Timing and reaction time

3 authors, including:

David A. Rosenbaum
Pennsylvania State University
184 PUBLICATIONS 9,882 CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Motor Planning View project

Project   Physical and Mental Effort View project

Timing and Reaction Time

Marc Grosjean, David A. Rosenbaum, and Catherine Elsinger
Pennsylvania State University

Because reaction time (RT) tasks are generally repetitive and temporally regular, participants may use timing strategies that affect response speed and accuracy. This hypothesis was tested in 3 serial choice RT experiments in which participants were presented with stimuli that sometimes arrived earlier or later than normal. RTs increased and errors decreased when stimuli came earlier than normal, and RTs decreased and errors increased when stimuli came later than normal. The results were consistent with an elaboration of R. Ratcliff's diffusion model (R. Ratcliff, 1978; R. Ratcliff & J. N. Rouder, 1998; R. Ratcliff, T. Van Zandt, & G. McKoon, 1999), supplemented by a hypothesis developed by D. Laming (1979a, 1979b), according to which participants initiate stimulus sampling before the onset of the stimulus at a time governed by an internal timekeeper. The success of this model suggests that timing is used in the service of decision making.

The aim of this research is to bridge two lines of investigation that so far have been largely independent: the study of timing, on the one hand, and the study of decision making and reaction time (RT), on the other. It is well known that people (and animals) find it natural to behave rhythmically and, despite minor variability in timing performance, generate time intervals with alacrity (Rosenbaum & Collyer, 1998). In light of this familiar propensity for rhythmicity, we thought it would be useful to test the hypothesis that because RT tasks are generally repetitive and temporally regular, participants might exploit temporal regularities to achieve desired combinations of speed and accuracy. Were this hypothesis confirmed, it would support the view that timing plays an explicit role in shaping RTs. Said another way, if we found that participants exploit timing in RT situations, the outcome would suggest that RTs are partly controlled with internal timing mechanisms and do not merely emerge as by-products of other information-gathering activities.

Before detailing our hypothesis and the methods we used to test it, we review three classes of observations that lend support to the general view that RTs may be affected by timing. The first class of observation pertains to the importance of endogenous and exogenous rhythms, the second concerns the presence of periodicities in series of RTs, and the third concerns the way temporal anticipations affect RTs.

Evidence for Time-Dependent Responding

Endogenous and Exogenous Rhythms

Biological agents, like physical systems in general, are highly prone to exhibit and then exploit self-sustaining oscillations (Cohen, Rossignol, & Grillner, 1988; Rosenbaum & Collyer, 1998; Turvey, 1990). That there are endogenous biological rhythms is well known. Heart rates, brain waves, and other biological events are highly periodic or quasi-periodic. Moreover, people easily reproduce instructed rhythms, generating required time intervals with considerable ease (see Rosenbaum & Collyer, 1998).

Exogenous rhythms are also pervasive. Transitions from day to night, from one season to the next, and from one musical note to another follow cyclical patterns. The presence of such patterns often induces corresponding changes in behavior, such as foot tapping when listening to music. The tendency to synchronize with external rhythms can play an important role in the perception of temporal structure (Fraisse, 1981).

Periodicities in Series of Reaction Times

Saying that there are endogenous and exogenous rhythms does not imply that cognitive processes exploit them or are exploited by them. Nonetheless, given how important rhythm is in organizing and constraining perceptual-motor behavior, it is reasonable to think that timing may play a role in RT tasks as well.

There is evidence that endogenous periodicities affect RTs. One line of research has concerned the correlational structure of sequences of RTs associated with ultradian rhythms (periods of less than 12 hr). To determine whether ultradian rhythms show up in RTs, Lovett Doust, Payne, and Podnieks (1978) asked participants to perform a simple RT task for 30 min during which an imperative stimulus was delivered every 60 s. An autocorrelogram of the sequence of RTs yielded a sinusoidal function, consistent with a cycle whose period was considerably less than 12 hr.

Whereas the ultradian cycles studied by Lovett Doust et al. (1978) were relatively simple, other, more complex patterns also appear in series of RTs. Gilden, Thornton, and Mallon (1995) and Gilden (1997) focused on so-called 1/f (or pink) noise. A time series is pink when the slope of its power spectral density function (using log–log coordinates) is close to −1. Pink noise characterizes the temporal unfolding of many physical phenomena and signifies long-range correlations in sequences of observations. Gilden provided evidence for pink noise in series of choice RTs. By contrast, Gilden et al. failed to find evidence for pink noise in series of simple RTs. In the choice RT task that Gilden used, stimulus presentation times were predictable, whereas in the simple RT task that Gilden et al. used, stimulus presentation times were unpredictable. Gilden et al. took the presence of pink noise in the choice RT task and the absence of pink noise in the simple RT task to indicate that 1/f noise reflects the involvement of decision processes. However, the fact that the stimulus presentation times were predictable in one task but not in the other leaves open the possibility that the presence or absence of long-range correlations in RT sequences, captured by pink noise, depends on temporal regularity per se. This possibility speaks again to the importance of temporal regularity in RT performance.

A final line of research that bears on the possible importance of underlying periodicities in RTs comes from studies seeking evidence for time-locked gating of information transmission. Early efforts to provide such evidence focused on the relationship between RTs and the frequency of physiological events, as captured by electroencephalographic (e.g., Surwillo, 1975) or electrocardiographic (e.g., Fernandez-Guardiola, Ayala, & Kornhauser, 1968) signals. More recent studies have relied on behavioral data alone. A representative study of the latter type was conducted by Dehaene (1993). His participants performed choice RT tasks in which difficulty (feature vs. conjunction search) and modality (visual vs. auditory) were manipulated. For all tasks, the stimuli were presented at regular intervals. To determine whether the RT distributions contained regularly spaced peaks and troughs, Dehaene high-passed filtered the envelopes of the histograms and submitted them to autocorrelation and fast Fourier transform (FFT) analyses. In all cases, the autocorrelograms showed sinusoidal patterns indicative of periodicity. The FFTs also revealed a dominant frequency component in the 20–100-Hz range. Collapsing across task difficulty and modality, Dehaene found that RT was linearly related to the oscillation period such that the higher the dominant frequency of the histogram, the quicker the response. On the basis of these results, he proposed that RTs are more frequently distributed at regular intervals than at random intervals after stimulus presentation, that information-processing cycles are initiated in phase with stimulus onsets or offsets, and that oscillation periods of RT distributions are systematically related to task difficulty.

Temporal Anticipation in Reaction Time

The studies just mentioned leave the participant "at the mercy" of his or her endogenous oscillators. That is, the studies do not indicate whether participants capitalize on temporal regularities in stimulus presentations to influence their RTs. It has long been recognized, however, that in RT experiments participants are adept at anticipating stimulus onsets.

Consider simple RT experiments. As Woodworth (1938) noted, simple RT distributions are mixtures of latencies for anticipations and detections. The main basis for this claim was that when there are fixed foreperiods between warning signals and reaction signals, participants often produce mean RTs close to 0 ms. Thus, they generate negative as well as positive RTs.

Special procedures have been used to discourage participants from anticipating. These include using catch trials, where the participant must decide between responding and not responding based on the identity or presence of a signal, and using distributed foreperiods, where the delay between the warning signal and reaction signal is unpredictable. Neither method is entirely successful, however. Simple RTs in catch-trial experiments vary with the probability of the catch trials and with the foreperiods used for the signals (Carterette, Friedman, & Cosmides, 1965; Gordon, 1967). Similarly, in experiments without catch trials, simple RTs depend on the distribution and durations of foreperiods (Karlin, 1959; Klemmer, 1957; Nickerson & Burnham, 1969). If participants in simple RT tasks could refrain from anticipating when catch trials were introduced or when foreperiods varied, one would not expect these effects.

Statistical techniques have been developed to distinguish anticipation responses from detection responses in simple RT experiments (Kornblum, 1973; Ollman & Billington, 1972). These methods have led to the conclusion that it is virtually impossible to prevent participants from anticipating in simple RT situations. The nature of this anticipation, according to Ollman and Billington (1972) and to Kornblum (1973), is that participants operate according to a deadline procedure. They respond either when an internal deadline is reached or when a stimulus is detected, whichever comes first. The accuracy of timekeeping (i.e., waiting for the deadline) is assumed to follow Weber's law, so that the variability of estimated durations is proportional to the mean of the durations. This leads to longer and more variable RTs as the expected delay grows (see also Schmidt, 1968, and Snodgrass, 1969).

Next consider choice RT tasks. Choice RTs, like simple RTs, are sensitive to distributions of foreperiods (Alegria & Bertelson, 1970), and participants in choice RT tasks adapt easily to deadlines (e.g., Link, 1971). When participants in choice RT experiments are told to maximize accuracy and ignore speed, their RT distributions are quantitatively rather than qualitatively different from RT distributions obtained when speed as well as accuracy is emphasized. Only the mean and variance increase under the accuracy-only instructions (Link & Tindall, 1971). On the basis of this result, Link and Tindall suggested that timekeeping plays a role in choice RT performance.

Recently, Lupker, Brown, and Colombo (1997) and Lupker, Taylor, and Pexman (1997) proposed that responses in word-naming tasks might be predominantly time based rather than signal or information based. In one experiment, Lupker, Brown, et al. examined naming latencies when word frequency, which was either high or low, was mixed or blocked. Lupker, Brown, et al. found that the difference in mean RTs between high- and low-frequency words was greater in the blocked than in the mixed conditions. They suggested that this outcome can be explained by assuming that participants adopted different time criteria for responding to words in the two frequency classes within the pure blocks but that participants adopted a single, intermediate time criterion for responding to the high- and low-frequency words in

the mixed blocks. This hypothesis makes RT a controlled rather than emergent property of performance (for an alternative interpretation of such effects, see Zevin & Balota, 2000).

There is precedent for this idea. Some RT models include a deadline to prevent infinitely long RTs (e.g., Nickerson, 1969; Swensson, 1972), although others do not because they are guaranteed to terminate processing with less than infinite times (e.g., random walk; see Feller, 1968). Lupker, Brown, et al. (1997) and Lupker, Taylor, et al. (1997) were the first, however, to suggest that deadlines play a role in governing RTs when reliance on stimulus information can provide a basis for responding with short or medium rather than only long latencies.

Laming (1979b) suggested another way that timing could shape choice latencies. He sought to explain why errors are generally faster than correct responses and why correct trials following a recent error usually have longer RTs than correct trials not following a recent error (Rabbitt, 1966). Laming (1979a, 1979b) suggested that participants anticipate the arrival of a stimulus by sampling information from the perceptual display before the stimulus appears. If the sampling continues and behaves as a random walk until a decision boundary is reached, the earlier the sampling begins relative to stimulus onset, the faster and more error prone the response will be. The time when stimulus sampling begins can be strategically controlled by the participant either by letting it start later than in the last trial (e.g., after an error), which reduces the chance of an error but also elevates RT, or by letting it start earlier than in the last trial, which reduces RT but increases the chance of an error. In either case, the onset of stimulus sampling occurs after a time interval has passed after the warning signal in discrete choice RT tasks, or after some landmark event such as the moment of the last response in serial choice RT tasks. The notion that participants trigger stimulus sampling after a delay is similar to the notion, offered by Ollman and Billington (1972) and by Kornblum (1973), that participants in simple RT tasks use an internal deadline procedure. The difference is that in Laming's (1979a, 1979b) account of choice RT performance, participants begin stimulus sampling when the deadline is reached, whereas in Ollman and Billington's and in Kornblum's account of simple RT performance, participants respond when an internal deadline is reached, unless the stimulus comes first and a response is made to it.

Laming (1979a) reviewed two main findings consistent with his premature sampling hypothesis. First, it had been observed that when there are variable foreperiods in choice RT experiments, short foreperiods yield long latencies and few errors, whereas long foreperiods yield short latencies and more errors (Bertelson & Tisseyre, 1968; Kadlac & Theios, as cited in Link, 1975). This outcome accords with the premature sampling hypothesis because one can assume that in these tasks participants start to sample the perceptual display at an intermediate time between the short and long foreperiods, which, for the reasons given above, led to a speed–accuracy trade-off.

The second major finding presented by Laming (1979a) in support of his premature sampling model was progressive lengthening or shortening of choice RTs over series of 2 to 3 trials. These effects can be quantified as positive autocorrelations that decrease exponentially with lag. The outcome fits with the model given the way participants are assumed to "gamble" with their premature sampling onsets (see Laming, 1979a, for details).

Task and Predictions

Although the notions of deadline setting and premature sampling reviewed in the last section are appealing, other factors, such as changes in perceptual or motor processing, could also account for timing effects in choice RT. Thus, the role of timing in decision making requires further investigation.

The experiments reported here were designed to provide a more in-depth investigation of the role of timing in choice RT performance. To pursue the investigation, we used a variant of the serial choice RT procedure. We presented stimuli at a fixed response–stimulus interval (RSI) until a critical trial at which point the stimulus arrived either earlier or later than normal. The question was what would happen on the critical trial. We considered five models, all of which assumed that the periodicity of stimulus arrivals would trigger some form of internal timekeeping that would influence performance. The models and their distinguishing predictions are reviewed in the next sections and are summarized in Table 1.

Steady Response Rate Model

The first model was an extension of Lupker's model (Lupker, Brown, & Colombo, 1997; Lupker, Taylor, & Pexman, 1997). It said that when participants are exposed to steady RSIs, they establish response deadlines. Consequently, participants generate responses at fairly steady rates, or, said another way, they try to maintain constant interresponse intervals (IRIs). The main prediction of this steady response rate model is shown in Figure 1 (also see Table 1). As a result of striving for constant IRIs, a change in RSI from the expected value results in a change of RT that is equal in magnitude but of opposite sign.

The steady response rate model also makes predictions about accuracy. The identity of the response that must be made when a response is due depends on whichever stimulus–response alternative has the most evidence at that time. Thus, when the response must come early relative to the onset of the stimulus (i.e., when the stimulus comes later than expected), less evidence than usual has

Table 1

Model Predictions for Mean Reaction Time (RT) and Mean Error Rate as a Function of Changes in Intertrial Interval (Shortened Versus Lengthened)

	Intertrial interval			
	Shortened		Lengthened	
Model	RT	Error rate	RT	Error rate
Steady response rate	↑ [a]	↓	↓ [a]	↑
Unelaborated diffusion	—	—	—	—
Diffusion with timed perceptual enhancement	↑	↑	↑ or —	↑ or —
Diffusion with timed motor enhancement	↑	—	↑ or —	—
Diffusion with timed decision enhancement	↑ [b]	↓	↓ [b]	↑

Note. ↑ = increase; ↓ = decrease; — = no change.
[a] Change maintained constant interresponse intervals. [b] Change did not maintain constant interresponse intervals.

Time
⟶

IRI

⟶ S ——RT—→ R ——RSI_C—→ S ——RT_C—→ R
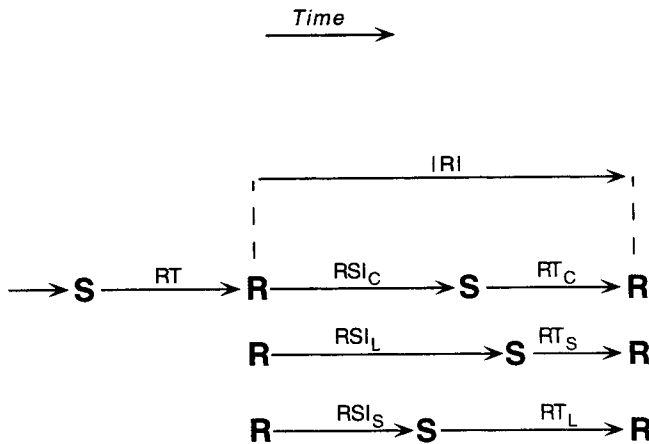
R ——RSI_L—→ S —RT_S→ R

R ——RSI_S→ S ——RT_L—→ R

Figure 1. Sequence of stimuli (S) and responses (R) in a serial reaction time (RT) task and predictions of the steady response rate model. Constant interresponse intervals (IRIs) lead to changes in RT that compensate for changes in response–stimulus interval (RSI). C = constant; L = lengthened; S = shortened.

accumulated, so accuracy is lower than usual. Similarly, when the response must come late relative to the onset of the stimulus (i.e., when the stimulus comes earlier than expected), more evidence than usual has accumulated, and accuracy is higher than usual. Considering accuracy and speed together, this model predicts instantaneous changes in speed–accuracy trade-off when RSIs suddenly change from their previous, expected values.

The steady response rate model also makes predictions about RT distributions. When participants are asked to tap a finger at constant rates, the resulting IRI distributions are approximately symmetric (Collyer & Church, 1998). The steady response rate model predicts, therefore, that RT distributions will be symmetric as well because the model defines each RT as an IRI minus the preceding intertrial interval. (The shape of a distribution is persevered, of course, if it is shifted by subtracting a constant.) In this connection, the steady response rate model also predicts that RT distributions will be laterally shifted but otherwise identical for the three RSI conditions. Previous studies of RT have shown that RT distributions are in fact positively skewed (see, e.g., Luce, 1986; Ratcliff, 1978; Woodworth, 1938). We are unaware of published distributions for serial choice RT tasks with fixed or variable RSIs.

Another prediction of the steady response rate model concerns correlations between generated IRIs. The prediction concerns lag k autocorrelations for responses separated by k intervening responses. An influential model of rhythmic tapping (Wing & Kristofferson, 1973) predicts that lag 1 autocorrelations for IRIs will lie between 0 and $-.5$ and that autocorrelations for higher lags will be 0. These predictions derive from three principal assumptions: (a) Successive responses are controlled with an internal timekeeper that is subject to random variation, (b) successive responses are subject to motor delays that are also subject to random variation, and (c) all delays are stochastically independent. According to the Wing and Kristofferson model, lag 1 autocorrelations will be negative because variation in one motor delay is independent of the next. Thus, if one motor delay happens to be longer than usual, the IRI it completes will also be longer than usual, but the IRI it

starts will be shorter than usual (and vice versa if a motor delay happens to be shorter than usual). Autocorrelations for lags greater than 1 will be zero because of the stochastic-independence assumption. The basis for predicting the range of lag 1 autocorrelations (0 to $-.5$) appears in Wing and Kristofferson and is reviewed in Wing (1980a) and Vorberg and Wing (1996). The predictions about IRI autocorrelations have had considerable influence in the human and animal timing literature. Because the steady response rate model treats responses in RT tasks as rhythmic tapping responses, it predicts the same pattern of autocorrelations for RTs.

Unelaborated Diffusion Model

The second model is one of the most influential models in RT research: the *diffusion* model of Ratcliff (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff, Van Zandt, & McKoon, 1999). The diffusion model is a sequential sampling model of the random walk class (see Luce, 1986). According to the diffusion model (see Figure 2), once a stimulus has been encoded, information accumulates continuously over time from a randomly varying starting point, z, at a drift rate, v, which varies randomly both within and across trials and is proportional to the discriminability of the stimulus. A decision is made when either the upper, a, or lower, 0, information boundary is reached. RT equals the decision time (T) plus the encoding and motor execution times.

The diffusion model successfully accounts for a wide range of phenomena in RT tasks, including latencies for correct and incorrect responses, shapes of RT distributions, and response probabilities. Arguably, it is the most complete theory of choice RT currently available, which makes it an excellent framework in which to evaluate possible timekeeping effects. Note, however, that our hypotheses concerning timing could be implemented in other random walk models as well.
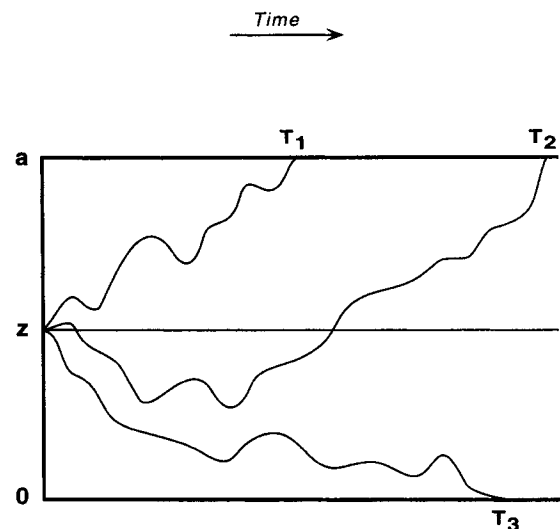
Time
⟶

Figure 2. Unelaborated diffusion model. z = starting point of the diffusion process; a = upper decision boundary; 0 = lower decision boundary; T denotes decision time; Subscripts 1–3 denote three sample paths. Path 1 leads to a fast positive response, Path 2 leads to a slow positive response, and Path 3 leads to a slow negative response.

What does the diffusion model say about sudden changes in intertrial interval? Interestingly, in its unelaborated form it predicts no effects (see Table 1). This is because none of the model's parameters depend on the timing of stimulus presentations. As a result, the model predicts that RTs will be the same regardless of whether stimuli come at regular or irregular intervals. Furthermore, it predicts that RT distributions will be positively skewed and otherwise identical no matter what the RSI. Accuracy will also be unaffected by RSIs.

Diffusion With Timed Perceptual Enhancement

The third model, like the remaining models to be discussed, was an elaborated version of the diffusion model. According to the third model, participants can enhance the gain of perceptual processing at or around the time the stimulus is expected.

As background for understanding the third model, recall that the unelaborated diffusion model asserts that the rate of drift depends on the discriminability of the stimulus. Parametric explorations of the model have shown that the drift rate can be related to stimulus discriminability for stimuli within a class (e.g., a variable number of asterisks that must be classified as "many" or "few"). This suggests that it is meaningful to view the drift rate as analogous to the gain of the perceptual system (Ratcliff & Rouder, 1998).

The contention in our third model is that participants affect this gain through either overt or covert behavior. An overt behavior that can affect gain is rocking forward in one's chair to get a closer look at the computer display as the time of the stimulus arrival draws nigh. A covert behavior that can affect gain is increasing phasic arousal (Posner, 1978). If participants engage in either sort of behavior, the drift rate can increase as the expected stimulus onset time approaches. This can result in a reduction of RT and, as discussed by Ratcliff and Rouder (1998), an increase in accuracy.

The specific predictions of this model for the present experiments are as follows (see Table 1). If the RSI is shorter than normal, RT will be longer than usual, and accuracy will be lower than usual. Recall that this is the opposite of the prediction of the steady response rate model, which predicted longer RTs than usual and higher accuracy than usual for earlier-than-normal stimuli. The prediction of the *timed perceptual enhancement* model for longer-than-normal RSIs depends on one's assumptions about what happens to gain after the expected stimulus arrival time passes. If the drift rate is assumed to decrease after reaching a zenith around the expected stimulus time, RT will be longer, and accuracy will be lower for late stimuli than for on-time stimuli. If the drift rate is assumed to remain at its zenith until the stimulus arrives, RT and accuracy will be the same when the stimulus arrives late and when it arrives on time.

Diffusion With Timed Motor Enhancement

The fourth model is similar to the third in that it concerns processes peripheral to the decision component of the diffusion model. According to this *timed motor enhancement* model, as the expected time of a response approaches, participants prime the motor system without committing to any particular response. Others have proposed that there may be generalized motor preparation of this sort (Alegria, 1975; Bertelson, 1967; Gottsdanker, 1975). If nonspecific motor preparation reaches its zenith at the expected

stimulus time, RTs will be longer than usual if the stimulus arrives early, but accuracy will be unaffected (see Table 1). No other model predicts equal accuracy for early and on-time stimuli. If the stimulus arrives late, the effect will depend on whether one assumes that motor preparation remains at its peak or ebbs. If one assumes that motor preparation remains at its peak, RT will be just as short if the stimulus comes late or if it comes on time, and accuracy will be unaffected. If one assumes that motor preparation ebbs, RT will slow if the stimulus comes late, but accuracy will be unaffected. Neither of these predictions was made by any of the models discussed so far.

Diffusion With Timed Decision Enhancement

The fifth and final model says that stimulus sampling begins prematurely, as proposed by Laming (1979a, 1979b) and discussed more recently by Rouder (1996). With respect to the diffusion model, before the stimulus is encoded, the drift rate, v, is assumed to have a mean of 0, so premature sampling leads to variability in the starting point, z, at the time that stimulus-based diffusion begins. The earlier the start of premature sampling, the wider the distribution of zs, and this, on average, will lead to faster and less accurate responses.

The *timed decision enhancement* model predicts that the earlier the stimulus arrives relative to its expected time, the longer and less error prone responses will be, whereas the later the stimulus comes relative to its expected time, the shorter and more error prone responses will be. As seen in Table 1, these predictions differ from the predictions of all the previous models except for the steady response rate model, which also predicts faster, less accurate responses for late stimuli and slower, more accurate responses for early stimuli. However, the steady response rate model predicts exact compensation in RTs for changes in RSIs, whereas the fifth model allows for less-than-perfect compensation. Because the steady response rate model makes a stronger prediction, it is easier to reject, which raises the question of whether it is just a straw man. It could be viewed as such. The purpose of presenting it was to have at least one time-based model that could serve as a boundary against which other, evidence-based, models could be compared.

General Method

We now report three experiments that were designed to test the predictions of the five models outlined above. In the first experiment, we induced temporal expectancies by using constant RSIs until a final, critical trial when the RSI suddenly changed. In the second and third experiments, we sought to replicate the results of the first experiment and control for potential confounds within it. Of the three experiments to be presented here, the first was the most important. It established the basic effects and helped eliminate all but one of the models. The second and third experiments helped establish the robustness of the findings of Experiment 1.

Participants

Participants were Pennsylvania State University undergraduates (mean age 21.8 years, range 17–56 years) who volunteered in exchange for course credit or $5. Twenty-four participants took part in each experiment. No participant took part in more than one experiment. All participants had normal or corrected-to-normal vision and were naive to the hypotheses.

Apparatus

A Macintosh IIsi microcomputer was used to present stimuli and record responses. Stimuli were displayed on a computer monitor in black on a white background. Responses were registered by keypresses on a Macintosh keyboard. The letters d and k were used for the left and right responses, respectively. These letters have raised "dimples," making it easy to rest the fingers on them without losing one's place on the keyboard while keeping the eyes on the display. All time intervals were computed in Macintosh clock ticks, with a resolution of 60 ticks per second (i.e., 1 tick every 16.7 ms). Although this sampling rate is lower than the conventionally used rate of 1 kHz (i.e., millisecond timing accuracy), it has been demonstrated that biases in measures of RT are negligible for clock intervals of less than 30 ms (Ulrich & Giray, 1989). The software used to run the experiment was written in Pascal.

Procedure

Participants performed a serial choice RT task similar to the classic line-length judgment task of Henmon (1906; cited in Link & Tindall, 1971). On each trial, a vertical and horizontal line appeared at the center of the visual display (see Figure 3). The vertical line was always presented in the same location and was about three times longer than the horizontal line. Together, the lines formed a cross that subtended a visual angle of about $9.34° \times 3.34°$. The horizontal line had two thirds of its length to the left or right of the vertical line. Participants were instructed to press the left key with the left index finger in the former case or the right key with the right index finger in the latter case. The cross remained on the screen until a response was recorded. If the participant accidentally pressed a wrong key (i.e., a letter other than d or k), the block terminated and was immediately repeated. This occurred less than 0.25% of the time across all three experiments.

One reason we adopted the left–right discrimination task was to be able, in future studies, to vary the difficulty of the discrimination by varying how much of the horizontal line lay to one side of the vertical meridian. Another reason was to be able to vary stimulus–response compatibility by varying which response went with which stimulus. Neither of these factors was manipulated here.

The position of the horizontal line varied from trial to trial in a pseudorandom fashion so that the same number of right and left stimuli were

Figure 3. Stimuli presented in the line-length discrimination task and their associated responses. Labeled arrows correspond to visual angles.

presented in each block. Each block consisted of 16 trials. At the end of each block, participants were shown a score equaling the mean RT for the block, multiplied by the quantity (1 + the number of errors). This formula was designed to place somewhat greater emphasis on accuracy than on speed. Participants were instructed to respond as quickly and as accurately as possible and were told to meet this instruction by trying to minimize the end-of-block score. Participants were told nothing about the timing of stimuli.

In all three experiments, we induced a temporal expectancy by maintaining a constant rate of stimulus presentation until a critical switch trial. On this trial, one of three things happened. In the control, or constant, condition, the stimulus arrived at the usual time. In the lengthened experimental condition, the stimulus arrived later than the usual time. In the shortened experimental condition, the stimulus arrived earlier than the usual time. The experimental trials always had changes of $\pm 1/3$ the value of the preswitch interval. We chose the value of $\pm 1/3$ to introduce a change that seemed, based on pilot work, to be as large but also as unobtrusive as possible. We used a ratio change because Weber's law characterizes the perception of time intervals (Getty, 1975; Ivry & Hazeltine, 1995). The switch always occurred between Trials 11 and 16 in all the experiments (for more details, see the *Method* sections that follow).

Each experiment had 61 blocks of trials. The 1st block was always run in the constant condition and was used to familiarize participants with the task. Data from this block were not analyzed. All participants then performed all three conditions in one of the six possible condition orders. Each condition was run for 10 consecutive blocks. This sequence was then repeated, making a total of 20 blocks per condition, or 960 trials for each participant. Counterbalancing was achieved by assigning each condition order to groups of 4 participants (24 participants = 6 condition orders × 4 participants). Each participant's session lasted about half an hour.

Data Analysis

The data were filtered in three ways. First, because of the variability in responses associated with the initiation of a block and our interest in temporal effects, which, presumably, take time to develop, we analyzed the last 10 preswitch trials as well as the switch trial. Second, to eliminate the effects of errors and the typical slowing of RTs in trials after errors (Rabbitt, 1966), we discarded error trials and trials immediately after an error. Finally, we restricted our analyses to responses with RTs between 100 and 1,000 ms.

To evaluate the predictions of the models, we analyzed mean RTs and arcsine-transformed mean error rates with three 2-way repeated measures analyses of variance (ANOVAs). The first ANOVA was restricted to the preswitch trials and allowed us to evaluate the effects of trial number for the last 10 preswitch trials by condition (lengthened, constant, or shortened). The second ANOVA dealt with the transition from the preswitch trials to the switch (critical) trial and allowed us to evaluate the effect of position (mean of the last 10 preswitch trials vs. the switch trial) by condition (lengthened, constant, or shortened). The third ANOVA had the same design as the second ANOVA but was restricted to the critical trial and the trial immediately before it. When either of the latter two ANOVAs yielded significant interactions, post hoc tests were performed to test for the effect of position at each condition level using paired t tests with alpha set at .05.

Because the models we considered made predictions about RT autocorrelations without regard to standard sequential effects (repetitions, alternations, and recovery from errors), we had to do some additional data filtering to compute the autocorrelations. To eliminate the slowing that typifies trials after errors, we restricted the autocorrelation analyses to blocks where no errors occurred. Further, to control for possible trends in the function relating RT to trial number, we detrended the sequence of RTs for each individual block by fitting a least squares straight line to RTs when they were plotted against trial number (for the last 10 preswitch trials).
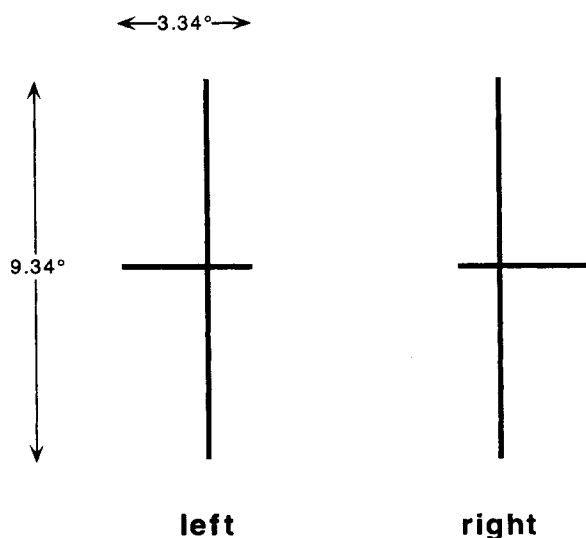
Detrending in this fashion has been used successfully in the rhythmic tapping literature (Vorberg & Wing, 1996) as well as the RT literature (Lovett Doust et al., 1978). We next obtained lag k autocorrelation values for the RT residuals (R) about the best fitting straight line for each block with the following formula: $\mathrm{cov}(R_i, R_{i+k})/\mathrm{cov}(R_i, R_i)$. The highest lag we investigated was 4. Effects of condition and lag (1–4) on autocorrelations were tested using a two-way repeated measures ANOVA. To determine whether the autocorrelations were systematically different from zero, we performed t tests on the mean autocorrelation for each lag taken across conditions within participants.

We also performed analyses to determine whether the results of the autocorrelation analysis could be ascribed to the effect of repeating versus alternating a stimulus–response trial type (Kirby, 1980). (With our design, we could not separate stimulus and response effects.) If such sequential effects existed, one would expect the correlation between successive RTs to differ for repetitions and alternations. We therefore computed correlations between RTs on trial n and trial $n - 1$ separately for repetitions and alternations within each condition, again using detrended sequences for the last 10 preswitch trials for each error-free block. Because of the pseudorandom presentation of stimuli, there was sometimes an insufficient number of repetition or alternation trials (< 3) in a block to compute the correlation. (Note that this is why we could not look at correlations for series of repetition or alternation trials separated by more than one response.) For all obtained mean correlations, we assessed the effects of trial type (repetition, alternation) and condition with a two-way repeated measures ANOVA.

Except for the data discarded because of errors, the number of trials removed did not differ across conditions for any experiment. All averaging and significance tests on correlations and autocorrelations were performed using Fisher's z' transformations, and all t tests were two-tailed.

Experiment 1

We tested the predictions of the five models by manipulating the delay between each response and the subsequent stimulus. We did this by holding the RSI constant until the last trial of each block. On this critical trial, the RSI was either shortened, lengthened, or held constant.

Method

In all three conditions, the last trial of a block had an RSI of 467 ms. In the constant condition, all the RSIs before the 16th trial were also 467 ms; in the lengthened condition, the first 15 trials had RSIs of 350 ms; and in the shortened condition, the first 15 trials had RSIs of 700 ms. Values in this range have been used widely in serial RT experiments (see Luce, 1986). Because the last RSI was always the same (i.e., 467 ms), any condition-based change in response speed and accuracy on the critical trial could be ascribed to the change in RSI per se rather than to its actual value.

Results

Reaction time. RTs were analyzed from the 94.2% of error-free trials and, of these, from the 99.93% of trials whose RTs were neither below the required minimum of 100 ms nor above the allowed maximum of 1,000 ms. Figure 4 shows the mean RTs from the three conditions of Experiment 1. As seen here, RTs increased with trial number until the critical (16th) trial and were slower overall in the lengthened condition than in the constant and shortened conditions. These impressions were confirmed in the first ANOVA, which was restricted to the preswitch trials. This ANOVA yielded a main effect of condition, $F(2, 46) = 6.82$, $p < .01$, and trial number, $F(9, 207) = 2.87$, $p < .001$, but no interaction between condition and trial number, $F(18, 414) = 0.97$, ns.

Figure 5 presents mean RT as a function of condition for the trials before and on the critical trial. As shown in Figure 5, RTs in preswitch trials were similar in the constant and shortened conditions but were slightly longer in the lengthened condition. On the
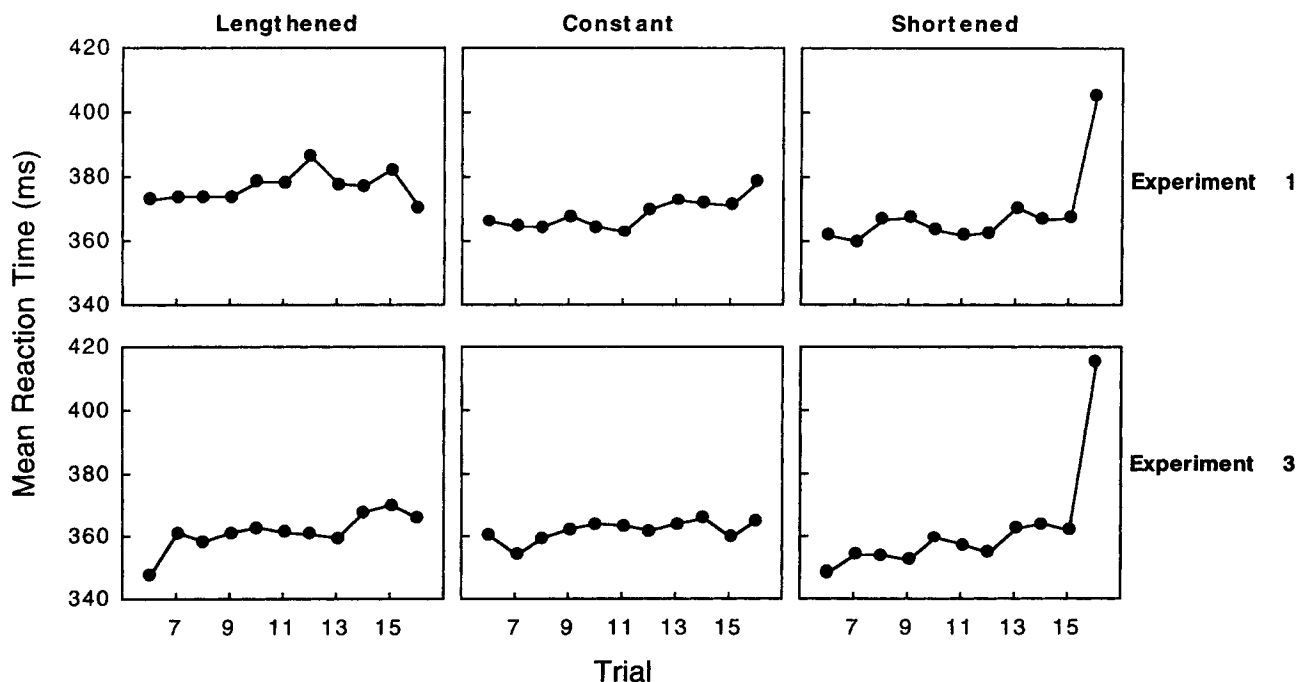

Figure 4. Mean reaction time as a function of trial number (6–16) for each condition (lengthened, constant, or shortened) of Experiments 1 and 3. Trial 16 is the critical trial.
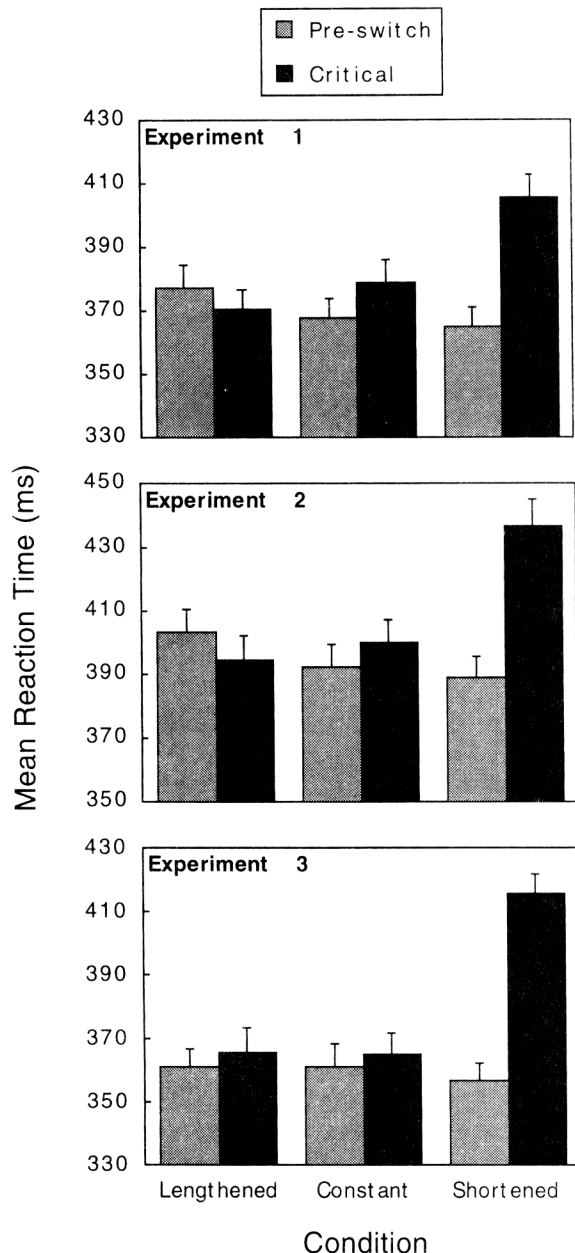
Figure 5. Mean reaction time as a function of condition (lengthened, constant, or shortened) for the trials before the critical trial (pre-switch: Trial 6 to the trial before the critical trial) and for the critical trial (critical) of the three experiments. Error bars represent 1 standard error (among participants).

critical trials, RTs decreased slightly in the lengthened condition, increased slightly in the constant condition, and increased substantially in the shortened condition. These impressions were mainly corroborated by the second ANOVA, which was concerned with the transition from the preswitch to the critical trials. This ANOVA yielded a main effect of position, $F(1, 23) = 50.74, p < .001$, a main effect of condition, $F(2, 46) = 6.27, p < .01$, and a significant Position $\times$ Condition interaction, $F(2, 46) = 29.88, p < .001$. Post hoc tests indicated that mean RT on the critical trial

differed from mean RT in the preswitch trials in the shortened and constant conditions but not in the lengthened condition. Overall, the results were identical in the third ANOVA, where only the 15th trial was compared with the critical trial. The post hoc tests based on this ANOVA revealed that mean RT on the critical trial differed from mean RT on the preceding trial only in the lengthened and shortened conditions.

To further understand the effect of the critical trial manipulation, we considered the RT distributions for the trials preceding and on the critical trial. As shown in Figure 6, the distributions were positively and similarly skewed. The critical trial manipulation had the effect of shifting the distributions along the abscissa so that the critical trial distribution was slightly to the left of the preswitch distribution in the lengthened condition and clearly to the right of the preswitch distribution in the shortened condition.

Error rate. Figure 7 shows mean error rate as a function of trial number for each condition. This figure suggests that there were no systematic changes in error rate before the critical trial. These impressions were confirmed in the first ANOVA, which showed that errors were approximately equally distributed over preswitch trials, with no significant effect of condition, $F(2, 46) = 1.80, ns$; trial number, $F(9, 207) = 1.90, ns$; or interaction between the two, $F(18, 414) = 1.10, ns$.

Figure 8 suggests, by contrast, that error rates differed in the last, critical trial. This impression was confirmed in the second ANOVA, which yielded a significant effect of condition, $F(2, 46) = 10.31, p < .001$, as well as a significant Position $\times$ Condition interaction, $F(2, 46) = 14.59, p < .001$, but no effect of position, $F(1, 23) = 1.48, ns$. Post hoc tests indicated that the critical trials in the lengthened and shortened condition had different error rates from the preswitch trials, whereas the last trial and earlier trials did not differ in the constant condition. The results were identical in the third ANOVA, where only the 15th trial was compared with the critical trial.

Speed–accuracy trade-off. To further assess the impact of the critical trial manipulation, we computed the magnitude of the change in RT and error rate between the critical trial and the trial immediately before it. We then examined how changes in RT were related to changes in error rate in a speed–accuracy trade-off space. The results, displayed in Figure 9, show the difference in mean RT between the critical trial and the immediately preceding trial as a function of the difference in mean error rate between the critical trial and the immediately preceding trial. The vertical and horizontal dashed lines denote no change in error rate or RT, respectively, whereas positive and negative values correspond to increases and decreases on the critical trial, respectively. It is seen in Figure 9 that when the RSI was lengthened, responses were faster and less accurate than when there was no change in RSI (constant condition). Conversely, when the RSI was shortened, responses were slower and more accurate than when there was no change in RSI. There was also a small increase in RT and a drop in error rate for the constant condition. Finally, the pattern was not completely symmetric. The change in RT was larger for the shortened condition than for the lengthened condition.

Sequential dependency. Mean autocorrelations of RT residuals averaged across conditions as a function of lag appear in Figure 10. To compute these autocorrelations, we restricted the analyses to the 65.1% of the blocks that had no errors. As seen in the figure, all values were negative and approached zero with
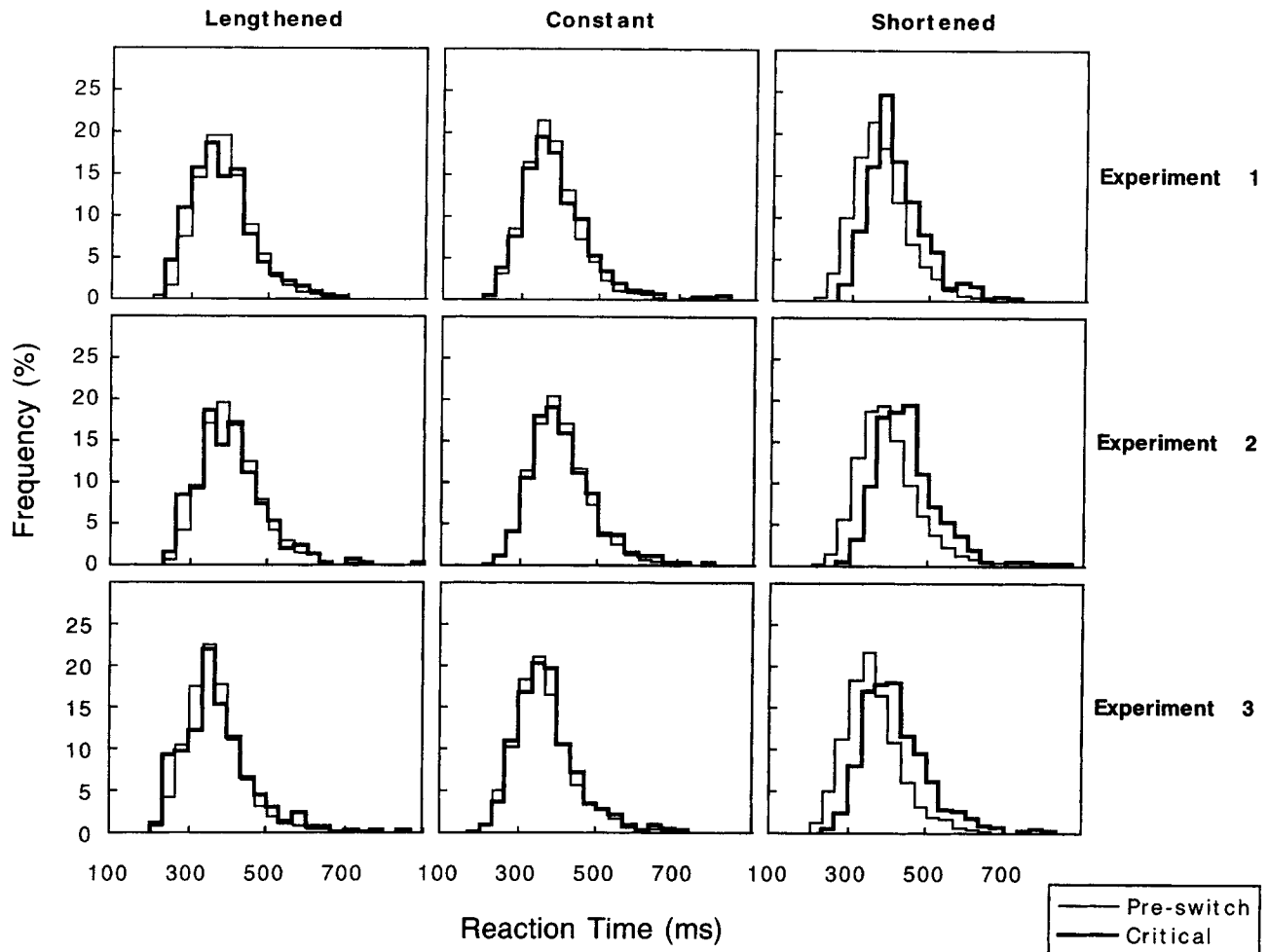
Figure 6. Reaction time distributions for the trials before the critical trial (pre-switch: Trials 6 to the trial before the critical trial) and the critical trial (critical) for each condition (lengthened, constant, or shortened) of the three experiments. The bin width is about 30 ms.

increases in lag. The ANOVA showed no effect of condition on autocorrelation, $F(2, 46) = 0.40$, *ns*; an effect of lag, $F(3, 69) = 45.73$, $p < .001$; and no interaction between condition and lag, $F(6, 138) = 0.45$, *ns*. The individual t tests showed that the mean autocorrelations differed significantly from zero at all lags (all ts < -10, all ps $< .001$).

Table 2 presents mean correlations between successive RT residuals as a function of trial type and condition. To arrive at these values, we restricted the analysis to the 86.98% and the 94.24% of the error-free blocks that had enough (> 2) repetitions and alternations, respectively, to permit reliable estimates. As seen in Table 2, all correlations were negative and lay in the range of $-.35$ to $-.1$. The ANOVA showed no effect of trial type, $F(1, 23) = 0.87$, *ns*; no effect of condition, $F(2, 46) = 1.31$, *ns*; and no interaction between the two, $F(2, 46) = 1.42$, *ns*.

Discussion

Experiment 1 yielded results that were inconsistent with four of the five models. The results contradicted the prediction of the steady response rate model because changes in RTs for early and late stimuli did not fully compensate for the earliness or the lateness of the stimuli. Thus, participants did not maintain constant IRIs. The results also contradicted the unelaborated diffusion model because that model predicted no effect of stimulus-timing changes. RTs and accuracy were in fact affected by the shortening or lengthening of RSIs. The timed perceptual enhancement model could also be rejected because that model predicted that when RSIs were shortened, accuracy would be lower than usual. We found, however, that accuracy improved when RSIs were suddenly shortened. The timed motor enhancement model could also be rejected because it predicted equal accuracy for early and late stimuli. Accuracy was, in fact, higher when stimuli came earlier than normal and was lower when stimuli came later than normal.

The only model that survived Experiment 1 was the timed decision enhancement model. As predicted by this model, RTs were shorter for the late stimuli than for the normally timed stimuli and were longer for the early stimuli than for the normally timed
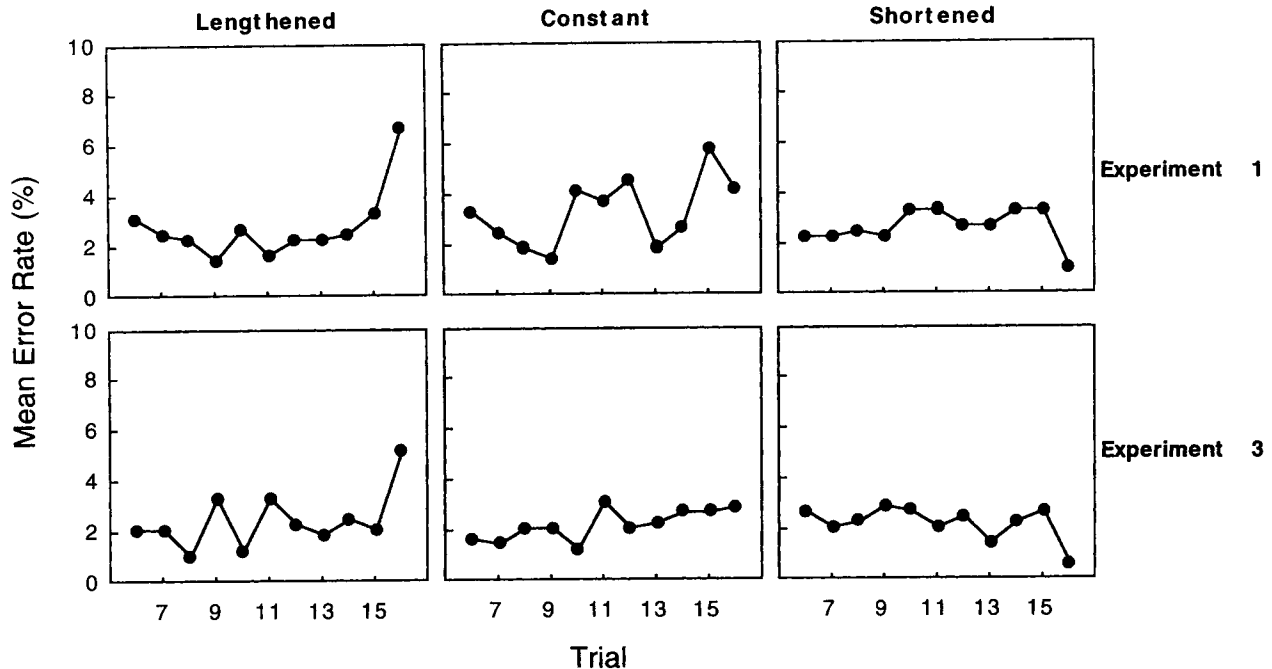
Figure 7. Mean error rate as a function of trial number (6–16) for each condition (lengthened, constant, or shortened) in Experiments 1 and 3. Trial 16 is the critical trial.

stimuli. These RT changes were not as large as the RSI changes that induced them, which the model allowed. Also as predicted by the timed decision enhancement model, accuracy was lower for the late stimuli than for the normally timed stimuli and was higher for the early stimuli than for the normally timed stimuli. RT distributions were positively skewed, as predicted by the diffusion model, of which the timed decision enhancement model is a token. The autocorrelation data were not explicitly predicted by the timed decision enhancement model but could be accommodated by it. One could explain the negative autocorrelations in RTs and the fact that the autocorrelations approached zero with increases in lag by supposing that participants adjusted their stimulus-sampling start times over trials in pursuit of optimal speed–accuracy relations. The fact that RT tended to increase with trial number is consistent with this view (Laming, 1979a). Our attainment of this effect replicates what has been found before in the serial RT literature (Willingham, Greenberg, & Thomas, 1997). In the periodic tapping literature, it has long been recognized that negative autocorrelations for lags greater than 1 may reflect adjustment of responses based on the timing of prior responses (Wing, 1980b). One way such adjustments could have been made in the current experiment would have been to change stimulus-sampling start times. Such corrections could have also yielded negative correlations between successive RTs regardless of whether errors were made or whether they contained repetitions or alternations, as we observed.

Three findings from Experiment 1 require some additional attention. First, although the preswitch versus postswitch pattern held when we compared performance on the critical trial both with mean performance on the prior 10 trials and with performance on the trial immediately before the critical trial, in

the former case the RT effect was somewhat obscured. This was most likely due to the fact that RTs increased with trial number. Second, changes in performance were larger for the shortened condition than for the lengthened condition. This was most likely due to the magnitude of the change in RSI. Because the change was made proportional to the initial RSI, the absolute change in RSI was larger for the shortened condition than for the lengthened condition. This could explain why the shift in speed–accuracy space was larger in the shortened condition than in the lengthened condition. Third, RTs on the preswitch trials were longer in the lengthened condition (when RSIs were short) than in the constant and shortened conditions (when RSIs were longer). This is consistent with the finding that when RSIs are manipulated in a blocked fashion, RT and RSI tend to be inversely related (e.g., Rabbitt, 1969).

Experiment 2

Experiment 2 was designed to replicate the first experiment and also to address the possibility that participants came to expect the change in RSI on the last trial of a block and altered their way of responding accordingly. The latter hypothesis is weakened by the fact that most participants did not report noticing a change in the timing of the last trial. However, we did not have a formal measure of the reliability of such reports, so we felt we needed to evaluate this hypothesis further. To do so, we used the same conditions as Experiment 1, but instead of always having the critical trial on the last trial of a block, the RSI switch occurred randomly in Trials 11–16. The predictions followed the same logic as in Experiment 1.
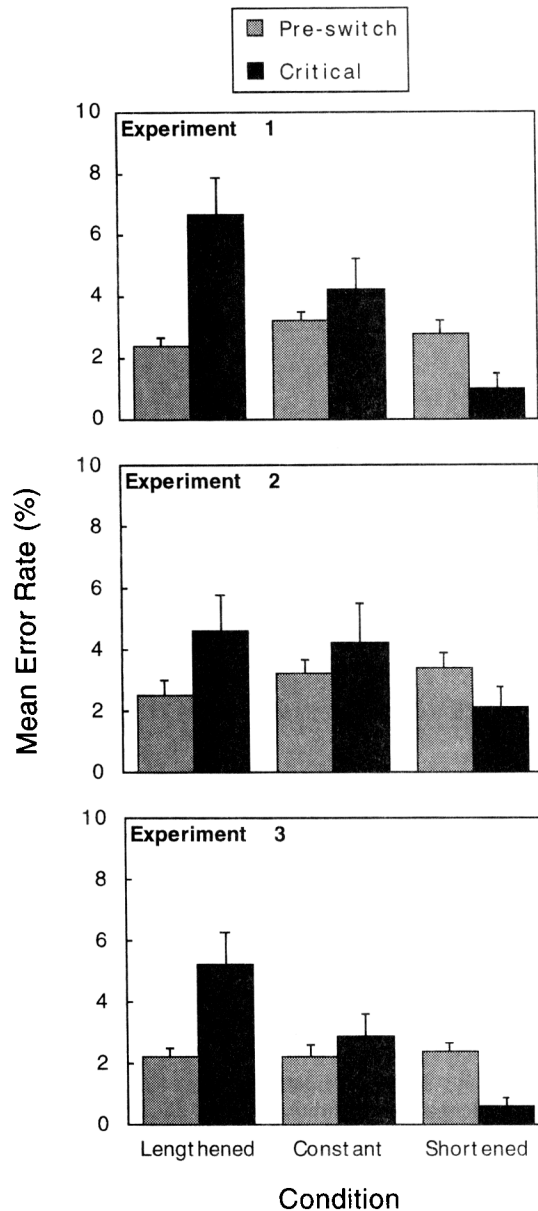
Figure 8. Mean error rate as a function of condition (lengthened, constant, or shortened) for the trials before the critical trial (pre-switch: Trial 6 to the trial before the critical trial) and for the critical trial (critical) of the three experiments. Error bars represent 1 standard error (among participants).

Method

Participants were asked to perform the same task as in Experiment 1. The only difference was that the switch in RSI occurred randomly in Trials 11–16. The data analysis was similar to that of the first experiment, except that because of the randomization of the position of the critical trial, it was impossible to analyze the preswitch trials in the same way. Specifically, the ANOVA designed to evaluate the effect of trial number was not performed for RTs or error rates. However, the two other ANOVAs were performed, although in one the preswitch trials corresponded to the mean of Trial 6 to the trial immediately before the critical trial, and in the other, the preswitch

trials corresponded only to the trial immediately before the critical trial. Because this experiment was mainly designed to control for a confound, we were not directly interested in when the critical trial occurred in the block, so we averaged the critical trials across the six possible switch positions (11–16). The autocorrelation and trial type correlation analyses were performed essentially as before, with the detrending of RTs for each block done for Trial 6 to the trial immediately before the critical trial.
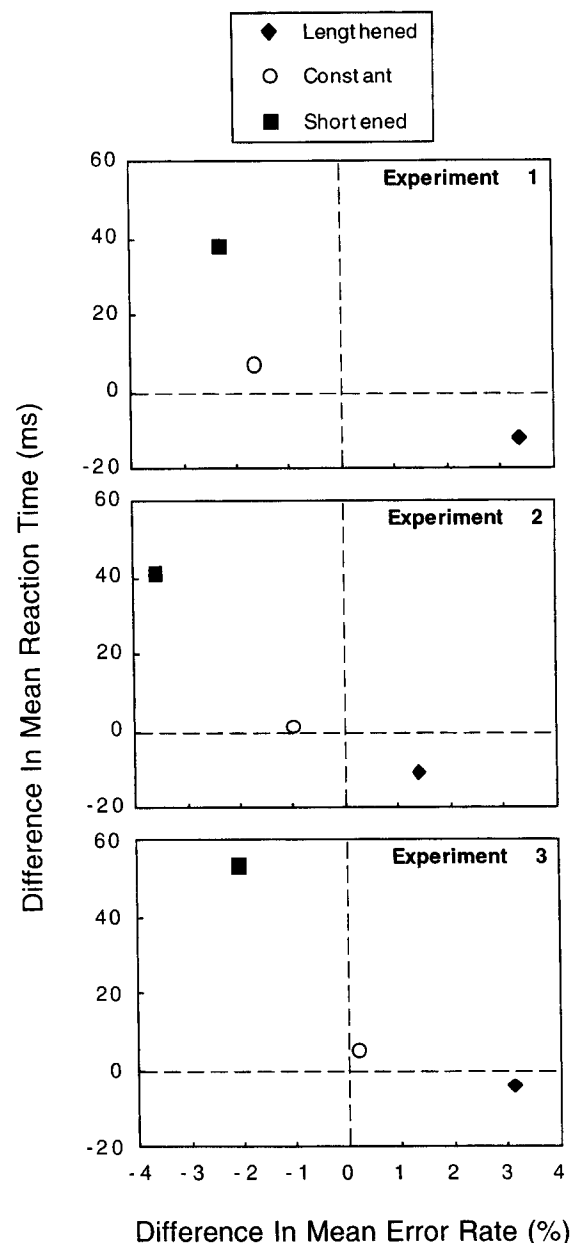
Figure 9. Difference in mean reaction time as a function of difference in mean error rate between the critical trial and immediately preceding trial. Vertical dashed lines represent no change in error rate. Horizontal dashed lines represent no change in reaction time. Positive values correspond to increases in the critical trial. Negative values correspond to decreases in the critical trial. Data are presented for each condition (lengthened, constant, or shortened) for the three experiments.

Figure 10. Mean autocorrelation of reaction time residuals averaged across conditions as a function of lag (1–4) for the three experiments.

Results

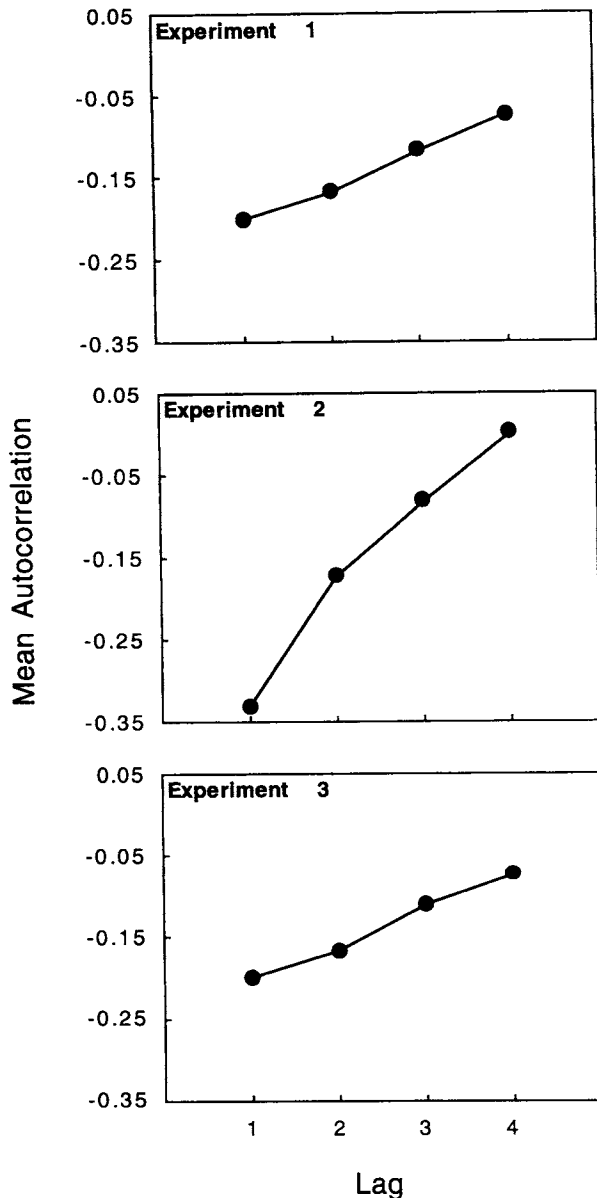*Reaction time.* RTs were analyzed from the 93.8% of error-free trials and, of these, from the 99.82% of trials whose RTs were neither below the required minimum of 100 ms nor above the allowed maximum of 1,000 ms. For the preswitch versus post-switch comparison, mean RT decreased in the lengthened condition, increased somewhat in the constant condition, and increased substantially in the shortened condition (Figure 5). Mean RTs for the preswitch trials were approximately the same over conditions. The ANOVA revealed a main effect of position, $F(1, 23) = 52.11$, $p < .001$, and condition, $F(2, 46) = 5.99$, $p < .01$, and a significant Position × Condition interaction, $F(2, 46) = 44.15$, $p < .001$. Post hoc tests indicated that the lengthened and short-

ened critical trial RT differed from the preswitch trial RTs. Overall, these results were identical in the ANOVA where only the trial immediately before the critical trial was compared with the critical trial, although after this ANOVA, the post hoc test no longer reached significance for the lengthened condition.

As seen in Figure 6, the distributions were all positively and similarly skewed, and the critical trial distribution was shifted to the right of the preswitch distribution for the shortened condition. The two distributions did not differ markedly for the constant and lengthened conditions.

Error rate. The error rates (Figure 8) were approximately equal for the preswitch trials and, as expected, showed the opposite pattern to RT on the critical trial for the lengthened and shortened conditions. The ANOVA showed an effect of position, $F(1, 23) = 8.88$, $p < .01$; no effect of condition, $F(2, 46) = 0.60$, ns; and a Position × Condition interaction that did not quite reach significance, $F(2, 46) = 3.03$, $p = .058$. The same pattern of results was obtained when only the trial immediately before the critical trial was compared with the critical trial, although none of the effects reached significance.

Speed–accuracy trade-off. As seen in Figure 9, when the RSI was shortened, responses became slower and more accurate than in the constant condition. Conversely, when the RSI was lengthened, responses became faster and less accurate than in the constant condition. The constant condition showed a slight drop in error rate but no change in RT. Moreover, the data pattern was asymmetric, with the changes in RT and error rate being larger for the shortened condition than for the lengthened condition.

Sequential dependency. The autocorrelation analysis was restricted to the 70.3% of blocks with no errors. The autocorrelations increased with lag from about $-.35$ at Lag 1 to about 0 at Lag 4 (Figure 10). The corresponding ANOVA showed no effect of condition on autocorrelation, $F(2, 46) = 0.07$, ns; an effect of lag, $F(3, 69) = 142.52$, $p < .001$; and no interaction, $F(6, 138) = 0.57$, ns. Individual t tests showed that the mean autocorrelations differed significantly from zero at all lags (all $ts < -7$, all $ps < .001$, except at Lag 4, $t < .46$, ns).

Table 2

Mean Correlation Between Successive Reaction Time Residuals for Repetition and Alternation Trials for Each Condition (Lengthened, Constant, Shortened) of the Three Experiments

	Trial type	
Condition	Repetition	Alternation
Experiment 1		
Lengthened	−.278	−.217
Constant	−.169	−.217
Shortened	−.302	−.210
M	−.249	−.215
Experiment 2		
Lengthened	−.461	−.447
Constant	−.276	−.423
Shortened	−.316	−.417
M	−.351	−.429
Experiment 3		
Lengthened	−.273	−.206
Constant	−.246	−.192
Shortened	−.264	−.183
M	−.261	−.194

Correlations between successive RT residuals were computed using the 60.51% and the 75.92% of the error-free blocks that had enough (> 2) repetitions and alternations, respectively, to allow for reliable estimates. (Note that the high percentage of blocks discarded for this analysis was due to the reduced number of preceding trials caused by the randomization of the critical trial.) As seen in Table 2, all the correlations occupied the range −.5 to −.25 and were slightly stronger for alternation trials than for repetition trials. However, the ANOVA showed no effect of trial type, $F(1, 23) = 1.59$, ns; no effect of condition, $F(2, 46) = 1.77$, ns; and no interaction between the two, $F(2, 46) = 0.71$, ns.

Discussion

The global pattern of results in the second experiment replicated what we observed in the first experiment. RTs and errors rates changed in the critical trials in a manner consistent with the timed decision enhancement model. Most important, this occurred even though in Experiment 2, in contrast to Experiment 1, the critical trial occurred at an unpredictable position. This outcome allows us to conclude that the pattern of results obtained before did not depend on participants' preparing for a particular critical trial.

Experiment 3

The third experiment was designed both to provide another replication of Experiment 1 and to determine whether the previous results depended on the fact that RSIs differed in preswitch trials but were identical in switch (critical) trials. In Experiment 3, we reversed this situation so RSIs were constant in the preswitch trials and then changed in the switch (critical) trials, at which point the final RSI became longer, shorter, or stayed the same as the RSIs before the switch. If the effects observed in Experiments 1 and 2 depended on having different initial RSIs that converged on a constant final RSI, they would change when the initial RSIs remain the same and the final RSIs varied.

Method

Participants were asked to perform the same serial discrimination RT task as in Experiments 1 and 2, except that RSIs on the preswitch trials were now the same across conditions but varied on the critical trials. To have the same constant condition as before, preswitch trials for all conditions had an RSI of 467 ms. In the constant condition, the critical trial RSI was also 467 ms, but in the lengthened condition, the critical trial RSI was 623 ms, and in the shortened condition, the critical trial RSI was 311 ms. One other change from the previous experiments was that the third experiment was conducted with a Macintosh computer equipped with a data acquisition system that allowed for millisecond timing (Strawberry Tree, Inc.; Bushe, Vaughan, & Rosenbaum, 1994).

Results

Reaction time. RTs were analyzed from the 95.4% of error-free trials and, of these, from the 99.8% of trials whose RTs were neither below the required minimum of 100 ms nor above the allowed maximum of 1,000 ms. As seen in Figure 4, mean RTs increased gradually in preswitch trials but did not differ across conditions. The first ANOVA confirmed that there was a main effect of trial number, $F(9, 207) = 3.85$, $p < .001$; no main effect

of condition, $F(2, 46) = 0.53$, ns; and no interaction between trial number and condition, $F(18, 414) = 0.92$, ns.

RTs were approximately constant over conditions in the preswitch trials. On the critical trial, RTs increased slightly in the lengthened and constant conditions but increased substantially in the shortened condition (Figure 5). The second ANOVA confirmed that there was a main effect of position, $F(1, 23) = 83.66$, $p < .001$, and condition, $F(2, 46) = 8.91$, $p < .001$, and a significant Position × Condition interaction, $F(2, 46) = 36.77$, $p < .001$. Post hoc tests showed that only the shortened critical trial's RTs differed from the preswitch trials' RTs. Overall, the results were identical in the third ANOVA, where only the 15th trial was compared with the critical trial. However, as seen in Figure 9, RT decreased slightly in the critical trial instead of increasing. (For a more detailed consideration of this pattern, see the following *Discussion* section.)

Distributions of RT were positively skewed, and the critical distribution was shifted to the right of the preswitch distribution for the shortened condition (Figure 6). The two distributions did not differ substantially in the constant and lengthened conditions.

Error rate. Errors were approximately equally distributed over preswitch trials, and there was no apparent effect of condition (Figure 7). The first ANOVA revealed no main effect of condition, $F(2, 46) = 0.30$, ns; no main effect of trial number, $F(9, 207) = 1.07$, ns; and no interaction between the two, $F(18, 414) = 0.74$, ns.

As seen in Figure 8, error rates in the three conditions were similar for the preswitch trials, whereas on the critical trial, they increased in the lengthened condition, increased slightly in the constant condition, and decreased in the shortened condition. The second ANOVA showed that there was a nonsignificant effect of position, $F(1, 23) = 2.45$; ns; a significant effect of condition, $F(2, 46) = 8.10$, $p < .01$; and a significant Position × Condition interaction, $F(2, 46) = 13.29$, $p < .001$. Only shortened errors differed from their corresponding preswitch error rates. For the most part, this result was identical in the third ANOVA, where only the 15th trial was compared with the critical trial. Here, the post hoc tests indicated that both the lengthened and shortened critical trial error rate differed from the 15th trial error rate.

Speed–accuracy trade-off. Responses became slower and more accurate in the shortened trials and faster and less accurate in the lengthened trials (Figure 9). However, the magnitude of the change was larger for the shortened trials than for the lengthened trials, and there was no substantial change in either RT or error rate in the constant condition.

Sequential dependency. The analysis of autocorrelations was restricted to the 72.8% of the blocks that were error free. As shown in Figure 10, autocorrelations were always negative and increased to zero with increasing lag. The ANOVA showed no effect of condition on autocorrelation, $F(2, 46) = 1.61$, ns; an effect of lag, $F(3, 69) = 38.39$, $p < .001$; and no interaction, $F(6, 138) = 0.67$, ns. The individual t tests showed that the mean autocorrelations differed significantly from zero at all lags (all $ts < -8$, all $ps < .001$).

Correlations between successive RT residuals were computed using the 86.64% and the 94.94% of the error-free blocks that had enough (> 2) repetitions and alternations, respectively, to allow for reliable estimates. The mean correlations occupied the range of −.3 to −.15 and were slightly stronger for repetition trials than for

alternation trials (Table 2). However, the ANOVA showed no effect of trial type, $F(1, 23) = 1.96$, ns; no effect of condition, $F(2, 46) = 0.12$, ns; and no interaction between the two, $F(2, 46) = 0.03$, ns.

Discussion

The results of Experiment 3 replicate what we found in Experiments 1 and 2. Although some of the analyses did not show significant effects, the global pattern of results was again confirmed.

Two aspects of the results from Experiment 3 deserve more discussion. First, when we compared performance on the critical trial with mean performance on the prior 10 trials, the change in RT appeared to go in the wrong direction for the lengthened condition. However, this was due to the tendency for RT to increase with trial number. When we contrasted the critical trial to the trial immediately preceding it, the change in RT was in the expected direction. Second, the aim of this experiment was to determine whether having identical preswitch RSIs but different switch RSIs would change the pattern of results. The pattern of results remained the same. RTs and errors rates tended to change on the critical trials in accordance with the predictions of the timed decision enhancement model. This outcome allows us to revise one of our previous statements. Earlier, we proposed that the change in RT on the critical trial might have been larger for the shortened condition because the corresponding change in RSI was greater than in the lengthened condition. In the third experiment, however, the magnitude of change in RSI was identical for both experimental conditions, yet the asymmetry was in the same direction. This result indicates that the magnitude of the change in RSI cannot explain the differential effect. An alternative and more appealing explanation is that the smaller shift in RT for the lengthened condition reflected a difference in the functions that characterize how response speed and response accuracy change with respect to the onset of premature stimulus sampling (for details, see Laming, 1979a, 1979b).

General Discussion

The purpose of this study was to bring together the study of timing, on the one hand, and the study of decision making and RT, on the other. The general hypothesis was that because RT tasks are often repetitive and temporally regular, participants may develop timing strategies that affect speed and accuracy. To evaluate the role of such strategies in choice RT performance, we used a serial choice RT task in which RSIs remained constant until a critical trial, when the stimulus came on earlier or later than before.

The results of the three experiments allowed us to conclude that participants' timing strategies affected response speed and accuracy and that sequential dependencies in series of RTs persisted after standard sequential effects (effect of errors and of stimulus–response repetitions vs. alternations) were removed. On the basis of these findings, we rejected four models and accepted one. The models we rejected claimed (a) that participants respond at steady rates, (b) that participants rely on processes akin to those described in the unelaborated diffusion model of Ratcliff and colleagues (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff et al., 1999), (c) that participants adjust the gain of perceptual processes in antici-

pation of stimulus inputs, and (d) that participants adjust the gain of motor output in anticipation of motor output. The third and fourth models were elaborations of Ratcliff's diffusion model.

The model we did not reject was an elaboration of the diffusion model, the timed decision enhancement model, which claimed that participants engage in premature stimulus sampling, as first suggested by Laming (1979a, 1979b). The start of the premature stimulus sampling is governed by an internal timekeeper that meters out delays after preceding stimuli and is adjusted according to the speed and accuracy of participants' recent performance. The data we have presented are consistent with the view that participants engaged in such time-related anticipation based on preceding RSIs.

In its original form, Laming's model (1979a) can account for the speed–accuracy trade-off we observed in our data. However, a feature of our results we did not discuss before but bears mentioning now is whether the trends we observed in the RTs are like those predicted by Laming. Recall that Laming found that series of RTs are characterized by positive autocorrelations that decrease exponentially with lag. In keeping with this result, we found a gradual lengthening of RT with trial number that could be ascribed to gradual delaying of the start of premature stimulus sampling. However, in contrast to Laming, we obtained negative rather than positive autocorrelations, although like Laming we found that the autocorrelations approached zero with increasing lag. The autocorrelations we obtained were consistent with a version of Wing and Kristofferson's (1973) model of rhythmic tapping. According to this version, the start of premature sampling is set by an internal clock that is adjusted on the basis of the timing of prior responses.

A possible reason for the discrepancy between our findings and Laming's (1979a) is that we used detrended series of RTs whereas Laming did not. To check this, we recomputed the autocorrelations without detrending. The outcome was a set of negative values with no systematic effect of lag. This suggests that the removal of potential nonstationarities in the time series did not account for the difference in patterns. Further work is needed to determine what other factors, such as the length of the time series or the way sequential effects were controlled for, might underlie the discrepancy between Laming's results and ours.

Another issue worth mentioning in relation to Laming's work (1979a) concerns the nature of the data on which he based his proposal. Laming focused on his own data for his autocorrelation analysis but referred to data published by other authors when considering speed–accuracy trade-offs. We, in contrast, have developed a new method for studying timing effects and have performed all of our analyses on the same data sets. This difference might also account for some or all of the discrepancies between Laming's findings and ours. Nonetheless, the applicability of Laming's model to our data is, for the most part, impressive. Indeed, the power of his approach can be viewed as that much greater in view of the fact that we have shown that his model stood up better than several other models we considered.

The remainder of this article concerns (a) three other possible accounts of our data, (b) the relation of the timed decision model to other evidence suggesting neither a purely perceptual nor purely motoric basis for timing, (c) the possibility that diffusion might serve as the basis for timing phenomena in general, and (d) the broader implications of our findings for RT researchers.

Other Possible Accounts of the Data

There are at least three accounts of our results other than the ones we considered so far. According to one, participants were simply surprised by the earlier-than-normal stimulus in the shortened condition. This would explain why RT was lengthened in this condition, and it might explain why the RT effect was more robust in the shortened condition than in the lengthened RSI condition. The surprise hypothesis is weakened, however, by the fact that accuracy increased when stimuli came earlier than expected. One would expect surprise only to have an adverse effect on accuracy, not a beneficial effect, so we doubt that surprise alone can explain our results.

A second explanation is that participants were affected differentially by the foreperiods corresponding to the individual RSIs within the critical trials. It is well known that RTs depend on foreperiods (see Luce, 1986). If one treats responses as warning events, RSIs can be functionally equated with foreperiods, in which case one might argue that RTs and error rates in the critical trials simply reflected the foreperiods operating in those trials. This interpretation is weakened by the fact that there was little or no effect of initial RSIs on RTs; the main effect of condition (i.e., of RSI) for the preswitch trials was only significant in Experiment 1. Thus, participants were largely unaffected by the foreperiods when the foreperiods were well established during the initial trials, which makes it unlikely that the change in RT on the last trial was shaped entirely by the foreperiod on just that trial. Furthermore, the foreperiod was actually the same in the critical trials of all three conditions in the first two experiments, yet the RTs in those three conditions differed depending on the context in which the RTs were obtained. Thus, the results cannot be accounted for by saying that the foreperiod in the critical trials determined the corresponding RTs.

The third possible account of our data is one we cannot reject. According to this account, rather than, or in addition to, engaging in premature stimulus sampling, participants began to decrease the distance between the decision boundaries in anticipation of the stimulus. This mechanism has been considered before (Viviani, 1979) and can be considered another time-related elaboration of the diffusion model. If decision boundaries are shifted in time, this strategy would be generally consistent with our claims about time-related anticipation. Detailed simulation work would be needed to evaluate this shifting-boundary model.

Amodal Timing

The model we have been led to claimed that participants use timing information to affect decision making. Neither a purely perceptual model (the timed perceptual enhancement model) nor a purely motoric model (the timed motor enhancement model) could account for the data. This finding complements a result obtained in the study of timing: Timing is neither purely perceptual nor purely motoric, but instead is amodal.

At least two studies have supported the amodality of timing. In one (Rosenbaum & Patashnik, 1980), participants made a left-index-finger response followed by a right-index-finger response, performing this task with two goals in mind. One was to produce a correct IRI between the two responses. The other was to minimize the simple RT for the first response after a *go* signal.

Feedback after each trial indicated how well the participant satisfied both goals. Mean RTs decreased with subsequent mean IRIs at a higher rate, the larger the variance of the IRIs. In another condition, the same participants pressed the left index finger as quickly as possible after the *go* signal and then judged whether vibration applied to the right index finger came earlier or later than a target interval. The simple RTs in this task were related to the target intervals and to the achieved precision of timing in virtually the same way as were the simple RTs in the interval production task. Thus, the RT results were the same when timing was to be expressed in interval production or interval judgment.

The other study that supported the amodality of timing was Ivry and Hazeltine's (1995) comparison of perception and production of time intervals. These authors asked participants either to judge or to reproduce single intervals defined by tone pairs. Ivry and Hazeltine found that Weber's law held in both tasks and that the value of the Weber constant was similar when participants produced or judged times. The similarity of the findings for perception and production in Ivry and Hazeltine's study, like the finding of Rosenbaum and Patashnik (1980), points to an abstract basis for timing and echoes the conclusion of the present work that participants' time-related anticipation is not restricted to changing the gain of perceptual processing (the diffusion with timed perceptual enhancement model) or motor processing (the diffusion with timed motor enhancement model), but instead is more abstract (the diffusion with timed decision enhancement model).

Diffusion-Based Timing

Another issue concerns the possibility that timing might be based on diffusion. As has often been said in the timing literature, there is no organ of time. The sense of time derives from experiences that unfold chronologically. This raises the possibility that the diffusion process could be used to control time intervals. By setting the decision boundary, a, further from the starting point, z, and either keeping the drift rate, v, the same or allowing it to vary, one might attempt to control the time between presentation of a stimulus and the decision to accept it. If the stimulus were one's own finger tapping on a key, one might be able to rely on this method to control the delay between when the finger is tapped and when it is tapped again.

This idea is appealing because it could lead to a unified account of timing and decision making. Nevertheless, the idea is difficult to accept because distributions of RTs are almost always positively skewed (Luce, 1986), whereas distributions of IRIs in tapping tasks are almost always symmetric (Collyer & Church, 1998). Distributions of time intervals are also almost always symmetric in interval-completion tasks where participants are supposed to respond at a target time after a stimulus (Gibbon, Church, & Meck, 1984). The fact that the shapes of the distributions are so reliably different in RT tasks and in timing tasks suggests that timing and decision making are fundamentally different. Yet the present study, like others (Kornblum, 1973; Laming, 1979a, 1979b; Lupker, Brown, & Colombo, 1997; Snodgrass, 1969), suggests nonetheless that timing is used to aid the speed of decision making. This indicates that timing and decision making are controlled with reference to one another even if they rely on somewhat different mechanisms.

The fact that timing and decision making are fundamentally different is somewhat surprising because one could imagine that diffusion (or some other evidence-based system) might be used to control timing. A step in that direction was attempted by Ratcliff (1978) in what he referred to as *time-controlled processing*. In the context of the diffusion model, he discussed the use of temporal boundaries instead of information boundaries to account for RT performance when response deadlines or response signals are used to manipulate speed–accuracy trade-offs. The possibility of time-controlled processing is attractive both because one hopes for a unified theory and because the diffusion model can produce distributions of RTs that can approximate symmetry. However, the inclusion of temporal boundaries changes the system from being information driven to being time driven. Thus, an exciting challenge for the future will be to see whether a true evidence-based system can account both for times that emerge during decision making and for times produced for their own sake.

Implications of This Work for Reaction Time Researchers

What can be gleaned from this research by those who use RT tasks? The immediate message is methodological. In keeping with previous findings, our work shows that RT researchers must pay close attention to the temporal predictability of the stimuli they use. This is because, as we have shown, participants engage in premature stimulus sampling (or timed boundary shifts) when temporal predictability is high. Our work does not imply that researchers must fundamentally revise their interpretations of Stroop, memory scanning, and other effects indexed by RT. On the other hand, our work does imply that effects observed in classical RT contexts may reflect a mixture of evidence-based decision making, on the one hand, and time-related anticipation, on the other. Because evidence-based decision making is, in most cases, the primary focus of researchers' interests, it is important for RT researchers to be mindful of the fact that participants in choice RT studies tend to use time-based anticipation when they can. So far, time-related anticipation has mainly been documented in simple RT tasks, where higher level processes are generally assumed to be less involved. We have shown that even in choice RT tasks, which are generally assumed to tap higher level processes, participants also rely on time-related anticipation. We can therefore recommend that if one wants to minimize the likelihood that time-related anticipation contributes to RTs, one should maximize temporal uncertainty as well as emphasize accuracy over speed.

Another message is that there are rich and interesting connections between timing and RT. Quite possibly, time-related anticipation can never be fully eliminated in RT tasks. A more useful venture than finding ways of minimizing time-related anticipation, therefore, might be to study the extent to which participants rely on anticipatory timing as a function of the cognitive processes they are otherwise using. Does anticipatory timing occur more readily in visual imagery experiments, for example, than in semantic decision experiments? If so, what accounts for these differences? The method introduced here provides a way of addressing this exciting new question.

References

Alegria, J. (1975). Sequential effects of foreperiod duration: Some strategical factors in tasks involving time uncertainty. In P. M. A. Rabbitt &

S. Dornic (Eds.), *Attention and performance V* (pp. 1–10). London: Academic Press.

Alegria, J., & Bertelson, P. (1970). Time uncertainty, number of alternatives, and particular signal response pair as determinants of choice reaction time. In A. F. Sanders (Ed.), *Attention and performance III.* Amsterdam: North-Holland.

Bertelson, P. (1967). The time course of preparation. *Quarterly Journal of Experimental Psychology, 19,* 272–279.

Bertelson, P., & Tisseyre, F. (1968). The time course of preparation with regular and irregular foreperiods. *Quarterly Journal of Experimental Psychology, 20,* 297–300.

Bushe, M. M., Vaughan, J., & Rosenbaum, D. A. (1994). Pascal external functions for Strawberry Tree's "Analog Connection Workbench." *Behavior Research Methods, Instruments, and Computer, 26,* 461–466.

Carterette, E. C., Friedman, M. P., & Cosmides, R. (1965). Reaction-time distributions in the detection of weak signals in noise. *Journal of the Acoustical Society of America, 38,* 531–542.

Cohen, A. H., Rossignol, S., & Grillner, S. (Eds.). (1988). *Neural control of rhythmic movements in vertebrates.* New York: Wiley.

Collyer, C. E., & Church, R. M. (1998). Inter-response intervals in continuation tapping. In D. A. Rosenbaum & C. E. Collyer (Eds.), *Timing of behavior: Neural, computational, and psychological perspectives* (pp. 63–87). Cambridge, MA: MIT Press.

Dehaene, S. (1993). Temporal oscillations in human perception. *Psychological Science, 4,* 264–270.

Feller, W. (1968). *An introduction to probability theory and its applications* (3rd ed.). New York: Wiley.

Fernandez-Guardiola, A., Ayala, F., & Kornhauser, S. (1968). EEG, heart rate, and reaction time in humans: Effect of variable vs. fixed interval repetitive stimuli. *Physiology and Behavior, 3,* 231–240.

Fraisse, P. (1981). Multisensory aspects of rhythm. In R. D. Walk & H. L. Pick (Eds.), *Intersensory perception and sensory integration* (pp. 217–248). New York: Plenum.

Getty, D. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics, 18,* 1–18.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Annals of the New York Academy of Sciences: Vol. 423. Timing and time perception.* (pp. 52–77). New York: New York Academy of Sciences.

Gilden, D. L. (1997). Fluctuations in the time required for elementary decisions. *Psychological Science, 8,* 296–301.

Gilden, D. L., Thornton, T., & Mallon, M. W. (1995). 1/f noise in human cognition. *Science, 267,* 1837–1839.

Gordon, I. E. (1967). Stimulus probability and simple reaction time. *Nature, 215,* 895–896.

Gottsdanker, R. (1975). The attaining and maintaining of preparation. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 33–49). London: Academic Press.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 3–18.

Karlin, L. (1959). Reaction time as a function of foreperiod duration and variability. *Journal of Experimental Psychology, 58,* 185–191.

Kirby, N. H. (1980). Sequential effects in choice reaction time. In A. T. Welford (Ed.), *Reaction times* (pp. 129–172). London: Academic Press.

Klemmer, E. T. (1957). Simple reaction time as a function of time uncertainty. *Journal of Experimental Psychology, 54,* 195–200.

Kornblum, S. (1973). Simple reaction time as a race between signal detection and time estimation: A paradigm and method. *Perception & Psychophysics, 13,* 108–112.

Laming, D. (1979a). Autocorrelation of choice-reaction times. *Acta Psychologica, 43,* 381–412.

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

Link, S. W. (1971). Applying RT deadlines to discrimination reaction time. *Psychonomic Science, 25,* 355–358.

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

Link, S. W., & Tindall, A. D. (1971). Speed and accuracy in comparative judgments of line length. *Perception & Psychophysics, 9,* 284–288.

Lovett Doust, J. W., Payne, W. D., & Podnieks, I. (1978). An ultradian rhythm of reaction time measurements in man. *Neuropsychobiology, 4,* 93–98.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 570–590.

Lupker, S. J., Taylor, T. E., & Pexman, P. M. (1997, November). *Strategic control of a time criterion in naming: New evidence and effects.* Paper presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia.

Nickerson, R. S. (1969). "Same"–"different" response times: A model and a preliminary test. *Acta Psychologica, 30,* 257–275.

Nickerson, R. S., & Burnham, D. W. (1969). Response times with nonaging foreperiods. *Journal of Experimental Psychology, 79,* 452–457.

Ollman, R. T., & Billington, M. J. (1972). The deadline model for simple reaction time. *Cognitive Psychology, 3,* 311–336.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.

Rabbitt, P. M. A. (1966). Errors and error corrections in choice-response tasks. *Journal of Experimental Psychology, 71,* 264–272.

Rabbitt, P. M. A. (1969). Psychological refractory delay and response–stimulus interval duration in serial, choice response tasks. In W. G. Koster (Ed.), *Attention and performance II* (pp. 195–219). Amsterdam: North-Holland.

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

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

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review, 106,* 261–300.

Rosenbaum, D. A., & Collyer, C. E. (Eds.). (1998). *Timing of behavior: Neural, psychological, and computational perspectives.* Cambridge, MA: MIT Press.

Rosenbaum, D. A., & Patashnik, O. (1980). A mental clock-setting process revealed by reaction times. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 487–499). Amsterdam: North-Holland.

Rouder, J. N. (1996). Premature sampling in random walks. *Journal of Mathematical Psychology, 40,* 287–296.

Schmidt, R. A. (1968). Anticipation and timing in human motor performance. *Psychological Bulletin, 70,* 631–646.

Snodgrass, J. G. (1969). Foreperiod effects in simple reaction time: Anticipation or expectancy? *Journal of Experimental Psychology Monographs, 79*(3, Pt. 2).

Surwillo, W. W. (1975). Reaction-time variability, periodicities in reaction time distributions, and the EEG gating signal hypothesis. *Biological Psychology, 3,* 247–261.

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

Turvey, M. T. (1990). Coordination. *American Psychologist, 45,* 938–953.

Ulrich, R., & Giray, M. (1989). Time resolution of clocks: Effects on reaction time measurement—Good news for bad clocks. *British Journal of Mathematical and Statistical Psychology, 42,* 1–12.

Viviani, P. (1979). A diffusion model for discrimination of temporal numerosity. *Journal of Mathematical Psychology, 19,* 108–136.

Vorberg, D., & Wing, A. M. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action* (pp. 181–262). London: Academic Press.

Willingham, D. B., Greenberg, A. R., & Thomas, R. C. (1997). Response-to-stimulus interval does not affect implicit motor sequence learning, but does affect performance. *Memory & Cognition, 25,* 534–542.

Wing, A. M. (1980a). The long and short of timing in response sequences. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 469–486). Amsterdam: North-Holland.

Wing, A. M. (1980b). Timing of movement phases of a repeated response. *Journal of Motor Behavior, 12,* 113–124.

Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics, 14,* 5–12.

Woodworth, R. S. (1938). *Experimental psychology.* New York: Holt.

Zevin, J. D., & Balota, D. A. (2000). Priming and attentional control of lexical and sublexical pathways during naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 121–135.