. Furthermore, the number of priming conditions and our instructions regarding speed-accuracy trade-offs varied.

The design of the experiments was intended to satisfy three general requirements. First, we wanted to obtain reasonably powerful tests from our maximum-likelihood statistical technique. Using valid, easily processed prime stimuli helped us in this respect, because it yielded rather large priming effects.<sup>17</sup> Giving subjects extended practice also helped, because it reduced the variance of the obtained reaction-time distributions, making them more easily discriminable from each other. Second, we wanted to obtain some data in support of both discrete and continuous models, thus starting to build a taxonomy for determining which models fit best under what circumstances. The present manipulations of stimulus-response compatibility and speed-accuracy trade-off strategies served this purpose, in that they seemed likely to have marked qualitative effects on the temporal properties of information processing (Fitts & Seeger, 1953; Ollman, 1977; Swensson, 1972). Third, we wanted to build on past research regarding discrete versus continuous models. In particular, much of McClelland's (1979) argument for the cascade model and against Sternberg's (1969) additive-factor method dealt with data and processes related to word recognition. This provided additional motivation for our selection of the words and nonwords as prime stimuli. The chosen stimulus-response mappings were likewise appropriate in that they followed studies by Leonard (1958), Miller (1982), and Rosenbaum (1980), who gave subjects preliminary information about which hand and/or finger to use in making manual responses.

We should emphasize that the present experiments did differ in significant respects from some previous studies of priming and word recognition. Because words and nonwords do not have strong a priori associations with arrows or keypress responses, the effects of prime stimuli on reaction times observed here may involve "controlled" rather than "automatic" priming (cf. Neely, 1976, 1977; Posner & Snyder, 1975). This

<sup>17</sup> Unlike some previous research (e.g., Neely, 1976; Posner & Snyder, 1975), our experiments did not include any invalid primes, because they might discourage subjects from paying attention to the valid primes (but see Yantis, 1985). Also, we used test stimuli and responses difficult enough that subjects would benefit significantly from the valid primes. There were no extremely compatible stimulus-response combinations, which might place reaction times on a floor level even without any priming (Fitts & Seeger, 1953; Leonard, 1959). However, higher compatibility existed between the test stimuli (arrows) and responses (keypresses) than between the prime stimuli (words and nonwords) and responses. This compatibility difference encouraged subjects to switch from processing the prime stimuli to processing the test stimuli as soon as they appeared, consonant with our ancillary assumptions.

could have important implications for theoretical interpretations of our results and comparisons of them with those of other investigators, who have examined automatic priming processes (see Experiment 1, Discussion).

## EXPERIMENT 1

Experiment 1 included two test stimuli and two responses. The test stimuli were right and left arrows, and the responses were keypresses by the right and left index fingers. Word and nonword primes provided valid information about which key would have to be pressed (word = right, nonword = left). There were three priming conditions: unprimed, partially primed, and completely primed, which corresponded to short, medium, and long priming intervals, respectively. (For a brief summary of this study, see Meyer et al., 1984.)

### *Method*

*Subjects.* Four undergraduate students at the University of Michigan participated as paid subjects. They were sampled from a pool of volunteers maintained by the Human Performance Center. Three (B.C., S.R., and N.Y.) were females and the other (J.F.) male. Each subject received approximately \$4 per session, including a salary of \$1.75 plus a bonus for good performance.

*Apparatus.* The subjects sat at a table in a moderately illuminated sound-attenuating booth. A minicomputer (Digital Equipment Corporation PDP-11/34) presented the stimuli and recorded the responses. Warning signals, prime stimuli, test stimuli, and feedback appeared on a display terminal (Hewlett-Packard 2621A). Responses were made on the terminal's keyboard, which was placed so that the subjects' arms rested comfortably on the table.

*Stimuli.* Sets of 120 English words and 120 pronounceable nonwords, each having four letters, were used as prime stimuli. The frequency of the words in normal text equaled or exceeded 32 occurrences per million (Kucera & Francis, 1967). The nonwords were constructed by altering individual letters of the words without violating English orthographic rules (Venezky, 1970). All of the primes appeared in uppercase characters. Each character subtended about 0.35° of visual angle in width and 0.5° in height at a viewing distance of 35 cm. The test stimuli were left and right arrows, which subtended a horizontal visual angle of 1.05°.

*Design.* For each subject, the experiment was divided into four 1-h sessions conducted on separate days within a 1-week span. The first and second sessions provided instructions and practice on various aspects of the adaptive response-priming procedure. The third and fourth sessions were used to collect the reported data.

Session 1 began with three different blocks of 30 instruction trials. During the first block, subjects had to make lexical decisions about whether strings of letters were words or nonwords. The events on these trials were similar to those in subsequent blocks, except that no arrows were presented as stimuli, and responses occurred directly to the letter strings. During the second block, the subjects performed a pure arrow-discrimination task for which keypress responses were made to right and left arrows without the benefit of any informative prime stimuli. The third block introduced the adaptive response-priming procedure with a combination of valid primes (words and nonwords) and test stimuli (arrows). Following these instruction blocks, there were seven additional practice blocks of 72 trials

with the priming procedure. Session 2 replaced the initial three instruction blocks with a warmup block of 20 trials involving the priming procedure, but was otherwise similar to Session 1.

Sessions 3 and 4 began with a warmup block of 20 trials followed by seven test blocks of 72 trials. Seventy-five percent of the trials per block were "regular" ones on which a test stimulus actually appeared. The other 25% were "catch" trials on which there was no test stimulus (right or left arrow). We included the catch trials to monitor and discourage premature responses (anticipations). The unprimed, partially primed, and completely primed conditions each occurred with one-third probabilities during the regular trials and during the catch trials. On the regular trials, right arrows always followed the word primes, and left arrows always followed the nonword primes. Each type of test stimulus had a probability of one-half, as did each type of prime. Given these constraints, the order of the trials was randomized. All of the words and nonwords were used at least once over the course of a session. This yielded a total of 504 trials (i.e., 378 regular and 126 catch) per session to be analyzed for each subject.<sup>18</sup>

*Procedure.* On each regular trial, four pound signs (####) first appeared in the center of the display screen as an initial warning signal. This signal had a duration of 500 ms, which was chosen to maximize the subjects' level of alertness, based on previous studies of general foreperiod effects (Bertelson, 1967; Posner & Boies, 1971). When the trial involved the partially primed or completely primed condition, the pound signs were replaced by a valid prime stimulus (word or nonword) for a variable duration (priming interval). When the trial involved the unprimed condition, no prime was displayed, in effect yielding a priming interval of zero. Next the display was erased, and two dashes appeared for 85 ms in the positions reserved previously for the second and third characters of the prime stimulus. The dashes served as a final warning signal that helped to equalize subjects' alertness regardless of the duration of the priming interval.<sup>19</sup> After the final warning signal, the two center dashes changed into an arrow (i.e., test stimulus) that abruptly moved one character position to the right or left. Specifically, a left arrow appeared as a less-than sign to the left of two dashes in character positions 1 and 2 (<- -), and a right arrow appeared as a greater-than sign to the right of the two dashes in character positions 3 and 4 (->). The arrow remained visible until either a response was recorded or 1500 ms elapsed. Re-

---

<sup>18</sup> A number of other ancillary trials were also included in each of these blocks. At the start of each block, three warm-up trials occurred to get the subject in a state of equilibrium before the ensuing test trials. Moreover, an extra recovery trial was inserted after each incorrect response, allowing subjects to regain their composure before proceeding.

<sup>19</sup> The effectiveness of the final warning signal is demonstrated by results of a control study run with four other subjects (Meyer et al., 1984). In the control study, the procedure was the same as for Experiment 1, except that all of the prime stimuli were uninformative (i.e., "neutral") rows of four X's rather than informative (valid) words and nonwords (cf. Neely, 1976; Posner & Snyder, 1975). The priming interval was short, medium, or long. A final warning signal (two centered dashes) immediately preceded the onset of each test stimulus for 85 ms. Despite the variable duration of the neutral prime, the mean reaction times were relatively constant. A difference of only 16 ms emerged between the means obtained with the short and long priming intervals. While the difference was statistically significant [ $t(3) = 4.6, p < .05$ ], it was much less than the 77 ms effect obtained with the valid primes in Experiment 1. This demonstrates that the present priming effects cannot be attributed only to general foreperiod effects and fluctuations of subjective alertness (cf. Bertelson, 1967; Posner & Boies, 1971). Other evidence against the presence of general foreperiod effects appears in Experiments 2 and 3.

sponses were made using the Z and slash (/) keys of the terminal keyboard. The subject pressed either the slash key with the right index finger for a right arrow, or the Z key with the left index finger for a left arrow. Subjects were instructed to use the prime stimuli for preparing fast responses to the test stimuli. However, they were also instructed that when an arrow appeared, they had to stop processing the prime and proceed with responding to the arrow. It was emphasized that producing a fast accurate response had highest priority. The catch trials were similar to the regular trials, except that the final warning signal persisted for an additional 750 ms, and did not change into an arrow. Subjects had to withhold their responses on catch trials.

After each trial, there was a feedback period followed by a 1-s intertrial interval. If the subject made an error on a regular trial, the message INCORRECT appeared for 500 ms. If an anticipation occurred on a catch trial, the message PLEASE RESPOND ONLY WHEN ARROWS APPEAR was displayed for 500 ms. If any key other than the Z or slash key was depressed, the message PLEASE USE Z OR / KEYS ONLY appeared.<sup>20</sup>

After each trial block, a chart was displayed to summarize the total correct responses, the number of errors, the mean reaction time, and the bonus point score for the block. Each correct response earned 100 points, each incorrect response lost 300 points, and each 10 ms of reaction time on a trial lost 1 point in a subject's overall score. For example, a correct response that took 500 ms earned 50 points. Subjects received a payment proportional to their net total of bonus points. The bonus points were converted to cash at a rate of 1¢ per 100 points. A typical bonus score for a session equaled about \$2 to \$2.50. There was a 1-min rest period between trial blocks, and an additional break of 3 min halfway through each session.

*Priming intervals.* As mentioned previously, the priming interval in the unprimed condition was set to zero by proceeding directly from the initial warning signal to the final warning signal, without displaying a valid prime stimulus (word or nonword). This precluded any specific priming effect, giving us an unprimed distribution of reaction times.

The completely primed condition involved a long priming interval of 805 ms. The long interval consisted of 720 ms for the presentation of the prime stimulus, plus another 85 ms for the final warning signal. We assume that subjects processed the prime stimulus fully during this interval, and that they were highly prepared when the test stimulus appeared, producing a completely primed distribution of reaction times. This is consistent with previous lexical-decision research, where responses to words and nonwords have taken about 500 ms (Meyer et al., 1975).

The partially primed condition involved a medium priming interval that was set by the staircase tracking algorithm (Appendix 3). Starting with the third instruction block of Session 1, initial estimates of the ideal medium interval  $d_m^*$  and temporal cutpoint  $T_x$  between the theoretical unprimed and completely primed distributions of reaction times were initialized at 200 ms and 350 ms, respectively. The initial cutpoint estimate was used as the tracking algorithm began adjusting the duration of the medium interval. On each subsequent block of Session 1, a new estimate of the cutpoint replaced the old one. The new estimate equaled the overall mean reaction time from the immediately preceding block.

At the start of each block in Sessions 2 through 4, the cutpoint was reestimated by taking the unprimed and completely primed reaction times from the immediately preceding five blocks and finding a value of  $T_x$  such that Eq. (9) held. The cutpoint estimate was updated

---

<sup>20</sup> This feedback was supplemented by additional information during the warm-up and practice trials. If the subject was correct on them, then the message CORRECT was displayed for 500 ms. Also, bonus points were displayed after each of these trials, allowing subjects to learn the benefit of speed and cost of errors.

regularly, accommodating any changes in subjects' performance due to practice effects. We replaced old samples of  $f_u(t)$  and  $f_c(t)$  with new ones after each trial block, keeping a backlog of data from the last 60 trials for estimation purposes. For example, the cutpoint on Block 2 of Session 3 was calculated from data on Block 1 of Session 3 combined with data from previous blocks of Session 2.

Concurrently, the duration of the ideal medium priming interval,  $d_m^*$ , was estimated on a trial-by-trial basis as outlined in Appendix 3. The tracking algorithm was applied separately to data obtained with word primes and with nonword primes, using reaction-time cutpoints ( $T_x$ ) derived separately for each of the two prime types. This allowed possible differences in performance for the different primes to be accommodated.

Subjects' reaction times typically stabilized by the end of the second session. Across subjects and trial blocks, the ideal medium priming intervals averaged 185 ms for word primes and 203 ms for nonword primes. The difference between these averages is consistent with results from previous research on lexical decisions, which have revealed slower processing times for nonwords than words (Meyer & Schvaneveldt, 1971).

*Data analysis.* Our analysis focused on data from Sessions 3 and 4. For the correct responses in those sessions, the means, standard deviations, and distributions of reaction times were estimated. We employed the maximum-likelihood statistical technique (Smith et al., 1982) discussed earlier to test the mixture predictions of the two-state discrete stage model and a three-state model [Eqs. (1) and (4)]. We also examined the rates of errors on regular trials and anticipations on catch trials. An error consisted of pressing the wrong key after the onset of the test stimulus. An anticipation consisted of pressing a key after the onset of the final warning signal but before the onset of a test stimulus.

The data were analyzed for each combination of prime stimulus and priming condition. We treated individual subjects and sessions separately. The results below have been taken from each subject's session for which the partially primed condition yielded mean reaction times closest to midway between the means obtained in the unprimed and completely primed conditions. This maximized the power of our statistical tests, which depend on the unprimed and completely primed distributions of reaction times being as far apart as possible. Data collected during Sessions 1, 2, and the ancillary (i.e., warm-up and recovery) trials of Sessions 3 and 4 were not analyzed. The analysis also excluded latencies of incorrect responses.

## Results

Table 1 shows the means and standard deviations of reaction times for correct responses on regular trials, the error rates on regular trials, and the anticipation rates on catch trials from each subject, prime type, and priming condition. Each row of the table includes contributions taken over 63 regular trials and 21 catch trials. Several general facts about the data should be noticed. The prime stimuli had significant facilitative effects on reaction times for all subjects. When there was complete priming, the effects averaged 77 ms [ $t(120) \geq 8.0$ ,  $p < .001$  in all cases]. When there was partial priming, the effects averaged 38 ms, half of the mean for the completely primed condition, but still highly significant [ $t(122) \geq 3.3$ ,  $p < .01$  in all cases]. Error rates on regular trials were low, never exceeding 4% and not varying significantly across subjects, prime stimuli, or priming conditions ( $p > .10$ ). There was no evidence of a speed-accuracy trade-off on regular trials. However, there were substantial

TABLE 1  
Results for Individual Subjects in Experiment 1

Subject	Prime type	Priming condition	Mean RT (ms)	$\sigma$ (ms)	Errors (%)	Anticipations (%)
B.C.	Word	Unprimed	304	23	0.0	0.0
		Partial	255	56	0.0	19.0
		Complete	240	26	0.0	0.0
	Nonword	Unprimed	297	23	0.0	0.0
		Partial	246	47	1.6	4.8
		Complete	234	35	0.0	14.3
S.R.	Word	Unprimed	293	32	0.0	0.0
		Partial	262	43	3.2	9.5
		Complete	234	47	1.6	9.5
	Nonword	Unprimed	295	34	0.0	0.0
		Partial	267	56	0.0	4.8
		Complete	231	36	1.6	4.8
J.F.	Word	Unprimed	306	26	0.0	4.8
		Partial	271	29	1.6	0.0
		Complete	230	59	1.6	14.3
	Nonword	Unprimed	298	23	0.0	0.0
		Partial	267	30	1.6	0.0
		Complete	196	64	1.6	23.8
N.Y.	Word	Unprimed	311	20	0.0	0.0
		Partial	267	36	0.0	4.8
		Complete	212	46	0.0	33.3
	Nonword	Unprimed	313	20	0.0	0.0
		Partial	278	37	1.6	9.5
		Complete	223	42	0.0	33.3

numbers of anticipatory responses on catch trials. The overall anticipation rate increased significantly ( $p < .05$ ) with the degree of priming, averaging 0.6, 6.6, and 16.6% across subjects for the unprimed, partially primed, and completely primed conditions. Anticipations always occurred at least 100 ms after the onset of the final warning signal, during the period when there would have been a test stimulus if the trial were a regular one. Thus, anticipations could not be detected as such on regular trials.

Some intriguing differences also occurred between the results of the individual subjects. Two of them (J.F. and N.Y.) made significantly more anticipatory responses than did the other two (B.C. and S.R.), especially in the completely primed condition (26.2% vs 7.2%;  $z = 4.25$ ,  $p < .001$ ). Likewise, the high-anticipation subjects had larger overall priming effects on reaction times than did the low-anticipation subjects [92 ms vs 63 ms;  $t(6) = 4.2$ ,  $p < .01$ ]. This pattern of individual differences has potentially interesting theoretical implications, as we discuss later.



*Subject B.C.* Figure 6 shows the distributions of reaction times produced by subject B.C. when the prime stimuli were words. The top, middle, and bottom panels contain results from the unprimed, partially primed, and completely primed conditions, respectively. Solid histograms represent observed relative frequencies of reaction times, and dashed curves represent maximum-likelihood estimates of the best-fitting frequency distributions derived from the mixture prediction [Eq. (1)] of the two-state discrete stage model (cf. Goodness-of-Fit Tests).

B.C.'s reaction times for the word primes look very similar to the mixture distributions illustrated in Fig. 1, supporting the two-state model.

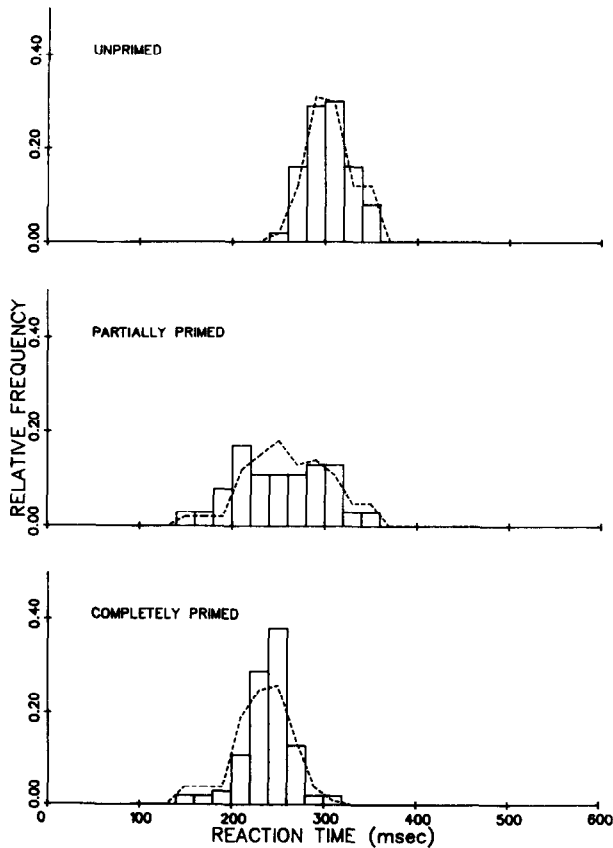


FIG. 6. Distributions of reaction times produced by subject B.C. in Experiment 1 when the prime stimuli were words. The top, middle, and bottom panels show the data obtained in the unprimed, partially primed, and completely primed conditions, respectively. Solid histograms represent relative frequencies of observed times, and dashed curves represent best-fitting frequency distributions that satisfy the mixture prediction of the two-state discrete stage model.

In the partially primed condition, the empirical distribution was flatter and wider than those in the unprimed and completely primed conditions. The lower tail of the partially primed distribution overlapped entirely with the lower tail of the completely primed distribution, and the upper tail of the partially primed distribution overlapped entirely with the upper tail of the unprimed distribution. Combined deviations of the fitted frequency distributions (dashed curves) from the actual data (solid histograms) were not statistically significant [ $\chi^2(13) = 21.5, p > .05$ ].<sup>21</sup> The estimated value of the mixture parameter  $\pi(d)$  was .64. This pattern is not consistent with the one expected from a prototypical cascade model (Fig. 3), raising doubts about whether the processing here involved a continuous growth of response activation over time.

The distributions of reaction times produced by B.C. when the prime stimuli were nonwords had a pattern similar to that for word primes (Fig. 7). The deviation of these distributions from the mixture prediction was insignificant [ $\chi^2(13) = 16.5, p > .20$ ], and the mixture parameter  $\pi(d)$  equaled .66.

Support for the mixture prediction is also evident from the reaction-time variances (Table 1). B.C. produced greater variances in the partially primed condition than in the unprimed and completely primed conditions. This is consistent with Eq. (3), which includes the inflation factor  $\pi(d)[1 - \pi(d)](M_u - M_c)^2$  for the variance,  $V_p(d)$ , expected under the medium priming interval.

*Subject S.R.* The results of S.R. followed the same pattern as those of B.C. Her reaction times had approximately equal or greater variances in the partially primed condition than in the unprimed and completely primed conditions (Table 1). The lower and upper tails of the partially primed distribution of reaction times overlapped entirely with those of the completely primed and unprimed distributions, respectively. Furthermore, the mixture prediction of the two-state discrete stage model was well fit for both word primes [ $\chi^2(6) = 4.6, p > .25$ ] and nonword primes [ $\chi^2(6) = 6.7, p > .25$ ]. In these goodness-of-fit tests, the mixture parameter  $\pi(d)$  had estimated values of .57 and .61, respectively.

*Subject J.F.* However, quite different results emerged in the case of subject J.F. There were three distinctive characteristics of the observed data. First, the variance of the reaction times increased monotonically as the degree of priming increased; partial priming did not yield the largest

<sup>21</sup> The goodness-of-fit is especially compelling because hidden artifacts may spuriously inflate the calculated  $\chi^2$  values, which are sensitive to certain types of nonindependence that violate multinomial probability assumptions. For example, if the mixture parameter  $\pi(d)$  varies across experimental trials, or if the density functions of the basis distributions [i.e.,  $f_u(t)$  and  $f_c(t)$ ] vary from trial to trial, then this could decrease the apparent goodness-of-fit of the two-state discrete stage model, even though the model is actually valid.

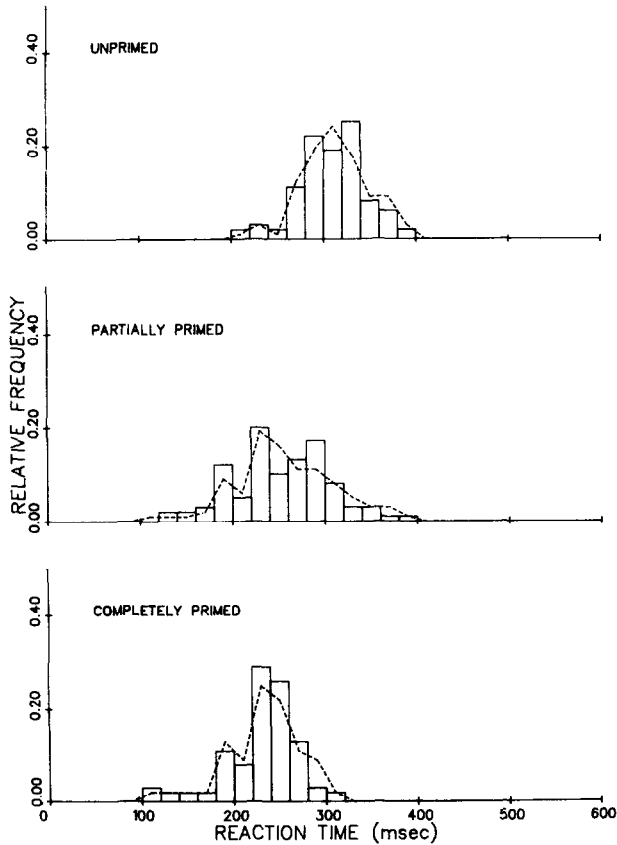


FIG. 7. Distributions of reaction times produced by subject B.C. when the prime stimuli were nonwords.

variance. Second, the completely primed distribution had two separate peaks. The left peak contained a substantial number of reaction times faster than 200 ms, contrasting with the data of subjects B.C. and S.R., who produced no such times, whereas the right peak contained times more like the ones of those subjects. Third, the partially primed distribution was not well fit by a mixture of the unprimed and completely primed distributions [ $\chi^2(13) = 36.7, p < .01$ ], contrary to the two-state model.

Furthermore, J.F.'s data for the nonword primes were similar to his data for the word primes. A monotonically increasing variance of reaction times again occurred as the degree of priming increased (Table 1), and the completely primed distribution had two separate peaks. Most of the reaction times under the left peak were less than 200 ms, whereas the right peak contained times greater than 200 ms. The partially primed

distribution had only a single peak, and it violated the mixture prediction of the two-state discrete stage model [ $\chi^2(13) = 43.8, p < .01$ ].

This pattern of results is inconsistent with a prototypical cascade model as well as the two-state discrete stage model. Although the cascade model does not make a finite-mixture prediction, it lacks any easy way of explaining the observed trend in the variances of the reaction times and the bimodality of the completely primed distribution.

*Subject N.Y.* The results from subject N.Y. looked much like those of J.F. Her reaction-time distributions in the partially primed condition did not have the largest variances (Table 1). Instead, the largest variances occurred with complete priming, and the completely primed distributions were bimodal. The left peaks of those distributions contained many times less than 200 ms, whereas the right peaks contained times comparable to those of subjects B.C. and S.R. in the completely primed condition. Significant deviations again emerged when we tested the mixture prediction of the two-state model against these data [ $\chi^2(11) = 38.2, p < .001$ , for word primes;  $\chi^2(11) = 36.8, p < .001$  for nonword primes].

*Test of the three-state model.* Although the results from J.F. and N.Y. are not consistent with the two-state discrete stage model, a three-state model may fit them reasonably well. The possible existence of three distinct preparatory states is suggested, in particular, by the bimodality of these subjects' completely primed distributions of reaction times. Left peaks of the latter distributions, which include many very fast (i.e., less than 200 ms) times, could correspond to an extremely high state of preparation. Right peaks could correspond to a more moderate state that lies between the extremely high state and the unprepared state that is reflected by the reaction-time distributions in the unprimed condition. If J.F. and N.Y. experienced all three states, depending on the length of the interval between the prime and test stimuli, but B.C. and S.R. experienced only the unprepared and moderately prepared states, then this would perhaps account for the differences observed between these two pairs of subjects.

One way of testing for three preparatory states would involve examining reaction-time distributions obtained under four or more different priming conditions (e.g., based on a short, a long, and two medium priming intervals). As outlined earlier, our maximum-likelihood statistical technique allows the mixture prediction of an  $n$ -state discrete model to be assessed whenever  $n + 1$  empirical distributions of reaction times are available ( $n \geq 2$ ). However, Experiment 1 included only three priming conditions (i.e., unprimed, partially primed, and completely primed), which makes it impossible to evaluate a three-state model for subjects J.F. and N.Y. using this approach.

We have therefore taken a somewhat different approach toward ana-

lyzing their data further. As part of the subsequent analyses, the completely primed distributions of reaction times were divided into two parts, based on an assumed correspondence between each part and a distinct underlying preparatory state. This division, together with the distributions of reaction times from the unprimed and partially primed conditions, gave us additional leverage for testing a three-state discrete model despite our original use of only three priming conditions.

Figure 8 shows how the analysis worked in the case of J.F. when the prime stimuli were words. Here we divided the completely primed distribution (solid histogram, lower panel) into one part that contained reaction times faster than 200 ms (left peak) and another part that contained times greater than 200 ms (right peak). A value of 200 ms was used as a dividing point because of two related considerations. First, the minimum

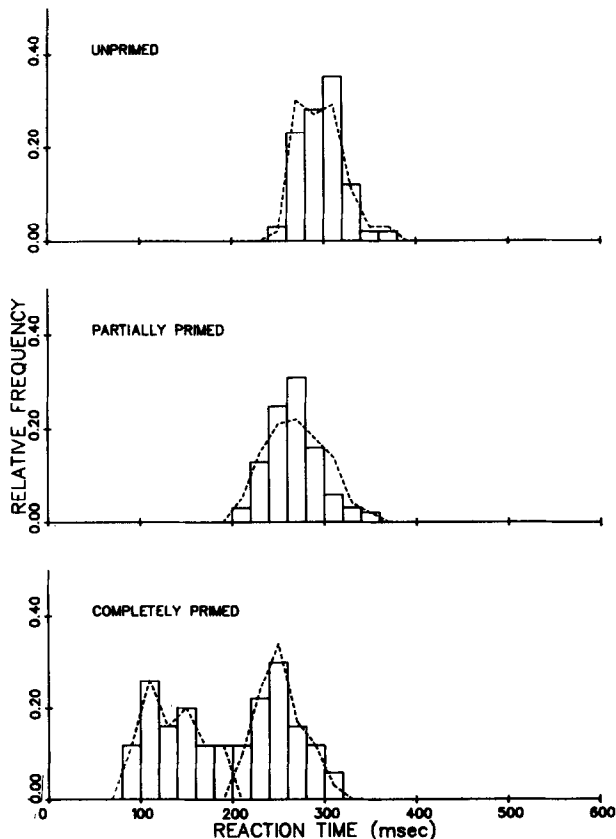


FIG. 8. Application of a three-state discrete model for the reaction-time distributions produced by subject J.F. when the prime stimuli were words.

of the partially primed distribution (middle panel, solid histogram) fell at about the same point, midway between the two peaks of the completely primed distribution, suggesting that 200 ms might represent a lower bound for a moderate state of preparation. Second, the reaction times of anticipations on catch trials had a maximum of about 200 ms, and their distribution appeared similar to the faster part (left peak) of the completely primed distribution on regular trials.<sup>22</sup> This suggested that the selected point of division might represent an upper bound for an extremely high state of preparation.

After dividing the completely primed distribution in this way, we normalized each of its two parts to have frequencies equal to one. Then a three-state discrete model was fit to these parts along with the partially primed distribution and unprimed distribution (top panel, solid histogram), testing a generalized mixture prediction [Eq. (4)] for them. The dotted curves in Fig. 8 represent the best-fitting frequency distributions that are consistent with the three-state model. Combined deviations of the fitted distributions from the data were not statistically significant [ $\chi^2(13) = 20.7$ ,  $p > .05$ ], unlike in the case of a two-state model. For example, the observed distribution under the partially primed condition appeared to closely match a mixture of the distributions associated with the unprimed condition and the right peak of the completely primed condition.

Other evidence of three discrete states of preparation emerged from a parallel analysis of J.F.'s reaction-time distributions for nonword primes (Fig. 9). Here again a value of 200 ms was used to divide the completely primed distribution into two parts, corresponding to its right and left peaks. Then we fit a three-state discrete model to these different parts along with the unprimed and partially primed distributions. Deviations from a generalized mixture prediction were not statistically significant [ $\chi^2(13) = 20.9$ ,  $p > .05$ ]. It appeared that the two parts of J.F.'s distribution for nonword primes in the completely primed condition came from two distinct states of preparation, only one of which (i.e., a moderately prepared state) was entered during trials with a medium priming interval.

The results of subject N.Y., who also had relatively large priming effects and many anticipatory responses, can be analyzed in the same way.

<sup>22</sup> The mean and standard deviation of J.F.'s anticipatory reaction times induced by the word primes on catch trials in the completely primed condition were 169 and 13 ms, respectively. On regular trials, the corresponding mean and standard deviation of the reaction times from the faster part (i.e., left peak) of the completely primed distribution (Fig. 8, bottom panel, solid histogram) were 133 and 24 ms, respectively. The similarity of these means and standard deviations is consistent with an assumed equivalence between the preparatory states that led to the anticipations on catch trials and the extremely fast responses on regular trials.

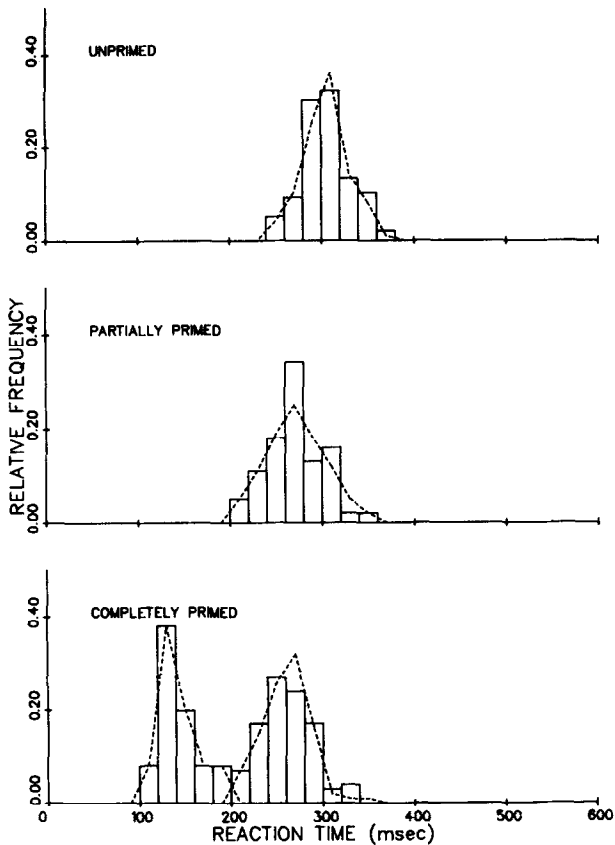


FIG. 9. Application of a three-state discrete model for the reaction-time distributions produced by subject J.F. when the prime stimuli were nonwords.

The mixture prediction of a three-state model provided a good fit to N.Y.'s reaction-time distributions for the nonword primes [ $\chi^2(10) = 17.0$ ,  $p > .05$ ], and a somewhat improved fit for the word primes [ $\chi^2(9) = 27.5$ ,  $p < .01$ ]. As in J.F.'s case, N.Y. apparently dealt with the test stimuli from either an unprepared, moderately prepared, or extremely prepared state, depending on the length of the interval between the presentation of the prime and test stimuli.

### Discussion

*Theoretical implications.* The results of Experiment 1 illustrate the ability of the adaptive response-priming procedure to provide interesting insights about the time course of human information processing. We found clear evidence of a limited number of discrete preparatory states whose probabilities of occurrence varied as a function of the length of

the priming interval. Two of the subjects (B.C. and S.R.) produced data consistent with the mixture prediction of the two-state discrete stage model, and the other two (J.F. and N.Y.) produced data consistent with a three-state model. There was no evidence to support a prototypical cascade model or other members of the continuous class that assume gradually increasing preparation over time. It appears, in particular, that the processing of words and nonwords as priming stimuli did not induce a continuous growth of response activation under the present circumstances. This outcome raises doubts about the generality of arguments against discrete stage models for word recognition (cf. McClelland, 1979).

Of course, our results must be interpreted cautiously. Experiment 1 involved a relatively restricted type of priming situation. The absence of additional preparatory states may have occurred here because of idiosyncrasies in the connections between the prime stimuli, test stimuli, and responses. Word primes do not ordinarily cue the impending occurrence of right arrows or right index-finger keypresses, and nonwords do not ordinarily cue the impending occurrence of left arrows or left index-finger keypresses. Perhaps this precluded the kind of automatic, as opposed to controlled, priming commonly associated with continuous activation mechanisms (e.g., see Neely, 1976). The apparent number of preparatory states may also have been limited by the relative ease of the simple arrow discrimination task (Swenson, 1972). More research is needed to determine exactly what constraints cause component mental processes to produce discrete or continuous outputs.

Nevertheless, in several respects, Experiment 1 sets the stage for such research, and constitutes an initial step toward a more general taxonomy of information-processing tasks. One of our subsequent studies (Experiment 3) is designed to take us a next step along that way. Another study (Experiment 2) focuses more closely on the nature of the three discrete states of preparation that we have discovered already in looking at the results from subjects J.F. and N.Y.

*Interpretation of the three preparatory states.* A possible interpretation of the three inferred preparatory states follows from certain facts about J.F. and N.Y.'s anticipatory responses on catch trials in the completely primed condition. The relative frequencies of their anticipations averaged .26, roughly equaling the relative frequencies of their reaction times that were less than 200 ms on regular trials in the completely primed condition. Also, their anticipations usually embodied the information conveyed by the prime stimuli. For example, when the prime stimuli on catch trials were words, the subsequent anticipations were almost invariably right keypresses, indicating expectations of a right-arrow test stimulus. On the other hand, nonword primes induced a strong tendency toward left key-



presses. These facts suggest that the fastest responses by J.F. and N.Y. in the completely primed condition were executed independently of the test stimuli, but took the identities of the prime stimuli into account.

Specifically, J.F. and N.Y. may have sometimes adopted a deadline strategy after fully processing the prime stimuli, if the priming interval was long enough. Under such a strategy, the termination of prime stimulus processing would start an internal timing mechanism that functions as an alarm clock for generating overt responses (Ollman & Billington, 1972). The timer would generate an interrupt after a predetermined lag, regardless of subsequent external stimulus events, leading to a response based on the identity of the prime stimulus (i.e., right keypresses for words and left keypresses for nonwords). No evaluation of the test stimulus would occur once the timing mechanism starts.

The deadline strategy accounts nicely for several aspects of J.F. and N.Y.'s performance. It explains why they produced approximately equal relative frequencies of very fast responses during regular and catch trials in the completely primed condition, because once the timing mechanism starts, further control of behavior due to the test stimulus is relinquished. If the occasional use of this mechanism were combined with a prior two-state process of preparation, as described by the two-state discrete stage model, then we should see evidence of three preparatory states, just as we did. The three states would include an unprepared state, a moderately prepared state in which the subject has programmed a response to a test stimulus through a discrete priming process but not relinquished control to the internal timer, and an extremely prepared state in which the timer has taken control of response production following the second state. Furthermore, this interpretation allows the results of B.C. and S.R. to be reconciled with those of J.F. and N.Y., assuming that the former subjects seldom if ever passed beyond the second (i.e., moderately prepared) state, yet otherwise performed much in the same way as the latter subjects (cf. Meyer et al., 1984).

The bias of some subjects toward anticipatory responses in Experiment 1 comes as no great surprise. Our instructions were specifically to "use the prime whenever possible to speed up the response to the test arrow." We discouraged errors on regular trials, but did not emphasize the importance of avoiding anticipations. Thus, it seems natural that at least some subjects (e.g., the more impulsive ones) would strive for an extreme state of preparation, not just a moderate one (Dickman, in press). Faced with a ratio of 75% regular trials to 25% catch trials, and reinforced by a substantial number of bonus points for fast regular-trial reaction times, the best way of maximizing the achieved point score might be to enter an extremely prepared state.

## EXPERIMENT 2

Experiment 2 was designed as a further test of the preceding ideas about the three discrete states of preparation revealed by Experiment 1. During the second experiment, we emphatically instructed subjects to avoid anticipations on catch trials. The negative feedback for committing such responses was increased accordingly. If a deadline strategy mediated the extremely prepared state found previously, then these more potent reinforcements against anticipations should reduce the apparent number of preparatory states, and all subjects' performance should accord more closely with the mixture prediction of the two-state discrete stage model.

To compensate for a decrease of the priming effects that might arise because of the altered instructions in Experiment 2, we made some other changes as well. In particular, the stimulus-response mapping for the adaptive response-priming procedure was modified. Subjects had to respond with their right index fingers when the test stimuli were right arrows, and with their right middle fingers when the test stimuli were left arrows. This mapping increased the spatial incompatibility between the test stimuli and responses, and it also increased response complexity by requiring keypresses with fingers on the same hand rather than homologous fingers on different hands. Both of these increases should tend to raise the latencies of unprimed responses (Fitts & Seeger, 1953; Shulman & McConkie, 1973), thus magnifying the potential benefit that the valid primes provide.<sup>23</sup>

We also increased the power of our statistical tests by including four priming conditions instead of just three. In Experiment 2, the conditions were unprimed, partially primed<sub>1</sub>, partially primed<sub>2</sub>, and completely primed. For the trials with partial and complete priming, words and non-words again served as the valid primes, cuing which responses should be made to the subsequent test stimuli. These two partially primed conditions involved different medium priming intervals. The first one had a medium<sub>1</sub> interval with a relatively short duration, and the second one had a medium<sub>2</sub> interval with a relatively long duration. This gave us greater potential for testing a three-state discrete model, should the need arise (cf. Experiment 1).

<sup>23</sup> With a relatively incompatible stimulus-response mapping, more complex processes are required to translate the test stimulus into an appropriate response, thereby increasing the observed reaction time when there is a neutral prime. However, the incompatibility of the mapping may not affect the reaction times when there is a valid prime and a long priming interval. Valid primes allow the translation between the test stimulus and response to be short circuited, yielding times no greater than would occur with a compatible mapping. Thus, the difference between the unprimed and completely primed reaction times should be greater.

Finally, another aim of Experiment 2 was to monitor and evaluate subjects' average level of alertness more precisely. We sought further evidence that the effect caused by manipulating the duration of the priming interval was not due mainly to variations in subjective alertness (cf. Bertelson, 1967; Posner & Boies, 1971). Between the initial and final warning signals in the unprimed condition, a neutral prime (row of four X's) was therefore displayed for a period of time that equaled either the medium<sub>1</sub>, medium<sub>2</sub>, or long priming interval on the trials with valid (i.e., word and nonword) primes. If there are substantial fluctuations of alertness over the priming interval, then manipulating the duration of the neutral prime should yield large reaction-time differences. By contrast, if the final warning signal eliminates general foreperiod effects, as we argued previously (Footnote 19), then the distributions of unprimed reaction times should not depend on the duration of the neutral prime. Also, the unprimed reaction-time distributions should have the same form as those obtained with the valid primes in the partially primed conditions when subjects happen to remain entirely unprepared.

### *Method*

*Subjects.* Four University of Michigan undergraduates served as paid subjects. Two of them (W.M. and J.G.) were males, and the other two (A.M. and S.M.) females. Each subject received approximately \$4 per session. None had been in Experiment 1.

*Apparatus and stimuli.* The apparatus and stimuli were identical to those in Experiment 1.

*Design.* The design was similar to that of Experiment 1 except for those changes needed to accommodate the four priming conditions and neutral prime stimulus. Four sessions were again run per subject, with the same number of trial blocks as before. During each test block, there were 64 trials, including 75% regular trials and 25% catch trials. A quarter of the trials were unprimed, a quarter were completely primed, and a half were partially primed. We used the medium<sub>1</sub> and medium<sub>2</sub> priming intervals on alternate blocks. Thus, over the course of a session, each of the four priming conditions occurred 25% of the time. The two types of test stimulus (i.e., right and left arrows) each had relative frequencies of .5 on the regular trials, and varied orthogonally with the duration of the priming interval. No test stimuli occurred on catch trials. Valid primes (i.e., words and nonwords) appeared equally often after the medium<sub>1</sub>, medium<sub>2</sub>, and long intervals. In the unprimed condition, the neutral prime always appeared instead. The duration of the neutral prime varied from trial to trial. On half of the unprimed trials, its duration was the same as for the valid primes presented with the medium interval assigned to that block. On the other half of the trials, the neutral prime had the same duration as the valid primes presented with the long priming interval. As before, word and nonword primes always cued the presentation of right and left arrows, respectively. Right and left arrows each followed the neutral prime on half of the trials, rendering it totally uninformative.

*Procedure.* The procedure was similar to that of Experiment 1. Responses were made with the index and middle fingers of the right hand, using the period and slash keys of the terminal keyboard. Subjects had to press either the slash key with their right middle finger when a left arrow appeared, or the period key with their right index finger when a right arrow appeared. Thus, the stimulus-response mapping was relatively incompatible (Fitts

& Seeger, 1953). The instructions and feedback strongly discouraged anticipations on catch trials.

*Priming intervals.* The long priming interval for the completely primed condition was again set at 805 ms. It included a duration of 720 ms for the display of the prime stimulus, plus 85 ms for the final warning signal. The two medium intervals for the partially primed conditions were set by the staircase tracking algorithm (Appendix 3). Across subjects and trial blocks, the medium<sub>1</sub> intervals used with word and nonword primes averaged 209 ms and 198 ms, respectively. The corresponding medium<sub>2</sub> intervals averaged 341 and 351 ms, respectively.

*Data analysis.* In analyzing the data, we first examined the reaction times for the unprimed condition to determine whether the duration of the neutral prime influenced performance through general foreperiod effects. No substantial differences emerged, indicating a reasonably constant level of subjective alertness.<sup>24</sup> Results in the unprimed condition were therefore combined across the various priming intervals used with the neutral primes.

The rest of the data analysis was similar to that of Experiment 1. We fit the mixture prediction of the two-state discrete stage model [Eq. (1)] to empirical distributions of reaction times from the unprimed, partially primed<sub>1</sub>, partially primed<sub>2</sub>, and completely primed conditions. For each subject, the fitted distributions came from the session that yielded partial-priming effects most appropriate (i.e., ones whose mean RTs fell midway between those from the extreme conditions) for maximizing the power of the goodness-of-fit tests. Because the fits were quite good, it was not necessary to test a three-state model, even though the total of four priming conditions used here would have allowed such tests.

## Results

Table 2 shows mean reaction times and standard deviations of correct responses on regular trials, error rates on regular trials, and anticipation rates on catch trials for each subject, type of prime stimulus, and priming condition. The mean reaction times decreased as the degree of priming increased. Overall, the priming effects averaged 25, 62, and 83 ms in the partially primed<sub>1</sub>, partially primed<sub>2</sub>, and completely primed conditions, respectively. The approximately equal separations among these effects is a successful product of the staircase tracking algorithm, which adjusted the medium<sub>1</sub> and medium<sub>2</sub> priming intervals to yield intermediate priming.

Statistical tests of the priming effects for each subject yielded highly significant outcomes. Out of 24 cases, there were 16 that reached the .001 significance level [ $t(90) \geq 3.9$ ]. Another 4 cases reached the .01 level [ $t(90) \geq 3.2$ ]. Of the 3 cases that failed to reach the .01 level, all occurred in the partially primed<sub>1</sub> condition, where we would expect the degree of priming to be relatively small, and the effects there still went in the right direction.

<sup>24</sup> Across subjects, the reaction times obtained with the neutral primes averaged 374, 370, and 373 ms when the priming interval was medium<sub>1</sub>, medium<sub>2</sub>, and long, respectively. Furthermore, the length of the priming interval did not markedly affect the error rates on regular trials or the anticipation rates on catch trials. The error rates averaged 0.8, 2.3, and 3.1% for the medium<sub>1</sub>, medium<sub>2</sub>, and long intervals, respectively; there were no anticipatory responses regardless of which interval was used.

TABLE 2  
Results for Individual Subjects in Experiment 2

Subject	Prime type	Priming condition	Mean RT (ms)	$\sigma$ (ms)	Errors (%)	Anticipations (%)
A.M.	Word	Unprimed	375	84	6.2	0.0
		Partial <sub>1</sub>	358	75	2.1	0.0
		Partial <sub>2</sub>	333	72	6.2	0.0
		Complete	283	67	2.1	12.5
	Nonword	Unprimed	378	56	0.0	0.0
		Partial <sub>1</sub>	340	63	2.1	0.0
		Partial <sub>2</sub>	316	62	2.1	6.2
		Complete	286	50	0.0	12.5
S.M.	Word	Unprimed	343	46	2.1	0.0
		Partial <sub>1</sub>	319	35	0.0	0.0
		Partial <sub>2</sub>	293	46	0.0	6.2
		Complete	277	54	0.0	0.0
	Nonword	Unprimed	372	47	0.0	0.0
		Partial <sub>1</sub>	363	65	6.2	0.0
		Partial <sub>2</sub>	342	55	2.1	0.0
		Complete	308	64	0.0	6.2
W.M.	Word	Unprimed	426	57	2.1	0.0
		Partial <sub>1</sub>	396	63	0.0	0.0
		Partial <sub>2</sub>	350	71	0.0	0.0
		Complete	306	80	0.0	0.0
	Nonword	Unprimed	427	53	0.0	0.0
		Partial <sub>1</sub>	388	56	2.1	0.0
		Partial <sub>2</sub>	335	77	0.0	0.0
		Complete	335	78	0.0	0.0
J.G.	Word	Unprimed	324	36	6.2	0.0
		Partial <sub>1</sub>	290	51	0.0	0.0
		Partial <sub>2</sub>	269	38	2.1	0.0
		Complete	263	39	0.0	6.2
	Nonword	Unprimed	333	50	0.0	0.0
		Partial <sub>1</sub>	325	71	2.1	0.0
		Partial <sub>2</sub>	287	43	0.0	0.0
		Complete	258	45	0.0	0.0

Error rates on regular trials were very low, averaging 2.1, 1.8, 1.6, and 0.3% in the unprimed, partially primed<sub>1</sub>, partially primed<sub>2</sub>, and completely primed conditions, respectively. A positive correlation was observed between the mean reaction times and error rates, indicating an absence of any substantial speed-accuracy trade-off. The anticipation rates on catch trials were also very low, exceeding zero only in the partially primed<sub>2</sub> and completely primed conditions, where they averaged 1.6 and 4.7%, respectively. This is a marked change from Experiment 1,

TABLE 3  
 Goodness-of-Fit Tests for the Mixture Prediction of the Two-State Discrete Stage Model  
 in Experiment 2

Subject	Prime type	$\pi(d_{m1})$	$\pi(d_{m2})$	$\chi^2$	<i>df</i>	<i>p</i>
A.M.	Word	.32	.43	15.9	14	>.30
	Nonword	.45	.70	15.5	14	>.30
S.M.	Word	.00	.52	21.7	16	>.10
	Nonword	.11	.56	11.8	14	>.50
W.M.	Word	.08	.51	16.8	16	>.30
	Nonword	.29	.91	21.5	16	>.10
J.G.	Word	.55	1.00	18.1	14	>.20
	Nonword	.11	.55	15.9	16	>.30

where some subjects (i.e., J.F. and N.Y.) had much stronger tendencies toward anticipations in the completely primed condition (Table 1).

Table 3 shows goodness-of-fit tests for the mixture prediction of the two-state discrete stage model, which was applied to the reaction-time distributions from each subject. There are eight tests, corresponding to the various combinations of subjects and prime types. Unlike some of the tests in Experiment 1, none in Experiment 2 yielded a statistically significant deviation from the two-state model [ $\chi^2(14) \leq 21.7$ ,  $p \geq .10$  in all cases], and most of the fits were quite good. The completely primed condition did not induce bimodal reaction-time distributions, and fewer than 1% of the responses on regular trials were faster than 200 ms, whereas such times had been occasionally prominent before (cf. Fig. 8 and 9).<sup>25</sup>

<sup>25</sup> In some cases, the estimates of the mixture parameters [ $\pi(d_{m1})$  and  $\pi(d_{m2})$ ] did have rather extreme values (i.e., close to zero or one), which is not desirable because it creates a degenerate situation for testing the mixture prediction. We therefore reran the goodness-of-fit tests shown in Table 3, including only those combinations of subjects and priming conditions for which the mixture parameters fell between 0.2 and 0.8. This raised the obtained significance levels somewhat, but still there were no statistically reliable deviations ( $p > .10$ ), further supporting the two-state discrete stage model. It should also be noted that the variances of reaction times in the partially primed conditions were not always larger than those in the unprimed and completely primed conditions (Table 2), which may appear to somewhat weaken the support for the two-state model. However, the model does not necessarily predict that the partially primed conditions should yield the largest variances, because the variances there depend on exactly what the values of the mixture parameters and means of the basis distributions are [Eq. (3)]. Some values of these parameters could lead the variances in the partially primed conditions to be intermediate (see Appendix 1).

*Discussion*

Like the two conservative subjects in Experiment 1 (B.C. and S.R.), each subject in Experiment 2 avoided anticipatory responses on catch trials and adhered to the mixture prediction of the two-state discrete stage model. The reaction-time data support the hypothesis that, in Experiment 1, the third state of extreme preparation exhibited by the two other subjects (J.F. and N.Y.) may have arisen through an anticipatory deadline strategy used on some trials of the completely primed condition. When we discouraged such anticipations, the third state disappeared and, instead, subjects took advantage of the prime stimuli in a more limited fashion, becoming either moderately prepared or else remaining entirely unprepared. There was no evidence that response preparation increased gradually during the priming interval, contrary to what a prototypical cascade model and some other members of the continuous class would predict.

Although the subjects in Experiment 2 did not enter an extremely high state of preparation, their priming effects were still relatively large. The average effect of complete priming was 83 ms for A.M., S.M., W.M., and J.G., whereas the conservative subjects of Experiment 1 exhibited an average effect of 63 ms. A larger priming effect presumably occurred here because the stimulus-response mapping was less compatible and thus allowed the valid prime stimuli to be more beneficial (Footnote 23). It is noteworthy that the decreased compatibility did not appear to induce a continuous process of preparation. This outcome, as we will discuss later, may help to interpret results of studies in which preparatory processes do seem to be more nearly continuous (Experiment 3).

What we found under the unprimed condition is also relevant to some other concerns surrounding the use of neutral prime stimuli. Based on an assumption that neutral primes neither facilitate nor inhibit subjects' performance, some investigators have estimated the "benefit" (facilitation) and "cost" (inhibition) of priming by comparing reaction times obtained with valid and invalid primes against those obtained with neutral primes (e.g., Neely, 1976, 1977; Posner & Snyder, 1975). Such cost and benefit estimates may reflect the relative contributions of controlled and automatic mental processes. However, critics have questioned whether the estimates are really meaningful (Antos, 1979; de Groot, Thomassen, & Hudson, 1982; Fischler & Bloom, 1979; Jonides & Mack, 1984; Schuberth & Eimas, 1977). Experiment 2 helps assuage these concerns. The distributions of partially primed reaction times obtained here conformed to an essentially perfect mixture of the unprimed and completely primed reaction times. It therefore appears that the neutral primes and the valid primes presented for a medium interval induced essentially the same ini-

tial unprepared state. This is what one would expect if, and only if, neutral primes are truly "neutral."

### EXPERIMENT 3

The purpose of Experiment 3 was to investigate the temporal properties of information processing in a somewhat more complex context than did Experiments 1 and 2. We again used the adaptive response-priming procedure, but increased the number of alternative test stimuli and responses. The relationship between the prime and test stimuli was also modified. These changes let us further assess the extent to which preparation is an all-or-none process as implied by the two-state discrete stage model.

The test stimuli for Experiment 3 consisted of upward vertical arrows presented at each of four different locations on a display screen. The responses consisted of keypresses made by the index and middle fingers of the right and left hands. We assigned the responses to the test stimuli through a mapping that was compatible in some respects and incompatible in others, as described below (see Method). This increased the magnitude of obtained priming effects, thereby enhancing the power of our goodness-of-fit tests.

Three priming conditions were included as in Experiment 1. For valid primes in the partially primed and completely primed conditions, we again used words and nonwords. The words cued responses by fingers on the right hand, and the nonwords cued responses by fingers on the left hand, but these primes were no longer completely informative about exactly what the test stimulus or response would be. There were also uninformative neutral primes, which we used in the unprimed condition. Because the prime stimuli never indicated exactly which test stimulus or response would occur, subjects could not fully prepare their responses ahead of time. This deterred anticipatory responses, allowing us to omit catch trials from the design.

Experiment 3 extends previous research by Rosenbaum (1980) and Miller (1982), who primed subjects with respect to response parameters such as the effector, the direction, and the extent of a movement. In their studies, there were four to eight test stimuli, with a different response assigned to each. Valid priming information was presented before (Rosenbaum, 1980) or as part of (Miller, 1982) the test stimuli. The results revealed that subjects could partially prepare the effector that would make a movement, even though certain aspects of the movement (e.g., its direction and/or extent) had not been completely determined yet. Miller (1982) found specifically that the hand used for making a manual response may be programmed before the selection of the type of finger (e.g., index or middle) that will produce the response. It was not clear



from his data, however, whether this hand preparation involves an all-or-none or a continuous process. The present procedure, which resembles Miller's (1982) in its ensemble of responses, helps resolve this issue. If the partially primed distributions of reaction times obtained here approximate perfect mixtures of the unprimed and completely primed distributions, then we would have evidence that hand preparation is a two-state discrete process. On the other hand, failure of the mixture prediction would implicate either a higher-order  $n$ -state ( $n > 2$ ) discrete process or a continuous process.

### Method

*Subjects.* Three University of Michigan undergraduates served as paid subjects. Two of them (L.K. and B.N.) were females, and the other (B.Z.) was male. Each subject received a payment of \$4 per session. None had been in Experiments 1 or 2.

*Apparatus.* The apparatus was the same as before.

*Stimuli.* The prime stimuli were identical to those in Experiment 2. The test stimuli were arrowheads formed from two vertically aligned carets ( $\blacktriangle$ ) on the display screen. Each test stimulus appeared above one of four adjacent locations arranged horizontally across the screen, with a separation of about  $0.4^\circ$  of visual angle between locations. The stimuli mapped in a crisscross fashion onto the four alternative keypress responses. Given this mapping, the left exterior and left interior test stimuli corresponded to movements of the right index and right middle fingers, respectively, whereas the right exterior and right interior test stimuli corresponded to movements of the left index and left middle fingers.

*Design.* The design was similar to those of Experiments 1 and 2. The relevant factors were priming condition (unprimed, partially primed, and completely primed), type of prime stimulus (word, nonword, and neutral), and response type. Each priming condition had a relative frequency of one-third. In the partially primed and completely primed conditions, half of the trials involved word primes, and the other half involved nonword primes. When the prime stimulus was a word, the subsequent test stimulus had a .5 probability of being a vertical arrow in the right exterior position, and a .5 probability of being a vertical arrow in the right interior position. When the prime stimulus was a nonword, the subsequent test stimulus had a .5 probability of being a vertical arrow in the left exterior position, and a .5 probability of being a vertical arrow in the left interior position. Thus, as mentioned earlier, the valid primes indicated which hand but not which specific finger to use for responding. In the unprimed condition, the prime stimulus was always neutral (i.e., a row of four X's). Each of the four types of test stimuli (arrows) and associated responses followed the neutral prime equally often, making it entirely uninformative. There were no catch trials.

*Procedure.* The procedure was similar to that of Experiment 2. On each trial, an initial warning signal (pound signs) appeared for 500 ms. Next a four-character prime stimulus was presented. The duration of the prime depended on the priming condition. To terminate the priming interval, the prime stimulus was removed, and four dashes (----) were displayed in a horizontal row for 85 ms where the characters of the prime had been. The dashes served as a final warning signal and as a set of fixation markers. After the final warning signal, one of the four dashes changed abruptly into a test stimulus. Subjects responded by pressing either the Z, X, period (.), or slash (/) key on the terminal keyboard, which rested under the left middle, left index, right index, and right middle fingers, respectively.

*Priming intervals.* For half of the trials in the unprimed condition, the neutral prime had a duration equal to the medium priming interval, and for the other half, it had a duration

equal to the long priming interval. The long interval, which was also used with valid primes in the completely primed condition, equaled 805 ms. The medium interval was set through the staircase tracking algorithm (Appendix 3). Separate medium intervals were tracked for each possible combination of prime stimulus and response type. The durations of the medium intervals averaged 361, 330, 284, and 313 ms across subjects for the left index, left middle, right index, and right middle fingers, respectively. There was a moderate effect of the prime type (i.e., word vs nonword) on the duration of the medium priming intervals. Word primes yielded medium intervals that were 47 ms less on the average than those for nonword primes, consistent with previous results (Experiment 1).

*Data analysis.* The data were analyzed as in Experiments 1 and 2. We again checked whether the duration of the neutral primes in the unprimed condition affected performance. No significant differences emerged.<sup>26</sup> Thus, as part of subsequent analyses, results from the medium and long priming intervals of the unprimed condition were combined for comparisons with results from the partially primed and completely primed conditions.

### Results

Tables 4, 5, and 6 summarize the mean reaction times and standard deviations of correct responses and the error rates for subjects B.Z., B.N., and L.K., respectively. Table 7 contains the goodness-of-fit tests used to compare these subjects' empirical reaction-time distributions against the mixture prediction of the two-state discrete stage model.

Before examining the results of individual subjects in more detail, several general facts should be highlighted. Error rates were very low, averaging 0.8, 1.4, and 0.4% in the unprimed, partially primed, and completely primed conditions. The mean reaction times decreased as the degree of priming increased, being 450, 389, and 350 ms, respectively, under the three conditions. The priming effects averaged 61 and 100 ms in the partially primed and completely primed conditions, and they were generally consistent across the individual subjects [ $t(77) \geq 2.0$ ,  $p < .05$ , in all but one case, i.e., subject L.K. for the partially primed condition involving nonword primes]. From each subject's reaction-time distributions, an interesting pattern may be induced. In particular, the goodness-of-fit between the two-state discrete stage model and the data appears to depend systematically on which combinations of test stimuli and responses are considered.

*Subject B.Z.* For subject B.Z., the two-state model's mixture prediction was supported when the test stimuli required responses by either of the two index fingers (Table 7). The fit in the case of the right index finger did not yield significant deviations [ $\chi^2(6) = 5.6$ ,  $p > .30$ ], and a similarly good fit occurred in the case of the left index finger [ $\chi^2(6) = 7.8$ ,  $p >$

<sup>26</sup> Across response types and subjects, reaction times averaged 452 and 448 ms, respectively, when the duration of the neutral prime was medium and long. Errors occurred on less than 1% of the trials in the unprimed condition regardless of how long the priming interval was. There were no anticipatory responses.

TABLE 4  
Results for Subject B.Z. in Experiment 3

Prime type	Response type	Priming condition	Mean RT (ms)	$\sigma$ (ms)	Errors (%)
Word	ri	Unprimed	441	67	0.0
		Partial	385	76	0.0
		Complete	339	65	0.0
	rm	Unprimed	549	80	0.0
		Partial	432	92	2.4
		Complete	368	96	0.0
Nonword	li	Unprimed	475	65	0.0
		Partial	399	73	2.4
		Complete	363	72	0.0
	lm	Unprimed	531	101	0.0
		Partial	430	79	0.0
		Complete	368	128	2.4

*Note.* ri = right index; rm = right middle; li = left index; lm = left middle.

.20]. Figure 10 illustrates this outcome. It shows B.Z.'s unprimed, partially primed, and completely primed reaction times (solid histograms) together with best-fitting frequency distributions (dashed curves) when the test stimulus was a left exterior arrow and the response was a key-press by the right index finger. The pattern is similar to what happened for subjects B.C. and S.R. in Experiment 1 (e.g., Figs. 6 and 7).

TABLE 5  
Results for Subject B.N. in Experiment 3

Prime type	Response type	Priming condition	Mean RT (ms)	$\sigma$ (ms)	Errors (%)
Word	ri	Unprimed	390	39	2.4
		Partial	329	48	0.0
		Complete	304	47	0.0
	rm	Unprimed	362	92	0.0
		Partial	329	87	0.0
		Complete	287	60	0.0
Nonword	li	Unprimed	375	67	0.0
		Partial	336	39	0.0
		Complete	330	35	0.0
	lm	Unprimed	401	98	2.4
		Partial	331	68	9.5
		Complete	272	63	2.4

*Note.* ri = right index; rm = right middle; li = left index; lm = left middle.

TABLE 6  
Results for Subject L.K. in Experiment 3

Prime type	Response type	Priming condition	Mean RT (ms)	$\sigma$ (ms)	Errors (%)
Word	ri	Unprimed	511	80	0.0
		Partial	434	88	0.0
		Complete	399	64	0.0
	rm	Unprimed	517	48	0.0
		Partial	456	84	0.0
		Complete	406	77	0.0
Nonword	li	Unprimed	425	72	0.0
		Partial	415	89	2.4
		Complete	379	37	0.0
	lm	Unprimed	422	83	4.8
		Partial	394	49	0.0
		Complete	383	65	0.0

*Note.* ri = right index; rm = right middle; li = left index; lm = left middle.

In contrast, the two-state model fit B.Z.'s reaction times much less well for the middle-finger keypresses. For example, Figure 11 shows what happened when the test stimulus was a right interior arrow and the response was a keypress by the left middle finger. Here the deviation of the actual data from the best-fitting frequency distributions that satisfy the model's mixture prediction was highly significant [ $\chi^2(6) = 30.6$ ,  $p <$

TABLE 7  
Goodness-of-Fit Tests for the Mixture Prediction of the Two-State Discrete Stage Model in Experiment 3

Subject	Prime type	Response type	$\pi(d)$	$\chi^2$	$df$	$p$
B.Z.	Word	ri	.59	5.6	6	>.30
		rm	.72	13.5	6	<.05
	Nonword	li	.72	7.8	6	>.20
		lm	.10	30.6	6	<.001
B.N.	Word	ri	.85	9.5	6	>.10
		rm	.83	11.8	6	<.10
	Nonword	li	.78	4.0	6	>.50
		lm	.28	21.6	6	<.01
L.K.	Word	ri	.74	7.5	6	>.20
		rm	.70	17.8	6	<.01
	Nonword	li	.21	4.7	6	>.50
		lm	.35	1.5	6	>.50

*Note.* ri = right index; rm = right middle; li = left index; lm = left middle.

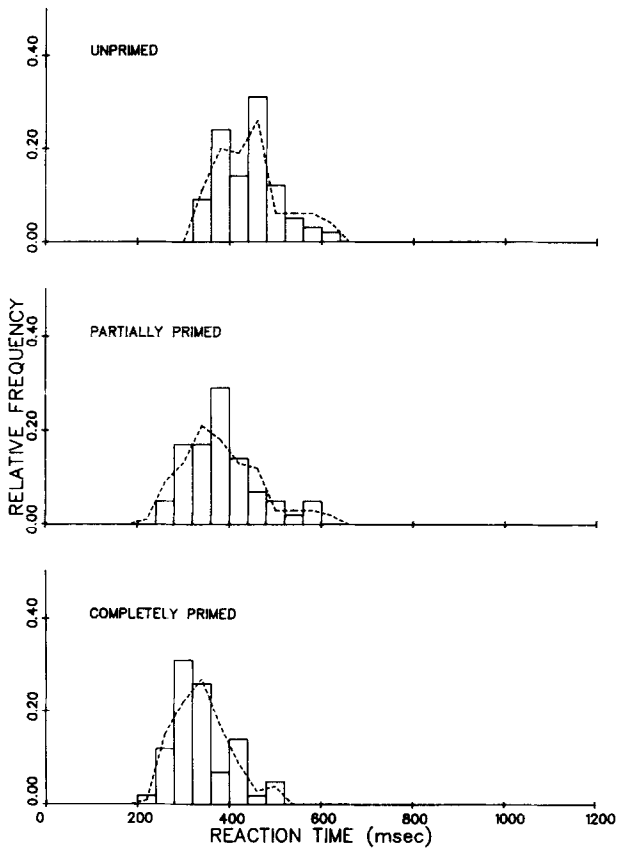


FIG. 10. Observed distributions of reaction times (solid histograms) and best-fitting frequency distributions based on the two-state discrete stage model (dashed curves) for subject B.Z. in Experiment 3 when the responses were keypresses by the right index finger and the condition was unprimed (top panel), partially primed (middle panel), or completely primed (bottom panel).

.001]. A significant deviation also occurred when the responses were right middle-finger keypresses [ $\chi^2(6) = 1.35, p < .05$ ].

It should be stressed that these deviations differ markedly from the failures of the two-state discrete stage model found for the less conservative subjects of Experiment 1 (i.e., J.F. and N.Y.). The middle-finger reaction times of B.Z. do not suggest a three-state discrete model of preparation (cf. Figs. 8 and 9). Instead, there is a pattern that appears to fit better with what a prototypical cascade model or other member of the continuous class would predict (e.g., Fig. 3). B.Z.'s partially primed distributions for the middle fingers had relatively small variances, were shaped similarly to those of the unprimed and completely primed distri-

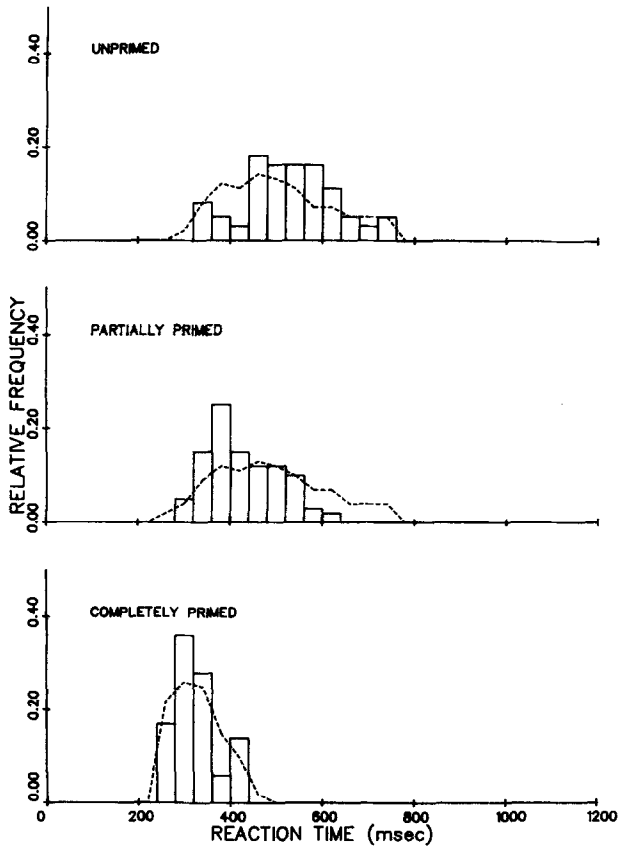


FIG. 11. Subject B.Z.'s observed reaction-time distributions and best-fitting frequency distributions for a two-state model when the responses were keypresses by the left middle finger.

butions, and did not overlap entirely with them. No strong evidence of bimodality occurred in any of the distributions for his middle fingers.

*Subjects B.N. and L.K.* An analogous pattern of good and bad fits to the two-state discrete stage model appears in the data of the other two subjects (Table 7). When they responded with the right and left index fingers, the data did not deviate significantly from the mixture prediction of the two-state discrete model. However, when they responded with the right and left middle fingers, considerably larger deviations emerged [ $\chi^2(6) \geq 11.8$ ,  $p < .10$ , in 3 of 4 cases]. Again the deviations tended toward what would be expected from a prototypical cascade model.<sup>27</sup>

<sup>27</sup> Only L.K.'s left middle-finger responses failed to follow the pattern of significant deviations exhibited by the other two subjects [ $\chi^2(6) = 1.5$ ,  $p > .50$ ]. This may have happened

*Discussion*

Experiment 3 produced some striking results. In all, there were six goodness-of-fit tests based on reaction-time distributions for index-finger responses. Every one gave relatively small  $\chi^2$  values ( $p > .10$ ), indicating a reasonably good fit to the mixture prediction of the two-state discrete stage model, and supporting the hypothesis of all-or-none preparation. For five of these tests, the fit was quite good ( $p > .20$ ). However, four of six tests based on reaction-time distributions from middle-finger responses produced rather large  $\chi^2$  values ( $p < .05$ ), failing to satisfy the mixture prediction. Three of these tests were highly significant ( $p < .01$ ). The middle-finger responses appeared to be more in line with predictions derived from a prototypical cascade model. They produced partially primed reaction-time distributions with single modes, moderate variances, and tails that did not extend as far as those of the unprimed and completely primed distributions, suggesting a form of continuous rather than all-or-none preparation.

This pattern cannot simply be attributed to differences in the lexical status of the prime stimuli or to the incompatibility of the stimulus-response mapping. Word and nonword primes each yielded some evidence for the two-state discrete model as well as some against it. Thus, another factor presumably modulates if and when preparation is a discrete as opposed to continuous process. That factor seems to depend on more than just a variation of stimulus-response compatibility, since a relatively incompatible mapping was also used in Experiment 2, but no violations of the mixture prediction occurred there.

A possible clue to understanding the present results stems from the fact that Experiment 3 included four alternative responses and partially informative, rather than completely informative, prime stimuli. For example, suppose that subjects try to simultaneously prepare responses by both the index and middle fingers of a hand, based on the cues extracted from such primes. Suppose also that the preparation process has limited capacity, and that readying each response consumes some of the available resources (cf. Gopher & Sanders, 1984). Then it might be relatively difficult, even impossible, to have more than one response completely prepared at a given moment. Perhaps preparation is easier if at least some responses are activated gradually, as postulated by a prototypical cascade model.<sup>28</sup>

---

because those responses enjoyed the smallest complete-priming effect (39 ms) in the entire data set of Experiment 3, resulting in relatively low power for our maximum-likelihood statistical technique.

<sup>28</sup> Given the above account, the choice of exactly which responses to prepare in an all-or-none fashion and which to prepare gradually could depend on additional factors. Ana-

In this vein, it is interesting to recall that Experiment 2 also included both middle-finger and index-finger responses. However, only two responses (i.e., right index and right middle) were required there, and the prime stimuli were completely informative, cuing a single response on each trial. The results supported the mixture prediction of the two-state discrete stage model for each type of finger, unlike in Experiment 3. Such an outcome implies that middle-finger responses do not necessarily preclude all-or-none preparation, when only one response has to be prepared at a time. Further research is therefore needed to determine exactly what the capacity requirements of preparing various finger movements are.

### GENERAL DISCUSSION

Experiments 1 through 3 together document the complexity of human information-processing dynamics. Our observations based on the adaptive response-priming procedure suggest that discrete preparatory states may occur under some circumstances but not others. We have tried initially to specify what some of the relevant determinants are. They include the number of alternative responses to be prepared and the speed-accuracy trade-off strategies of individual subjects, which influence biases toward or away from anticipatory responses. Some groundwork has thus been laid for moving beyond the present application of our procedure to a more general consideration of priming phenomena and their time course.

As intended, the present results bear on the issue of when it is appropriate to use conventional reaction-time methods for studying information-processing dynamics. We have reported evidence of discrete preparatory states when subjects must contend with only two response alternatives. A two-state discrete stage model holds fairly well there, especially if subjects refrain from producing anticipatory responses. This bolsters the argument that measurements of reaction times and additive-factor analyses (Sternberg, 1969) are useful when an experimental task involves a limited number of responses. However, the argument now seems less persuasive for tasks that incorporate greater numbers of responses. With nonbinary choice-reaction tasks, there may be a stronger tendency toward continuous rather than discrete processing, which would increase the likelihood of complex speed-accuracy trade-offs and de-

---

tomical studies suggest that the execution of the middle-finger responses may require more complex muscular control than does the execution of index-finger responses (Marble, 1960). As a result, more capacity might be needed to prepare responses by the middle fingers than to prepare responses by the index fingers. Perhaps these different requirements bias subjects toward using an all-or-none preparation process for the index fingers and a continuous process for the middle fingers.



crease the potential applicability of conventional reaction-time methods. Speed-accuracy trade-off methods could have more utility where multiple responses are required (cf. Wickelgren, 1977).

### *Power of the Statistical Analyses*

Another issue still of importance concerns the power of our statistical analyses. In Experiments 1 and 2, we obtained a good fit between a two-state discrete stage model and the data of some subjects, whereas the data of other subjects were fit better by a three-state model. In Experiment 3, we again obtained a good fit of the two-state model for certain subsets of data (i.e., those produced when subjects responded with their index fingers), but the cascade model seemed to fit better for others (i.e., those produced when subjects responded with their middle fingers). This indicates that our analyses were sufficiently powerful to discriminate between interesting alternative models, given the amount of data collected here. Nevertheless, it would be helpful to further assess the degree of power.

We have therefore performed some stochastic computer simulations with a continuum of models that range from a two-state discrete model to a prototypical cascade model.<sup>29</sup> These simulations built on our previous efforts to compare and contrast the predictions of the cascade model with those of the two-state model (Figs. 1-3). They provide a firmer basis for specifying exactly how much power the present data analyses had against various members of the discrete and continuous classes. Such a specification is particularly important in light of our previous exposition of the two classes, where we pointed out that some cases of continuous models may be able to closely mimic the properties of discrete models, and that conversely, some cases of discrete models may be able to closely mimic the properties of continuous models (see Theoretical Caveats). By finding the "twilight zone" where one class fades into the other from a standpoint of empirical discriminability, one can judge whether or not a set of experiments has enough power for one's intended purposes.

Figure 12 outlines how our simulations worked. As before (Fig. 2), we began with a prototypical cascade model in which the presentation of a prime stimulus caused activation to grow in an exponential approach from an initial base level  $b_p$  to an asymptote  $a_p$ . This growth is represented by the leftmost continuous curve of the figure, whose parameters are similar to those chosen previously. We further assumed that on top of the prime

<sup>29</sup> A simulation approach was followed because straightforward analytical solutions do not currently exist for predictions by the cascade model when the rate parameters of the activation-growth curves vary randomly (Ashby, 1982).

activation, the test stimulus caused additional activation to arise. The test-stimulus activation, which is not shown in Fig. 12, had the same general form as in Fig. 2. "Reaction time" was defined to be the amount of time from the onset of the test stimulus until the overall activation crossed the final response threshold. The value of the asymptotic activation,  $a_p$ , induced by the prime stimulus fluctuated randomly in a Gaussian fashion from one simulated trial to the next, yielding reaction-time distributions rather than just constant times.

Next we let the form of the activation due to the prime vary in other respects. A random fluctuation occurred in the total amount of time,  $\tau_9$ , between the onset of the prime stimulus and the moment when its activation-growth curve reached 90% of asymptote. Also, there was a parameter  $\theta$  that represented a ratio of two quantities: (i) the amount of dormant time during which activation remained at the base level after the onset of the prime, and (b) the time  $\tau_9$  taken for prime activation to reach 90% of asymptote, measured from when the prime was presented. For example, the middle prime-activation growth curve in Fig. 12 has a value of  $\theta$  equal to one half.

The parameter  $\theta$  may be viewed as a quantitative indicator of the degree to which a particular model along the simulated continuum is quasi-discrete or quasi-continuous. As  $\theta$  approaches 1, the prime-activation growth curves become more and more like a step function, mimicking the all-or-none preparation associated with the two-state discrete stage

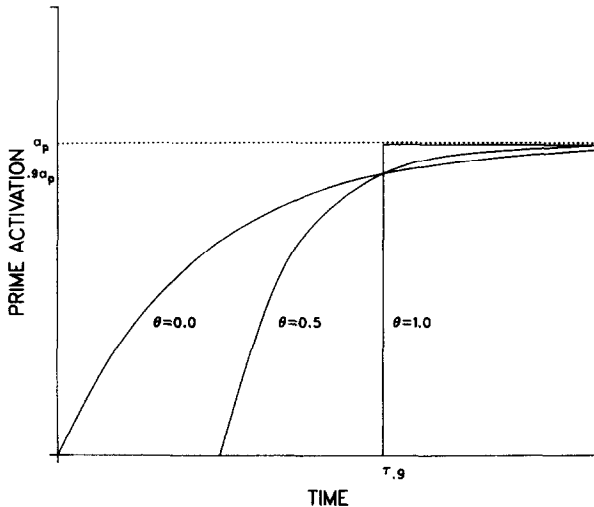


FIG. 12. Schematic representation of the stochastic computer simulation used to assess the statistical power of the reported goodness-of-fit tests.

model. Models for which the values of  $\theta$  are relatively near to 1 have a marked quasi-discrete quality for two reasons. First, the large amount of dormant time incorporated in them constitutes a period during which the processing system lacks partial information for all practical purposes, just as in the two-state model. Second, as  $\theta$  approaches 1, the rate of growth in prime activation ( $r_p$ ) becomes extremely large, approaching an instantaneous transition from the initial base level to the asymptote. By contrast, models for which the values of  $\theta$  are near 0 have a marked continuous quality.

The simulations let us assess the effect of the priming interval on the form of the obtained reaction-time distributions as a function of  $\theta$ . Specifically, our aim was to see how well the mixture prediction of the two-state discrete stage model would fit the simulated data, given different  $\theta$ -values and samples of simulated reaction times with approximately the same means, coefficients of variation, and numbers of observations ( $N = 100$ ) per distribution as in Experiments 1 through 3. To the extent that the mixture prediction can be rejected by the present goodness-of-fit tests for values of  $\theta$  greater than 0, this would demonstrate high power against a range of instantiations of the cascade model.

Figure 13 shows the results of the simulations. On the horizontal axis is  $\theta$ . On the vertical axis is a measure of power,  $\beta$ , the probability of rejecting the mixture prediction of the two-state discrete model. When  $\theta$  equals 0, which corresponds to a complete absence of dormant time, the

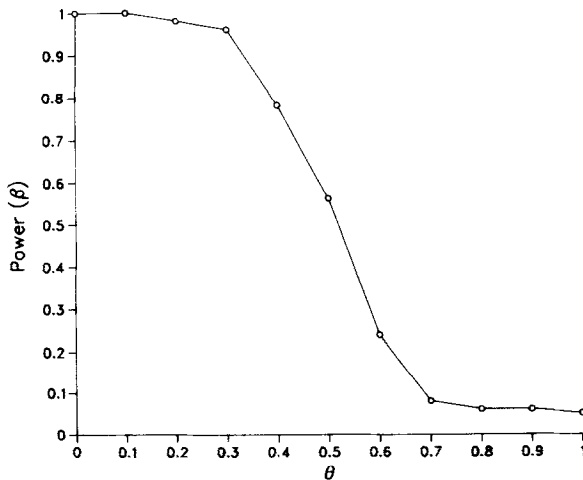


FIG. 13. A power curve that shows the probability ( $\beta$ ) of rejecting the mixture prediction of the two-state discrete stage model as a function of the parameter  $\theta$ , which represents the relative amount of dormant time before activation starts to grow after the onset of a prime stimulus under the cascade model (cf. Fig. 12).

power is very high ( $\beta > .99$ ), indicating high discriminability between the two-state model and a prototypical cascade model. The power continues to remain relatively high ( $\beta > .5$ ) until  $\theta$  exceeds .5. When  $\theta$  reaches .7 or greater, representing a large amount of dormant time, the power has dropped to about the .05 significance level associated with a Type I error [i.e., rejecting the null (two-state) hypothesis when it is true]. In this region, ( $.7 \leq \theta < 1$ ), we cannot discriminate various instantiations of the cascade model from the two-state model. However, the lack of power is not especially troublesome here. Matters would have to be stretched considerably to claim that cases with  $\theta$  greater than .7 preserve much of the original spirit of the cascade model. Our confidence in the strength of the present tests is therefore enhanced.

### *Criticisms of the Adaptive Response-Priming Procedure*

Of course, some criticisms may be directed against the response-priming procedure. To apply the procedure for our purposes, certain assumptions were needed about how subjects process the prime and test stimuli under the imposed task demands. These assumptions were not completely justified beforehand, and their validity could be questioned. This limits possible conclusions from the reported experiments, but we believe that the limitations are not too severe, and that there are some counterarguments to help disarm the criticisms.

*Self-termination of prime processing.* One assumption made earlier concerns the self-termination of prime processing. According to this assumption, subjects begin processing each prime stimulus immediately when it is presented, and persist until the onset of the test stimulus, at which moment the processing of the prime stops. In retrospect, the assumed self-termination has some firm basis. On roughly half the trials in the partially primed condition of Experiments 1 and 2, subjects performed the same as in the completely primed condition, but on the other half, performance was about the same as in the unprimed condition. This seems most likely to have happened because the processing of the prime stopped as soon as the test stimulus appeared. If the processing of the prime extended beyond the onset of the test stimulus, then subjects would fail to exhibit an unprepared state like the one observed in the partially primed condition.

*Utility of partial information.* A related point concerns the utility of partial information extracted before a prime stimulus has been completely processed. Our theoretical development assumed that subjects could and would use whatever partial information became available from processing a prime stimulus during the priming interval, even though their attention had to be shifted to the test stimulus as soon as it appeared. Some of the results from Experiment 3 support this assumption. Evidence

emerged there against a two-state discrete stage model for the preparation of responses by the middle fingers (Fig. 11). The reaction-time distributions of middle-finger responses were more in line with what a prototypical cascade model would predict. If partial information were generally ignored or unavailable, then such an outcome should not have occurred, given that partial information is part and parcel of the cascade model.

### *Directions for Future Research*

There are a number of potentially fruitful directions for future research with the adaptive response-priming procedure and analysis of reaction-time mixture distributions.

*Valid versus invalid prime stimuli.* Although we have used only valid prime stimuli in our experiments, the adaptive response-priming procedure may be easily extended to investigate the processing of invalid primes. As mentioned previously, some theorists have proposed that invalid primes inhibit subjects' ability to deal quickly with subsequent test stimuli (Neely, 1976, 1977; Posner & Snyder, 1975). This inhibition could involve different mechanisms than does the facilitation produced by valid primes. For example, "controlled" mechanisms could underlie inhibition, while "automatic" mechanisms underlie facilitation. It is not yet known for certain which of these mechanisms produce discrete outputs and which produce continuous outputs. Our approach provides a natural way to pursue the matter further.

*Other substantive applications.* The adaptive response-priming procedure may also have other substantive applications where theoretical controversies over information-processing dynamics remain unresolved. For example, some theorists have characterized mental imagery in terms of transformations (e.g., rotation and zooming) that gradually adjust the orientation or size of visual images from one setting to another (Kosslyn, Pinker, Smith, & Schwartz, 1979; Shepard, 1975), while others have characterized the transformation process as being discrete in nature (Just & Carpenter, 1976). These alternative accounts can be assessed by observing how subjects perform with prime stimuli composed of visual forms in various orientations and sizes, testing for the presence of reaction-time mixture distributions (Goldberg, Meyer, Yantis, & Langolf, 1985).

Similarly, this approach may be applied to issues regarding the nature of selective visual attention and spreading activation in semantic memory. Some models of attention and spreading activation have postulated continuous processes and others have postulated discrete processes.<sup>30</sup> As in

<sup>30</sup> Tsai (1983) has suggested that shifting attention from one spatial location to another is analogous to the movement of a spotlight sweeping continuously across a spatial region.

the case of mental imagery, these models may be tested by presenting appropriate prime stimuli through the adaptive response-priming procedure and analyzing the resultant reaction-time distributions for the presence of predicted mixtures (Yantis, 1985).

*Tracing the psychometric priming function.* Finally, it may be instructive to trace the psychometric priming function,  $\Psi(d)$ , in more detail than we have done so far. By definition,  $\Psi(d)$  represents the probability that a reaction time observed under conditions of partial priming will not exceed a temporal cutpoint,  $T_x$ , when the interval between the prime and test stimuli equals  $d$ . We used  $\Psi(d)$  previously to illustrate the objective of our staircase tracking algorithm for determining the ideal medium priming intervals in the partially primed condition (Fig. 5). However, this function could have other uses as well.

The detailed shape of  $\Psi(d)$  can reveal more about the way that a given prime stimulus is processed. Suppose, in particular, that a two-state discrete stage model holds under a selected set of priming conditions. Then the amount of time from the presentation of the prime to the transition between the unprepared and fully prepared states may be represented by a random variable  $\tau$ , which has a cumulative distribution function  $G(d)$  and a probability-density function  $g(d)$ . Given this representation, several facts follow. First, the mixture parameter  $\pi(d)$  associated with a medium priming interval  $d$  would simply equal  $G(d)$ , because the probability that a subject enters the prepared state in the partially primed condition is determined strictly by whether or not the transition time  $\tau$  happens to have a value less than  $d$ . Second, the psychometric priming function would be linearly related to the distribution function  $G(d)$ , as follows:<sup>31</sup>

---

On the other hand, Shaw (1984) proposed a model of attention that can account for similar phenomena without appealing to the spotlight metaphor. Extant data do not provide a strong basis for choosing between these two alternatives.

<sup>31</sup> The derivations of Eqs. (10) and (11) are straightforward. From the definition of  $G(d)$ , we have  $G(d) = \pi(d)$ , where  $\pi(d)$  represents the mixture parameter associated with the two-state discrete model and a medium priming interval  $d$ . This relation together with Eq. (1) of the text yields the following equality:

$$P(t < T_x | d) = G(d)F_c(T_x) + [1 - G(d)]F_u(T_x),$$

where  $t$  is reaction time and  $T_x$  is the reaction-time cutpoint of the staircase tracking algorithm. Furthermore, by definition  $\Psi(d) = P(t < T_x | d)$ . Combining the latter two equalities and rearranging terms yields Eq. (10). Differentiating both sides of Eq. (10) with respect to  $d$ , we arrive immediately at Eq. (11) for the slope,  $\Psi'(d)$ , of the psychometric priming function. Looking at Eq. (11), it can be seen that  $\Psi'(d)$  depends directly on the difference between  $F_u(T_x)$  and  $F_c(T_x)$ , that is, on the amount of separation between the unprimed and completely primed distributions of reaction times. As their separation increases, the slope of  $\Psi(d)$  increases, making it easier to accurately locate the ideal medium priming interval,  $d_m^*$  (Appendix 3).

$$\Psi(d) = [F_c(T_x) - F_u(T_x)]G(d) + F_u(T_x), \quad (10)$$

where  $F_c(T_x)$  and  $F_u(T_x)$  are defined as before [Eq. (9)]. Third, the slope (first derivative)  $\Psi'(d)$  of the psychometric priming function would be proportional to the probability-density function  $g(d)$ , that is,

$$\Psi'(d) = [F_c(T_x) - F_u(T_x)]g(d). \quad (11)$$

With these equations, we can work backward from the observed shape of the psychometric priming function to infer the shape of the distribution of transition times for completely processing the prime stimuli. The transition-time distribution  $g(d)$  would reveal details of how subjects encode a prime stimulus, use it to retrieve stored information, and prepare motor programs. Moreover, because the subjects do not have to respond overtly to the primes, the shapes of  $\Psi(d)$  and  $G(d)$  could provide purer reflections of these initial processing stages than do ordinary reaction times, which include an additional response-execution component (Sternberg, 1969). Analysis of the shape of  $\Psi(d)$  therefore represents a significant complement to standard reaction-time techniques (cf. Sabol & DeRosa, 1976; Yellott & Hildreth, 1969). An investigator may obtain data for this analysis simply by varying the duration of the priming interval over a reasonably wide range.

## APPENDIX 1

### Derivation of Reaction-Time Variances for an $n$ -State Discrete Model

Suppose that  $x$  is a continuous random variable,  $y$  is a discrete random variable whose sample space consists of the set of elements  $\{y_i; i = 1, 2, \dots, n\}$ , and  $z$  is a nonrandom variable. Suppose also that  $x$  is a stochastic function of  $y$ , and that  $y$  is a stochastic function of  $z$ . Then from the axioms of conditional probability theory, we have

$$E(x|z) = \sum_{i=1}^n E(x|y_i)P(y_i|z), \quad (1.1)$$

$$V(x|z) = E(x^2|z) - [E(x|z)]^2 = \left[ \sum_{i=1}^n E(x^2|y_i)P(y_i|z) \right] - [E(x|z)]^2, \quad (1.2)$$

where  $E(x|z)$  is the mean (expected value) of  $x$  given  $z$ , and  $V(x|z)$  is the corresponding variance.

Now consider an  $n$ -state discrete model ( $n \geq 2$ ). Let  $x$  be reaction time  $t$ ,  $y_i$  be preparation state  $i$  ( $i = 1, 2, \dots, n$ ), and  $z$  be the duration  $d$  of the medium priming interval in the partially primed condition of our

adaptive response-priming procedure. Then, by definition,  $M_p(d) = E(x|z)$ ,  $V_p(d) = V(x|z)$ ,  $M_i = E(x|y_i)$ , and  $\pi_i(d) = P(y_i|z)$ , where  $M_p(d)$ ,  $V_p(d)$ ,  $M_i$ , and  $\pi_i(d)$  are as in the text. Substituting these expressions into Eqs. (1.1) and (1.2) yields

$$M_p(d) = \sum_{i=1}^n \pi_i(d)M_i, \quad (1.3)$$

$$V_p(d) = \left[ \sum_{i=1}^n \pi_i(d)E(t^2|\text{state } i) \right] - [M_p(d)]^2. \quad (1.4)$$

Equation (2) in the text is equivalent to Eq. (1.3) with  $n = 2$ . Moreover,  $E(t^2|\text{state } i) = V_i + M_i^2$ , where  $V_i$  is the reaction-time variance under state  $i$ . Thus, Eq. (1.4) can be rewritten as

$$\begin{aligned} V_p(d) &= \left[ \sum_{i=1}^n \pi_i(d)(V_i + M_i^2) \right] - [M_p(d)]^2 \\ &= \left[ \sum_{i=1}^n \pi_i(d)V_i + \sum_{i=1}^n \pi_i(d)M_i^2 \right] - [M_p(d)]^2. \end{aligned} \quad (1.5)$$

The middle term in Eq. (1.5) may be rewritten as

$$\begin{aligned} \sum_{i=1}^n \pi_i(d)M_i^2 &= \sum_{i=1}^n \pi_i(d)[M_i - M_p(d) + M_p(d)]^2 \\ &= \sum_{i=1}^n \pi_i(d)[M_i - M_p(d)]^2 + 2M_p(d) \sum_{i=1}^n \pi_i(d)[M_i - M_p(d)] \\ &\quad + \sum_{i=1}^n \pi_i(d)[M_p(d)]^2 \\ &= \sum_{i=1}^n \pi_i(d)[M_i - M_p(d)]^2 + [M_p(d)]^2, \end{aligned} \quad (1.6)$$

where the last line of this equation results from applying Eq. (1.3). Substituting Eq. (1.6) into Eq. (1.5) gives

$$V_p(d) = \sum_{i=1}^n \pi_i(d)V_i + \sum_{i=1}^n \pi_i(d)[M_i - M_p(d)]^2, \quad (1.7)$$

which is the same as Eq. (6) in the text.

Eq. (3) in the text is simply a special case of Eq. (1.7) with  $n = 2$ . To see this, note that with  $n = 2$ , we have  $\pi_1(d) = \pi(d)$ ,  $\pi_2(d) = [1 - \pi(d)]$ , and

$$\begin{aligned} M_1 - M_p(d) &= M_1 - \{\pi(d)M_1 + [1 - \pi(d)]M_2\} \\ &= [1 - \pi(d)](M_1 - M_2), \end{aligned} \quad (1.8)$$



$$\begin{aligned} M_2 - M_p(d) &= M_2 - \{\pi(d)M_1 + [1 - \pi(d)]M_2\} \\ &= -\pi(d)(M_1 - M_2), \end{aligned} \quad (1.9)$$

where  $\pi(d)$  is the mixture parameter of the two-state model. Substituting Eqs. (1.8) and (1.9) back into Eq. (1.7) yields

$$\begin{aligned} V_p(d) &= \pi(d)V_1 + [1 - \pi(d)]V_2 + \pi(d)[1 - \pi(d)]^2(M_1 - M_2)^2 \\ &\quad + [1 - \pi(d)][\pi(d)]^2(M_1 - M_2)^2 \\ &= \pi(d)V_1 + [1 - \pi(d)]V_2 + \pi(d)[1 - \pi(d)](M_1 - M_2)^2. \end{aligned} \quad (1.10)$$

But  $V_1 = V_u$ ,  $M_1 = M_u$ ,  $V_2 = V_c$ , and  $M_2 = M_c$ . Thus, Eq. (1.10) is equivalent to Eq. (3) in the text.

From Eq. (1.10), we see that the variance of the partially primed distribution under the two-state model may easily be larger than the variances of the two basis distributions. For example, suppose that  $f_u(t)$  and  $f_c(t)$  have the same coefficient of variation,  $\kappa$ , as defined by the following equation:

$$\kappa = S_u/M_u = S_c/M_c, \quad (1.11)$$

where  $S_u$  and  $S_c$  are the standard deviations of the unprimed and partially primed distributions. Suppose also that  $S_c < S_u$ , which we expect because  $M_c$  will generally be less than  $M_u$ . Then  $V_p(d)$  will be greater than  $\max(V_u, V_c)$  whenever

$$M_c < [1 - \pi(d) - \kappa^2]M_u/[1 - \pi(d) + \kappa^2]. \quad (1.12)$$

Inequality (1.12) means that with  $\pi(d) = .5$  and  $\kappa = .15$ ,  $V_p(d)$  will exceed both  $V_u$  and  $V_c$  whenever  $M_c < .91M_u$ , a common occurrence under our experimental conditions.

The proof of the relationship between the coefficient of variation, mean reaction times, and reaction-time variances is straightforward. We wish to show that if Eq. (1.11) and inequality (1.12) hold, then  $V_p > \max(V_u, V_c)$  where  $\max(V_u, V_c) = V_u$ . This may be done by manipulating inequality (1.12). Specifically, inequality (1.12) implies that  $[1 - \pi(d) + \kappa^2]M_c < [1 - \pi(d) - \kappa^2]M_u$ , and rearranging the latter expression gives

$$(M_u + M_c)\kappa^2 < [1 - \pi(d)](M_u - M_c). \quad (1.13)$$

Multiplying both sides of inequality (1.13) by  $\pi(d)(M_u - M_c)$  and simplifying the result leads to  $\kappa^2\pi(d)(M_u^2 - M_c^2) < \pi(d)[1 - \pi(d)](M_u - M_c)^2$ , which may be rearranged to obtain

$$\kappa^2\pi(d)M_u^2 < \kappa^2\pi(d)M_c^2 + \pi(d)[1 - \pi(d)](M_u - M_c)^2. \quad (1.14)$$

Then adding  $[1 - \pi(d)]\kappa^2M_u^2$  to both sides of inequality (1.14) and combining terms yields

$$\kappa^2M_u^2 < \pi(d)\kappa^2M_c^2 + [1 - \pi(d)]\kappa^2M_u^2 + \pi(d)[1 - \pi(d)](M_u - M_c)^2. \quad (1.15)$$

Because Eq. (1.11) stipulates that  $\kappa^2 M_u^2 = V_u$  and  $\kappa^2 M_c^2 = V_c$ , the left side of inequality (1.15) is the same as  $V_u$ , and, given Eq. (1.10), the right side is the same as  $V_p$ . Inequality (1.15) therefore implies  $V_p > V_u = \max(V_u, V_c)$ , which is the claimed result.

## APPENDIX 2

### Derivation of Reaction-Time Distributions from a Prototypical Cascade Model

Our derivation of the probability-density functions for reaction times predicted by a prototypical cascade model extends Ashby's (1982) analysis. We proceed by assuming that a prime stimulus is presented during a priming interval of duration  $d$ , and that the prime is then replaced with a test stimulus to which a response must be made. Reaction time  $t$  is measured from the onset of the test stimulus until a response threshold is crossed.<sup>32</sup>

Given the preceding assumptions, the response activation  $A_s(T|d)$  at any time  $T$  after the onset of the test stimulus may be expressed under the model by the following set of equations (cf. Ashby, 1982; McClelland, 1979):

$$A_s(T|d) = b_p + a_p E_p(d) + a_s E_s(T), \quad T \geq 0 \quad (2.1)$$

$$E_p(d) = 1 - \sum_{i=1}^m K_{pi} \exp(-r_{pi}d), \quad d \geq 0 \quad (2.2)$$

$$K_{pi} = \prod_{j \neq i}^m \frac{r_{pj}}{r_{pj} - r_{pi}}, \quad (2.3)$$

$$E_s(T) = 1 - \sum_{i=1}^n K_{si} \exp(-r_{si}T), \quad T \geq 0 \quad (2.4)$$

$$K_{si} = \prod_{j \neq i}^n \frac{r_{sj}}{r_{sj} - r_{si}} \quad (2.5)$$

These expressions generalize Eqs. (7) and (8) of the text, which are simplified approximations to the response-activation functions. The integers  $m$  and  $n$  denote the numbers of cascaded processes for dealing with the prime stimulus and the test stimulus, respectively. The terms  $r_{pi}$  ( $i = 1, 2,$

<sup>32</sup> For simplicity, the present derivation treats response-execution time after the threshold crossing as zero (cf. McClelland, 1979). Our results can be easily generalized, however, to cases where the execution time equals any arbitrary positive constant.

$\dots, m)$  and  $r_{si}$  ( $i = 1, 2, \dots, n$ ) are fixed rate parameters used in processing the prime stimulus and the test stimulus. Correspondingly,  $K_{pi}$  and  $K_{si}$  are constant weighting parameters based on the relative magnitudes of the rates associated with the different processes in cascade. The terms  $b_p$ ,  $a_p$ , and  $a_s$  represent the base-level activation, asymptotic activation induced by the prime stimulus, and asymptotic activation induced by the test stimulus. It is assumed that they have normal distributions with respective means of 0,  $\mu_{ap}$ , and  $\mu_{as}$ , and variances of  $\sigma_{bp}^2$ ,  $\sigma_{ap}^2$ , and  $\sigma_{as}^2$ . Thus, response activation  $A_s(T|d)$  is also a normally distributed random variable with mean  $\mu_{ap}E_p(d) + \mu_{as}E_s(T)$  and variance  $\sigma_{bp}^2 + \sigma_{ap}^2[E_p(d)]^2 + \sigma_{as}^2[E_s(T)]^2$ .

Now, conditional on the activation function exceeding the response threshold (Ashby, 1982), the cumulative distribution function of reaction times can be expressed as follows:

$$F(T|d) = P(t \leq T \mid d \text{ and } t < \infty) = P[A_s(T|d) \geq R \mid A_s(\infty|d) \geq R], \quad (2.6)$$

where  $t$  is reaction time. Furthermore, because  $A_s(T|d)$  is a normal random variable, we can rewrite Eq. (2.6) as

$$F(T|d) = \frac{1 - \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{W(T|d)} \exp(-x^2/2) dx}{1 - Z(Y)}, \quad (2.7)$$

where

$$W(T|d) = \frac{R - \mu_{ap}E_p(d) - \mu_{as}E_s(T)}{\{\sigma_{bp}^2 + \sigma_{ap}^2[E_p(d)]^2 + \sigma_{as}^2[E_s(T)]^2\}^{1/2}},$$

$$Y = \frac{R - \mu_{ap}E_p(d) - \mu_{as}}{\{\sigma_{bp}^2 + \sigma_{ap}^2[E_p(d)]^2 + \sigma_{as}^2\}^{1/2}},$$

and  $Z(Y)$  is the standard normal integral evaluated at  $Y$ . The quantity  $[1 - Z(Y)]$  equals  $P(t < \infty)$ .

The probability-density function associated with  $F(T|d)$  is obtained directly by differentiating the left side of Eq. (2.7) with respect to  $T$  and evaluating the result for  $T = t$ . This yields

$$f(t|d) = \frac{V \left[ \sum_{i=1}^n K_{si} r_{si} \exp(-r_{si}T) \right] \exp\{-[W(T|d)]^2/2\}}{\sqrt{2\pi d} \{1 - Z(Y)\} [\sigma_{bp}^2 + \sigma_{ap}^2[E_p(d)]^2 + \sigma_{as}^2[E_s(T)]^2]^{3/2}}, \quad (2.8)$$

where

$$V = \mu_{as} \{ \sigma_{bp}^2 + \sigma_{ap}^2[E_p(d)]^2 \} + \sigma_{as}^2 E_s(T) [R + \mu_{ap}E_p(d)].$$

Equation (2.8) was used to generate the three density functions shown

in Fig. 3. For these functions, the mean base-level activation  $\mu_{bp}$  equaled 0.0, and the base-level variance  $\sigma_{bp}^2$  equaled 0.1. The asymptote of the prime-stimulus activation,  $a_p$ , had a mean  $\mu_{ap}$  of 2.0 and variance  $\sigma_{ap}^2$  of 0.2. The chosen values of  $d$  in the unprimed (top panel), partially primed (middle panel), and completely primed (bottom panel) conditions of Fig. 3 correspond to the prime-stimulus activation reaching 0, 40, and 95% of its asymptotic value, respectively [i.e.,  $E_p(d) = 0.0, 0.4, \text{ and } 0.95$ ]. The activation induced by the test stimulus involved four cascaded processes whose rates equaled 4, 16, 24, and 30, respectively. The asymptote of the test-stimulus activation,  $a_s$ , had a mean  $\mu_{as}$  of 10.0 and variance  $\sigma_{as}^2$  of 1.0. The response threshold  $R$  was set at 7.0. The parameters used here were similar to ones selected for illustrative purposes by McClelland (1979) and Ashby (1982) in their studies of the cascade model. Note that with  $d = 0$  and  $\sigma_{bp}^2 = 1.0$ , Eq. (2.8) is identical to the density function derived by Ashby [1982, Eq. (4)].<sup>33</sup> Our analysis extends his work by incorporating an additional contribution due to the presentation of a prime stimulus before the test stimulus.

### APPENDIX 3

#### Details of Staircase Tracking Algorithm

##### *Experiments 1 and 3*

To select the duration of the ideal medium priming interval,  $d_m^*$ , in Experiments 1 and 3, we began with an initial rough estimate,  $d_m$ , chosen on an informal basis (e.g., 200 ms for word primes). This estimate was then adjusted iteratively over successive trials of the partially primed condition. Each reaction time,  $t$ , obtained with the medium priming interval was compared against the estimated cutpoint,  $T_x$ , mentioned in the text. If  $t$  exceeded  $T_x$ , then  $d_m$  was increased by an increment  $\delta$  (e.g., 10 ms). If  $t$  fell below  $T_x$ , then  $d_m$  was decreased by  $\delta$ . In effect, the slower reaction times were treated as indicating too little priming, and the faster reaction times were treated as indicating too much priming. The medium interval was lengthened or shortened accordingly to produce an intermediate outcome. Eventually, the adjustment converged on the ideal,  $d_m^*$ , which had the property that there was a .5 probability of the partially primed reaction times exceeding the cutpoint,  $T_x$  (Levitt, 1971).

The rate at which this algorithm converges to the ideal medium priming interval depends on the relation between the probability-density functions of the unprimed and completely primed reaction times. Convergence is

<sup>33</sup> In part of his analysis, Ashby (1982) claimed that the cascade model predicts exceedingly large coefficients of variation (e.g., as large as .4) in reaction time. This claim stemmed from setting the base-level variance  $\sigma_{bp}^2$  to unity. With smaller values of  $\sigma_{bp}^2$ , the model predicts coefficients of variation more like what are typically observed in reaction-time experiments (e.g., .1 to .2; cf. Fig. 3).

fastest and most accurate when  $f_u(t)$  and  $f_c(t)$  do not overlap with each other, allowing the psychometric priming function (Fig. 5) to range from zero to one. Disjoint density functions permit observed reaction times to be classified unambiguously as coming from one distribution or the other; in signal-detection terms (Green & Swets, 1966), no false alarms or misses occur then. To the extent that these density functions do overlap, the range of  $\Psi(d)$  is compressed, requiring more reaction-time observations for the tracking algorithm to reach the ideal medium interval. It is therefore desirable to have an overall priming effect as large as possible.

### *Experiment 2*

We modified the tracking algorithm for Experiment 2, where there were two partially primed conditions, with medium<sub>1</sub> and medium<sub>2</sub> priming intervals, respectively. To track the ideal medium<sub>1</sub> priming interval, an estimated duration ( $d_{m1}$ ) was initialized at 200 ms. Subsequently, the estimate was decreased by a decrement  $\delta$  whenever a trial with a medium<sub>1</sub> interval produced a reaction time less than the cutpoint,  $T_x$ , which was defined as before. The estimate was increased by an increment  $\delta$  whenever two successive trials with medium<sub>1</sub> intervals each produced reaction times greater than  $T_x$ . We calculated the cutpoint  $T_x$  as before [Eq. (9)]. The increments and decrements ( $\pm\delta$ ) equaled 10 ms each. This rule, which had a more stringent requirement for increments than for decrements, targeted the medium<sub>1</sub> interval on a shorter ideal duration,  $d_{m1}^*$ , than did the one-up and one-down rule of Experiments 1 and 3. In particular, whenever  $1 - F_u(T_x) > .71$ ,  $d_{m1}^*$  had the property that there was a .71 probability of the reaction times in the partially primed<sub>1</sub> condition exceeding  $T_x$  (Levitt, 1971). By contrast, Experiment 1's rule yielded a corresponding probability of .5.

To track the ideal medium<sub>2</sub> priming interval, we used the complement of the rule chosen for adjusting the medium<sub>1</sub> interval. The estimated duration ( $d_{m2}$ ) was initialized at 350 ms. Whenever a trial with a medium<sub>2</sub> interval produced a reaction time greater than the cutpoint,  $T_x$ , the estimate was increased by an increment  $\delta$ . It was decreased by a decrement  $\delta$  whenever two successive trials with medium<sub>2</sub> intervals both produced reaction times less than  $T_x$ . This rule, which had a less stringent requirement for increments than for decrements, targeted the medium<sub>2</sub> interval on a longer ideal duration,  $d_{m2}^*$ , than did the rule of Experiments 1 and 3. In particular, whenever  $1 - F_c(T_x) < .29$ ,  $d_{m2}^*$  had the property that there was a .29 probability of the reaction times in the partially primed<sub>2</sub> condition exceeding  $T_x$  (Levitt, 1971).

### *Relation to Two-State Discrete Stage Model*

The performance of the staircase tracking algorithm is closely related

to the two-state discrete stage model. If the model holds, then in the limit, the tracking algorithm must satisfy the following equation:

$$F_p(T_x|d_m^*) = \pi(d_m^*)F_c(T_x|d_m^*) + [1 - \pi(d_m^*)]F_u(T_x|d_m^*), \quad (3.1)$$

where  $T_x$  is the chosen temporal cutpoint, and  $F_p(T|d_m^*)$  is the cumulative distribution function of reaction times generated in the partially primed condition associated with the targeted ideal medium priming interval. This stems directly from Eq. (1) of the text involving the simple mixture prediction. Rearranging these terms, we have an expression for the mixture parameter:

$$\pi(d_m^*) = \frac{F_p(T_x|d_m^*) - F_u(T_x)}{F_c(T_x) - F_u(T_x)} \quad (3.2)$$

Consequently,  $\pi(d_m^*)$  may range anywhere from zero to one, depending on the amount of difference between  $F_p(T_x|d_m^*)$  and  $F_u(T_x)$  relative to the amount of difference between  $F_c(T_x)$  and  $F_u(T_x)$ . If the probability-density functions of the unprimed and completely primed distributions do not overlap at all, then  $\pi(d_m^*)$  would equal  $F_p(T_x|d_m^*)$ .

In Experiments 1 and 3, for example, the algorithm was implemented such that  $F_u(T_x) = 1 - F_c(T_x)$ , and  $F_p(T_x|d_m^*) = .5$ . Substituting these last equalities back into Eq. (3.2) yields a value of  $\pi(d_m^*)$  equal to a half. The algorithm was therefore set to produce a 50–50 mixture of the unprimed and completely primed reaction times in the partially primed condition, assuming that the two-state discrete stage model held.

In Experiment 2, the tracking algorithm was implemented such that  $F_{p1}(T_x|d_{m1}^*) = .29$  for the partially primed<sub>1</sub> condition, and  $F_{p2}(T_x|d_{m2}^*) = .71$  for the partially primed<sub>2</sub> condition. The mixture parameter expected under the medium<sub>1</sub> interval was therefore .29 or less, and the mixture parameter expected under the medium<sub>2</sub> interval was .71 or greater.

The up-and-down rule of the tracking algorithm can be changed to achieve other values of the mixture parameter as well. For example, suppose that we seek an ideal medium priming interval,  $d_m^*$ , for which the mixture parameter  $\pi(d_m^*)$  equals a specified probability  $p$ . Suppose further that we increase our estimate of  $d_m^*$  by an increment  $\rho\delta$  whenever the medium priming interval yields a reaction time greater than  $T_x$ , and that we decrease our estimate of  $d_m^*$  by a decrement  $(1 - \rho)\delta$  otherwise, where the parameter  $\rho$  equals  $pF_c(T_x|d_m^*) + (1 - p)F_u(T_x|d_m^*)$ , and  $T_x$  is the same cutpoint as before [Eq. (9)]. Then the tracking algorithm would converge directly on the desired ideal.

Moreover, the reaction-time cutpoint  $T_x$  used in the algorithm has additional meaning under the two-state model. If the model is valid, and if the unprimed and completely primed distributions of reaction times have symmetric probability-density functions with equivalent shapes [i.e.,  $f_u(t$

+  $\epsilon$ ) =  $f_c(t)$ ,  $t \geq 0$ , where  $\epsilon$  is a positive constant of translation], then  $T_x$  is the temporal value that satisfies Falmagne's (1968) fixed-point property [i.e.,  $f_p(T_x|d) = f_u(T_x) = f_c(T_x)$ ].

## REFERENCES

- Antos, S. (1979). Processing facilitation in a lexical-decision task. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 527-545.
- Ashby, F. G. (1982). Deriving exact predictions from the cascade model. *Psychological Review*, 89, 599-607.
- Bertelson, P. (1967). The time course of preparation. *Quarterly Journal of Experimental Psychology*, 19, 272-279.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon.
- Broadbent, D. E. (1971). *Decision and stress*. New York/London: Academic Press.
- de Groot, A. M. B., Thomassen, A. J. W. M., & Hudson, P. T. W. (1982). Associative facilitation of word recognition as measured from a neutral prime. *Memory & Cognition*, 10, 358-370.
- Dickman, S. J. (in press). Impulsivity and perception: Individual differences in the processing of local and global dimensions of stimuli. *Journal of Personality and Social Psychology*.
- Donders, F. C. (1969). Over de snelheid van psychische processen: Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool, 1868-1869, Tweede reeks, II, 92-100 (W. G. Koster, Trans.). In W. G. Koster (Ed.), *Attention and performance* (Vol. 2, pp. 412-431). Amsterdam: North-Holland.
- Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction time, and human information processing. *Journal of Mathematical Psychology*, 2, 312-329.
- Eriksen, C. W., & Eriksen, B. A. (1972). Visual backward masking as measured by voice reaction time. *Perception & Psychophysics*, 12, 5-8.
- Eriksen, C. W., & Schultz, D. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, 25, 249-263.
- Everitt, B. S., & Hand, B. J. (1981). *Finite mixture distributions*. New York: Chapman & Hall.
- 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. (1968). Note on a simple fixed-point property of binary mixtures. *British Journal of Statistical and Mathematical Psychology*, 21, 131-132.
- Fischler, I., & Bloom, P. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior*, 18, 1-20.
- Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46, 199-210.
- Goldberg, J. H., Meyer, D. E., Yantis, S., & Langolf, G. D. (1985). Temporal processing during the mental rotation of disoriented letters. In A. Mital (Ed.), *Advances in ergonomics/human factors* (Vol. 1). Amsterdam: North-Holland.
- Gopher, D., & Sanders, A. F. (1984). S-Oh-R: Oh resources! oh resources! In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes*. New York: Springer-Verlag.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Hayes-Roth, B. (1977). Evolution of cognitive structures and processes. *Psychological Review*, 84, 260-278.

- Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin*, *96*, 29–44.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441–480.
- Kosslyn, S. M., Pinker, S., Smith, G. E., & Schwartz, S. P. (1979). On the demystification of mental imagery. *Behavioral and Brain Sciences*, *2*, 535–581.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown Univ. Press.
- LaBerge, D., Van Gelder, P., & Yellott, J. (1970). A cueing technique in choice reaction time. *Perception & Psychophysics*, *7*, 57–62.
- Leonard, J. A. (1958). Partial advance information in a choice reaction task. *British Journal of Psychology*, *49*, 89–96.
- Leonard, J. A. (1959). Tactual choice reactions. *Quarterly Journal of Experimental Psychology*, *11*, 76–83.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*, 467–477.
- Link, S. W. (1975). The relative judgment theory of two-choice reaction time. *Journal of Mathematical Psychology*, *12*, 114–136.
- Marble, H. C. (1960). *The hand: A manual and atlas for the general surgeon*. Philadelphia: Saunders.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*, 287–330.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.
- Meyer, D. E., & Irwin, D. E. (Nov., 1981). *On the time-course of rapid information processing*. Paper presented at the meeting of the Psychonomic Society, Philadelphia. Available as Tech. Rep. No. 43, Cognitive Science Program, University of Michigan, June, 1982.
- Meyer, D. E., Osman, A., & Yantis, S. (1982, November). *RT mixture distributions: Evidence for finite-state models of response preparation*. Paper presented at the meeting of the Psychonomic Society, Minneapolis.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.
- Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes. *Science (Washington, D.C.)*, *192*, 27–33.
- Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1972, November). *Activation of lexical memory*. Paper presented at the meeting of the Psychonomic Society, St. Louis.
- Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance* (Vol. 5, pp. 98–118). New York/London: Academic Press.
- Meyer, D. E., Yantis, S., Osman, A., & Smith, J. E. K. (1984). Discrete versus continuous models of response preparation: A reaction-time analysis. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 69–94). Hillsdale, NJ: Erlbaum.
- Miller, J. (1982). Discrete vs. continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 273–296.
- Miller, J. (1983). Can response preparation begin before stimulus recognition finishes? *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 161–182.



- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*, 165–178.
- Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence of facilitatory and inhibitory processes. *Memory & Cognition*, *4*, 648–654.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.
- Ollman, R. (1966). Fast guesses in choice reaction time. *Psychonomic Science*, *6*, 155–156.
- Ollman, R. (1977). Choice reaction time and the problem of distinguishing task effects from strategy effects. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 99–113). Hillsdale, NJ: Erlbaum.
- Ollman, R. T., & Billington, M. J. (1972). The deadline model for simple reaction times. *Cognitive Psychology*, *3*, 311–336.
- Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). New York: Halstead.
- Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, *78*, 391–408.
- Posner, M. I., & Synder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance* (Vol. 5). New York/London: Academic Press.
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*, 160–174.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, *109*, 444–474.
- Rosenbaum, D. A., & Kornblum, S. (1982). A priming method for investigating the selection of motor responses. *Acta Psychologica*, *51*, 223–243.
- Sabol, M. A., & DeRosa, D. V. (1976). Semantic encoding of isolated words. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 58–68.
- Schmitt, J. C., & Scheirer, C. J. (1977). Empirical approaches to information processing: Speed–accuracy tradeoff functions or reaction time? *Acta Psychologica*, *41*, 321–325.
- Schuberth, R. E., & Eimas, P. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 27–36.
- Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, *18*, 105–139.
- Schweickert, R. (1983). Latent network theory: Scheduling of processes in sentence verification and the Stroop effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 353–383.
- Shaw, M. L. (1984). Division of attention among spatial locations: A fundamental difference between detection of letters and detection of luminance increments. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention & performance* (Vol. 10, pp. 109–121). Hillsdale, NJ: Erlbaum.
- Shepard, R. N. (1975). Form, formation, and transformation of internal representations. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*, (pp. 87–122). Hillsdale, NJ: Erlbaum.
- Shulman, H. G., & McConkie, A. (1973). S–R compatibility, response discriminability, and response codes in choice reaction time. *Journal of Experimental Psychology*, *98*, 375–378.

- Smith, E. E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, *69*, 77–110.
- Smith, J. E. K., Meyer, D. E., Yantis, S., & Osman, A. (1982, August). *Finite-state models of reaction time: Estimation of latency distributions*. Paper presented at the meeting of the Society for Mathematical Psychology, Princeton.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 276–315.
- Sternberg, S. (1973, November). *Evidence against self-terminating memory search from properties of reaction-time distributions*. Paper presented at the meeting of the Psychonomic Society, St. Louis.
- Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*, 251–260.
- Swensson, R. G. (1972). The elusive tradeoff: Speed versus accuracy in visual discrimination tasks. *Perception & Psychophysics*, *12*, 16–32.
- Theios, J., & Smith, P. G. (1972). Can a two-state model account for two-choice reaction-time data? *Psychological Review*, *79*, 172–177.
- Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Erlbaum.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge Univ. Press.
- Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 523–530.
- 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.
- Venezky, R. L. (1970). *The structure of English orthography*. The Hague: Mouton.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*, 67–85.
- Yantis, S. (1985). *Dynamics of retrieval from lexical memory*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.
- Yantis, S., Osman, A., & Meyer, D. E. (May, 1982). *Temporal properties of high-speed information processing*. Paper presented at the meeting of the Midwestern Psychological Association, Minneapolis.
- Yellott, J. I. (1967). Correction for guessing in choice reaction time. *Psychonomic Science*, *8*, 321–322.
- Yellott, J. I. (1971). Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology*, *8*, 159–199.
- Yellott, J. I., & Hildreth, J. (1969, August). *Stochastic models for perceptual latency measured with a headstart procedure*. Paper presented at the meeting of the Society for Mathematical Psychology, Ann Arbor, MI.

(Accepted May 21, 1985)