Journal of Mathematical Psychology **45**, 334–354 (2001)
doi:10.1006/jmps.2000.1319, available online at http://www.idealibrary.com on IDEAL®

Testing Evidence Accrual Models by Manipulating Stimulus Onset

Jeffrey N. Rouder

University of Missouri–Columbia

Many models of response time assume that subjects accrue stimulus "evidence" samples in time (e.g., random walk models, counter models). In this paper, the concept of one stimulus dominating another is used to construct a test of the whole class of evidence accrual models. For an example of dominance, consider stimuli that are presented either virtually instantaneously (stepped) or in a gradually increasing manner (ramped). Ramped stimuli are presented such that the ramped portion precedes the stepped onset of stepped stimuli. In this case ramped stimuli dominate stepped stimuli. In this paper the class of evidence accrual models is formalized. It is shown that under appropriate assumptions evidence accrual models do predict more accurate responses to dominating stimuli. However, this result does not hold for response latencies. There are anomalous cases where an evidence accrual model (the accumulator model of Vickers (1970, *Ergonomics* **13**, 37–58)) predicts slower mean correct response latencies to dominating stimuli. It is shown through extensive computer simulation that these anomalous cases occur only when response criteria are so asymmetric that there are exceedingly extreme response biases. For experiments where response biases are not exceedingly extreme, random walk and accumulator models predict more accurate and quicker correct responses to dominating stimuli. In sum, manipulating the time course of stimuli in accordance with the concept of dominance can provide empirical tests of the class of evidence accrual models. © 2001 Academic Press

This work was supported by National Science Foundation Grant BCS-9817561. I am indebted to James Townsend for a critical review of a different manuscript. In that review, he pointed out the need for a formal treatment of stimulus onset and evidence accrual. In addition, Townsend has made a number of substantive comments that have improved the exposition of the mathematics presented here. Stephen Link, David Meyer, Sotoru Suzuki, Pablo Gomez, and Duncan Luce have provided numerous helpful comments.

Address correspondence and reprint requests to Jeff Rouder, Department of Psychology, 210 McAlester Hall, University of Missouri, Columbia, MO 65211.

0022-2496/01 $35.00

INTRODUCTION

Testing Evidence Accrual Models by Manipulating Stimulus Onset

Consider a discrimination paradigm in which an observer must classify members from a set of stimuli into one of two categories. The stimuli differ on a unidimensional variable, and the value of the stimuli on this variable is referred to as the strength. An example of such a paradigm is luminance discrimination, in which an observer is required to classify the luminance of a target patch as being darker or lighter than a neutral grey background. Considerable effort has gone into proposing and testing models that account for both the accuracy and latency of responses in these paradigms.

Evidence accrual models have often been used to explain two-choice discriminations. These models assume that observers extract evidence samples from stimuli in a sequential manner. Evidence samples are assumed to be noisy and a single sample is not, in general, sufficient to decide in favor of any one of the response alternatives with a reasonable level of confidence. The observer accrues the samples together until the total evidence from the stimulus is sufficient to make a response. A common feature of evidence accrual models is that evidence from several time intervals is accrued, integrated, or summed in some manner. Evidence accrual models have also been referred to as sequential sampling models, e.g., Smith, 1990. The class of evidence accrual models includes random walk models (Link & Heath, 1975), diffusion models (Ratcliff, 1978), accumulator models (LaBerge, 1963; Vickers, 1970), and Ornstein–Uhlenbeck "leaky" diffusion models (Busemeyer and Townsend, 1993; Smith, 1995).

Many researchers have not tested evidence accrual models as a class, but have attempted to show that one member of the class fits certain aspects of the data better than another member (e.g., Heath, 1984; LaBerge, 1994; Laming, 1979; Link, 1992; Pickett, 1967; Ratcliff & Rouder, 1998, Ratcliff, Van Zandt, & McKoon, 1999; Sanders & Ter Linden, 1967; Vickers, Caudry, & Willson, 1971; Vickers & Smith, 1985). In this paper, I outline a method for testing the class as a whole by varying the strength of the stimulus over the course of a trial. In one condition, the strength of the stimulus *abruptly* changes from a neutral value to a value that can be detected. In another condition, the strength *gradually* changes from a neutral value to one that can be detected.

To formalize time-course manipulations, let $s(t)$ be the value of the strength at time t. The function $s(t)$ is termed a *strength function*. Panel A of Fig. 1 shows a strength function which has zero strength between the time of 0 and the time of t_0. The time of zero corresponds to the warning signal, or the definitive start of the trial to the observer. After t_0, the stimulus has positive strength y. The interval between the time of 0 and the time of t_0 is the foreperiod. In the above luminance-discrimination example, the target patch has the same luminance as the background during the foreperiod. Then, after the foreperiod, the luminance of the target patch is instantly raised I_0 candelas per meter squared (cd/m^2) above that of the background. Panel B of Fig. 1 corresponds to the case in which the luminance of the target is lowered I_0 cd/m^2 below that of the background.

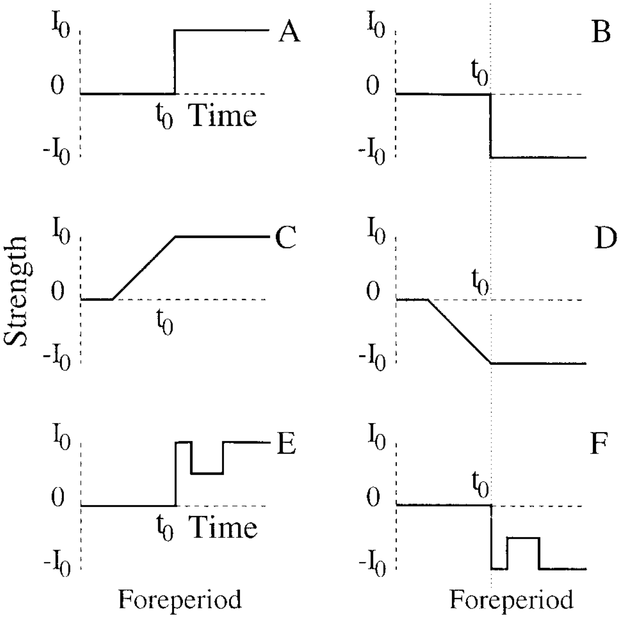

FIG. 1. Six strength functions: (A) a positive-stepped stimulus, (B) a negative-stepped stimulus, (C) a positive-ramped stimulus, (D) a negative-ramped stimulus, (E) a positive-notched stimulus, and (F) a negative-notched stimulus. The ramped stimuli dominate the stepped stimuli and the stepped stimuli dominate the notched stimuli.

The stimuli depicted in Panels A and B have abrupt changes and are hence termed *stepped stimuli.* Panels C and D depict stimuli in which the change is gradual; *ramped stimuli.* In Fig. 1, the ramp onsets of the ramped stimuli precede the step onsets of the stepped stimuli. This arrangement is deliberate. Ramped stimuli are either as extreme or more extreme than stepped stimuli. In the foreperiod (before t_0) the ramped stimuli are more extreme than their stepped counterparts, whereas after the foreperiod, the ramped and stepped stimuli are equally extreme. In this sense the ramped stimuli *dominate* the stepped stimuli. Panels E and F show another alternative to stepped stimuli: notched stimuli. These stimuli are dominated by stepped stimuli; that is, at all time points the stepped stimuli are as extreme or more extreme than their notched counterparts.

One plausible conjecture is that evidence accrual models predict the best performance to dominating stimuli. In particular, evidence accrual models may predict better performance to ramped stimuli than to stepped stimuli, and in turn, better performance to stepped stimuli than to notched stimuli. The goal of this paper is to formalize the class of evidence accrual models and to evaluate the veracity of this conjecture about dominance in evidence accrual models. If the conjecture is true, then using differing time courses yields a method of testing evidence accrual models.

The approach taken here is to define an evidence accrual model as a set of deterministic functions of an underlying stochastic process. The underlying stochastic process represents evidence from the stimulus. For example, a bounded random walk can be thought of as a model which (deterministically) computes the running sum of an evidence stochastic process. A response is produced when this sum becomes sufficiently extreme (Luce, 1986, defines random walks in this manner).

FROM STIMULUS TO EVIDENCE

To formalize the situation[1], let g be a strength function and let $\vec{\mathbf{G}}$ be a discrete time stochastic process corresponding to the evidence when g is presented. We refer to $\vec{\mathbf{G}}$ as the *evidence stochastic process*. For the case of discrete time, $\vec{\mathbf{G}}$ is a sequence of random variables and $\mathbf{G_i}$ is a member random variable representing the evidence sample at time i. We say that g *induces*[2] $\vec{\mathbf{G}}$. Figure 2 shows an example of an induction from a stepped stimulus (Panel A) into evidence samples (Panel B). It is important to note that there is no accrual of evidence yet in $\vec{\mathbf{G}}$. Each $\mathbf{G_i}$ is the amount of evidence gathered in the time interval i, but not the total amount of evidence. The construction of a nonaccumulative evidence stochastic process differs from other constructions in which the stochastic process represents the total accumulation of evidence gathered, such as in the diffusion model or the Ornstein–Uhlenbeck process. An example of this total evidence stochastic process from a random walk is shown in Panel C. The curve in Panel C is simply the running sum of the evidence samples in Panel B.

To proceed, an assumption about how stimuli induce evidence is needed:

Definition 1 (Monotonic Induction). Let g and h be stimulus strength functions. Let g and h induce stochastic processes $\vec{\mathbf{G}}$ and $\vec{\mathbf{H}}$, respectively. Let $F_{\mathbf{G_i}}$ and $F_{\mathbf{H_i}}$ be the distribution functions of the stochastic processes $\vec{\mathbf{G}}$ and $\vec{\mathbf{H}}$, respectively, at time i. The induction is called *monotonic* if

$$g \geqslant h \Rightarrow F_{\mathbf{G_i}}(y) \leqslant F_{\mathbf{H_i}}(y), \qquad \forall i, y. \tag{1}$$

In the above equation $g \geqslant h$ refers to the conventional meaning of dominance for functions: $g \geqslant h \Rightarrow g(i) \geqslant h(i)\ \forall i$. The ordering of distribution functions indicates that there is stochastic dominance of $\mathbf{H_i}$ by $\mathbf{G_i}$. If an induction is monotonic, then dominant strength functions induce dominant evidence stochastic processes.

One psychologically plausible monotonic induction is $\mathbf{G_i} = g(i) + \boldsymbol{\varepsilon}_\mathbf{i}$, where the $\boldsymbol{\varepsilon}_\mathbf{i}$ are independent and identically distributed. This case is the basis for stationary stochastic models such as random walks. Other monotonic inductions are possible. For example, several researchers have used linear filters to model early visual processing (e.g., Sperling, 1964; Graham, 1989). Linear filters with nonnegative impulse response functions can also be used as monotonic inductions, e.g., $\mathbf{G_i} = l_g(i) + \boldsymbol{\varepsilon}_\mathbf{i}$, where l denotes a linear filter operating on the strength function g at time i. An example of such a filter and its effects on stepped and ramped stimuli can be seen in Panel A of Fig. 3. A short proof that inductions through linear filters with nonnegative impulse-response functions are monotonic is provided in the Appendix.

[1] Random variables will be a denoted by a bold letter. Discrete time-stochastic processes (i.e., sequences of random variables) will be denoted by a vector arrow over a bold letter. Realizations of stochastic processes will be denoted by a vector arrow over a standard letter. Sets of realizations will be denoted by a script letter.

[2] For convenience, it is assumed that there is a countably infinite number of intervals i, with each interval corresponding to an evidence sample $\mathbf{G_i}$. This assumption does not provide a limitation—it is included as a minor technical consideration to insure that $\mathbf{G_i}$ is always well defined and never defective.

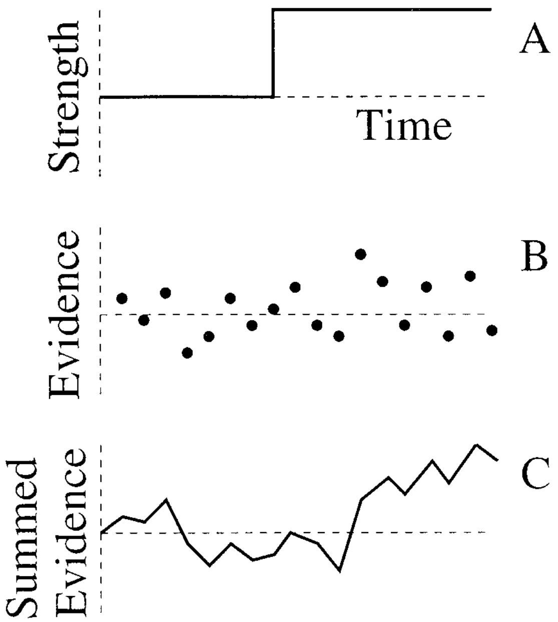

FIG. 2. An induction from a stepped stimulus (Panel A) to an evidence stochastic process (Panel B). Panel C shows the summation of evidence as specified by a random walk. Once the evidence stochastic process is realized (Panel B), the random walk summation shown in Panel C is deterministic.

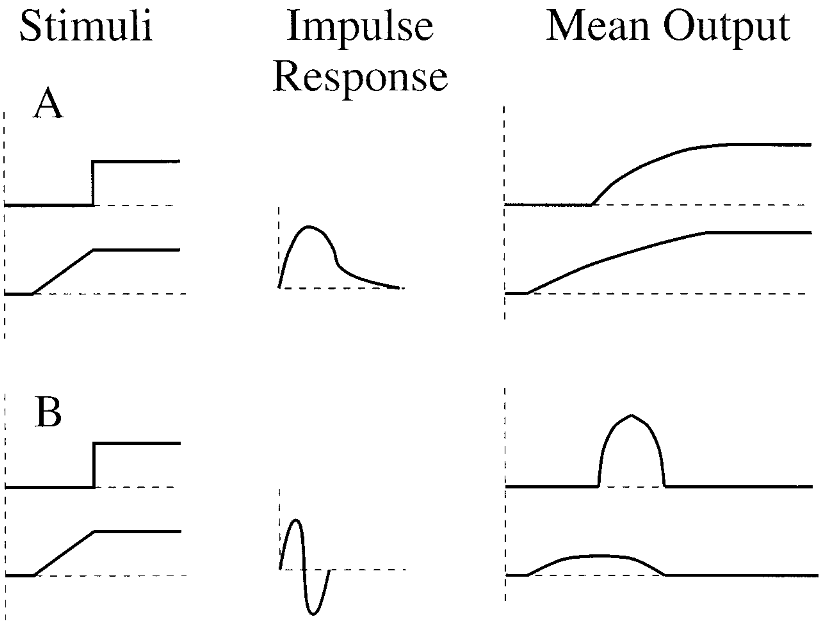

FIG. 3. Monotonic and nonmonotonic inductions. Panel A shows the case for a linear filter with a nonnegative impulse response function. The output for ramped stimuli dominates that for stepped stimuli. If the impulse response violates nonnegativity, the output from stepped stimuli may not dominate that from ramped stimuli, as depicted in Panel B.

There are psychologically plausible inductions which are not monotonic. Consider the case of a stochastic process $\vec{\mathbf{G}}$ which is induced by the strength function g such that $\mathbf{G_i} = g(i) + \boldsymbol{\varepsilon}_\mathbf{i}$, where $\boldsymbol{\varepsilon}_\mathbf{i}$ are independent but not identically distributed. Let each $\boldsymbol{\varepsilon}_\mathbf{i}$ be normally distributed with a mean of zero and a variance that is linearly related to stimulus strength. As strength increases the variance increases but the values of the smaller quantiles decrease, hence violating the monotonic induction assumption. Second, consider the case where the induction is through a linear filter, but the impulse function is not always nonnegative. An example of this filter and its effects on stepped and ramped stimuli can be seen in Panel B of Fig. 3. The induction through the linear filter of Panel B violates the monotonic induction assumption.

EVIDENCE ACCRUAL

In this paper evidence accrual models are defined as deterministic functions of realizations of the evidence stochastic process, $\vec{\mathbf{G}}$. On each trial there is a realization of the stochastic process $\vec{\mathbf{G}}$, and this realization is a sequence of real values denoted by $\vec{G}$. The ith member of the sequence is denoted by a real value G_i. For example, a bounded random walk can be defined as follows.

Definition 2. A bounded random walk model is composed of two real-valued deterministic functions, $L(\vec{G}, c_+, c_-)$ (the decision latency) and $X(\vec{G}, c_+, c_-)$ (the choice). These are functions of realizations of an evidence process, $\vec{G}$, and boundaries $c_+ > 0 > c_-$. L and X are defined in terms of the running sum of the evidence, $y(\vec{G}, k)$,

$$y(\vec{G}, k) = \sum_{i=0}^{k} G_i, \tag{2}$$

where k is a natural number. The latency L is the smallest k such that the running sum y exceeds one of the two boundaries. The response choice X is 1 if the sum exceeds the positive boundary, and it is 0 otherwise.

$$L(\vec{G}, c_+, c_-) = \min_k \{k\colon y(\vec{G}, k) > c_+ \texttt{ or } y(\vec{G}, k) < c_-\}. \tag{3a}$$

$$X(\vec{G}, c_+, c_-) = \begin{cases} 1 & y(\vec{G}, L) > c_+, \\ 0 & \texttt{otherwise.} \end{cases} \tag{3b}$$

Evidence accrual models consist of two deterministic functions of evidence realizations, one function for the alternative choice and the other for the decision latency. A formal definition is given below after a few points about notation. Realization $\vec{G}$ is said to dominate $\vec{H}$ if, for each value of i, $G_i \geqslant H_i$, and the dominance relation is written in the conventional form, $\vec{G} \geqslant \vec{H}$. $X(\vec{G})$ is a boolean variable that is defined as $X(\vec{G}) = 1$ if the realization $\vec{G}$ leads to a positive response and as $X(\vec{G}) = 0$ otherwise. $L_+(\vec{G})$ is the decision latency if the response is positive.

$L_+(\vec{G})$ is undefined if the response is not positive. As a matter of definition, we say that a model is an evidence accrual model if it meets the following two conditions.

DEFINITION 3 (Evidence Accrual Model).

$$\vec{G} \geqslant \vec{H} \Rightarrow [X(\vec{H}) = 1 \Rightarrow X(\vec{G}) = 1], \tag{4a}$$

$$[\vec{G} \geqslant \vec{H} \text{ and } X(\vec{H}) = 1] \Rightarrow L_+(\vec{G}) \leqslant L_+(\vec{H}). \tag{4b}$$

Simply put, for evidence accrual models, dominating realizations give rise to more accurate and quicker correct responses[3]. Evidence accrual, as defined here, is not probabilistic. It is a set of properties that a model possesses when conditioned on any realization, even those with measure zero.

LEMMA 1. *Bounded random walk models are evidence accrual models.*

Proof[4] *of Lemma* 1. Suppose $\vec{G} \geqslant \vec{H}$ and $X(\vec{H}) = 1$. We establish that Eqs. (4a) and (4b) hold; i.e., $X(\vec{G}) = 1$ and $L_+(\vec{G}) \leqslant L_+(\vec{H})$. Because $X(\vec{H}) = 1$ we know by definition (Eqs. (3a) and (3b)) that

$$y(\vec{H}, L_+(\vec{H})) \geqslant c_+$$

$$\forall t < L_+(\vec{H}), \qquad c_- < y(\vec{H}, t) < c_+.$$

By Eq. (2) and that $\vec{G} \geqslant \vec{H}$, we know

$$y(\vec{G}, k) = \sum_{i=0}^{k} G_i \geqslant \sum_{i=0}^{k} H_i = y(\vec{H}, k), \qquad \forall k. \tag{5}$$

Thus, for some $\tau \leqslant L_+(\vec{H})$ we may conclude

$$y(\vec{G}, \tau) \geqslant c_+,$$

$$\forall k < \tau, \qquad c_- < y(\vec{H}, k) \leqslant y(\vec{G}, k) < c_+.$$

Therefore, the minimum of such a τ defines $L_+(\vec{G})$ and so $L_+(\vec{G}) \leqslant L_+(\vec{H})$ and $X(\vec{G}) = 1$. ∎

DEFINITION 4. A bounded accumulator model is composed of two real-valued deterministic functions, $L(\vec{G}, c_+, c_-)$ (the decision time) and $X(\vec{G}, c_+, c_-)$ (the

[3] The term "correct response" refers specifically to a positive response. As a matter of convention, it is assumed that the correct response to the stimulus is the positive response: e.g., the presented stimulus is positive. All of the arguments in this paper apply to negative stimuli as well and there is no loss of generality from considering only positive stimuli. In fact, the arguments developed here apply to cases in which neither of the two answers is "correct," such as observers responding whether they agree with a public opinion statement.

[4] This proof is from Duncan Luce (personal communication).

choice). These are functions of realizations of an evidence process, $\vec{G}$, and boundaries $c_-, c_+ > 0$. Let functions I_+ and I_- be defined as

$$I_+(x) = \begin{cases} x, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

$$I_-(x) = \begin{cases} 0, & x > 0, \\ -x, & x \leqslant 0. \end{cases}$$

L and X are defined in terms of the positive and negative accumulators $y_+(\vec{G}, k)$ and $y_-(\vec{G}, k)$, respectively,

$$y_+(\vec{G}, k) = \sum_{i=0}^{k} I_+(G_i), \tag{6a}$$

$$y_-(\vec{G}, k) = \sum_{i=0}^{k} I_-(G_i). \tag{6b}$$

The functions L and X are defined as

$$L(\vec{G}) = \min_k \{k\text{: } y_+(\vec{G}, k) > c_+ \texttt{ or } y_-(\vec{G}, k) > c_-\}. \tag{7a}$$

$$X(\vec{G}) = \begin{cases} 1 & y_+(\vec{G}, L) > c_+, \\ 0 & \texttt{otherwise.} \end{cases} \tag{7b}$$

LEMMA 2. *Bounded accumulator models are evidence accrual models.*

Proof of Lemma 2. Suppose $\vec{G} \geqslant \vec{H}$ and $X(\vec{H}) = 1$. We establish that Eqs. (4a) and (4b) hold; i.e., $X(\vec{G}) = 1$ and $L_+(\vec{G}) \leqslant L_+(\vec{H})$. Because $X(\vec{H}) = 1$ we know by definition that

$$\begin{aligned} & y_+(\vec{H}, L_+(\vec{H})) \geqslant c_+, \\ \forall t < L_+(\vec{H}), \quad & y_+(\vec{H}, t) < c_+, \\ \forall t < L_+(\vec{H}), \quad & y_-(\vec{H}, t) < c_-. \end{aligned}$$

By Eqs. (6a) and (6b) and that $\vec{G} \geqslant \vec{H}$, we know that

$$\begin{aligned} y_+(\vec{G}, k) \geqslant y_+(\vec{H}, k), \quad & \forall k, \\ y_-(\vec{G}, k) \leqslant y_-(\vec{H}, k), \quad & \forall k. \end{aligned}$$

Thus, for some $\tau \leqslant L_+(B)$ we may conclude that

$$\begin{aligned} & y_+(A, \tau)) \geqslant c_+, \\ \forall t < \tau, \quad & y_+(A, t) < c_-, \\ \forall t < \tau \quad & y_-(A, t) < c_-. \end{aligned}$$

Therefore the minimum of such a τ defines $L_+(\vec{G})$ and so $L_+(\vec{G}) \leqslant L_+(\vec{H})$ and $X(\vec{G}) = 1$. ∎

DOMINANCE AND PERFORMANCE

The critical questions are about the effects of strength function dominance on a model's predictions for performance. Evidence accrual is defined in terms of realizations of stochastic processes. The remaining step is to integrate out all of the realizations. Theorem 3 relates accuracy for evidence accrual models to strength dominance. It states that (under mild technical conditions) monotonic induction and evidence accrual are sufficient for more accurate responses to the dominating strength functions.

THEOREM 3. *Let g and h be strength functions that monotonically induce evidence stochastic processes* $\vec{\mathbf{G}}$ *and* $\vec{\mathbf{H}}$, *respectively. Let the distribution functions* $F_{\mathbf{G_i}}$ *and* $F_{\mathbf{H_i}}$ *be absolutely continuous and strictly increasing. Let* $P(g)$ *be the probability of a correct response to a stimulus with strength function g. If* $g \geqslant h$, *then for all evidence accrual models* $P(g) \geqslant P(h)$.

Proof of Theorem 3. Because each $F_{\mathbf{H_i}}$ and each $F_{\mathbf{G_i}}$ is strictly increasing and absolutely continuous, there exist inverse functions $F_{\mathbf{H_i}}^{-1}$ and $F_{\mathbf{G_i}}^{-1}$ which are monotonically increasing and absolutely continuous. The evidence process $\vec{\mathbf{G}}$ can be expressed as a function of $\vec{\mathbf{H}}$ as

$$\mathbf{G_i} = F_{\mathbf{G_i}}^{-1} \circ F_{\mathbf{H_i}}(\mathbf{H_i}),$$

where $\circ$ denotes function composition.

Because each $F_{\mathbf{H_i}}$ and $F_{\mathbf{G_i}}$ is strictly increasing and absolutely continuous, the sample space for $\mathbf{G_i}$ and $\mathbf{H_i}$ is the real line. Without any loss, we can assume that the evidence processes are countably infinite in length. The sample space for the evidence processes $\vec{\mathbf{G}}$ and $\vec{\mathbf{H}}$ is the countable infinite cross product of reals; $\lim_{N \to \infty} \mathscr{R}^{(N)}$. We denote this limit as $\mathscr{R}^{(\infty)}$. Let $\vec{F}_{\mathbf{H}}$ be a function that maps sequences of real values into sequences of reals such that, for any $\vec{e} \in \mathscr{R}^{(\infty)}$,

$$\vec{F}_{\mathbf{H}}(\vec{e}) = (F_{\mathbf{H_1}}(e_1), F_{\mathbf{H_2}}(e_2), \ldots).$$

Then,

$$\vec{\mathbf{G}} = \vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{H}}(\vec{\mathbf{H}}). \tag{8}$$

The probability of a correct response for stimulus g is

$$P(g) = \mathrm{E}[X(\vec{\mathbf{G}})] = \int_{\vec{e} \in \mathscr{R}^{(\infty)}} X(\vec{e})\, dF_{\vec{\mathbf{G}}}(\vec{e}), \tag{9}$$

where $dF_{\vec{\mathbf{G}}}(\vec{e})$ is the probability measure of realization $\vec{e}$. Substituting Eq. (8) into the expression for $P(g)$ yields

$$P(g) = \mathrm{E}[\,X(\vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{H}}(\vec{\mathbf{H}}))\,] = \int_{\vec{e} \in \mathscr{R}^{(\infty)}} X(\vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{H}}(\vec{e}))\, dF_{\vec{\mathbf{H}}}(\vec{e}). \tag{10}$$

Likewise, the evidence process $\vec{\mathbf{H}}$ can be expressed as

$$\vec{\mathbf{H}} = \vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{G}}(\vec{\mathbf{H}}),$$

and

$$P(h) = \mathrm{E}[\,X(\vec{F}_{\mathbf{G}}^{-1} \circ F_{\mathbf{G}} \circ \vec{\mathbf{H}})\,] = \int_{\vec{e} \in \mathscr{R}^{(\infty)}} X(\vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{G}}(\vec{e}))\, dF_{\vec{\mathbf{H}}}(\vec{e}). \tag{11}$$

To complete the proof we show $P(g) > P(h)$. From Eqs. (10) and (11)

$$P(g) - P(h) = \int_{\vec{e} \in \mathscr{R}^{(\infty)}} (X(\vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{H}}(\vec{e})) - X(\vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{G}}(\vec{e})))\, dF_{\vec{\mathbf{H}}}(\vec{e}). \tag{12}$$

By monotonic induction, $g \geqslant h$ implies $F_{\mathbf{G_i}}(y) \leqslant F_{\mathbf{H_i}}(y), \forall i, y$. $F_{\mathbf{G_i}}^{-1}$ is absolutely continuous and strictly increasing. Hence, for every y and i, $F_{\mathbf{G_i}}^{-1} \circ F_{\mathbf{H_i}}(y) \geqslant F_{\mathbf{G_i}}^{-1} \circ F_{\mathbf{G_i}}(y)$, and likewise, $\vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{H}}(\vec{e}) \geqslant \vec{F}_{\mathbf{G}}^{-1} \circ \vec{F}_{\mathbf{G}}(\vec{e}), \forall \vec{e}$. Furthermore, by the definition of evidence accrual, X is monotonically increasing. Hence the integrand in Eq. (12) is always positive, proving the theorem. ▮

Although monotonic induction and evidence accrual are sufficient for more accurate responses to stimuli with dominant strength functions, they are not sufficient for quicker correct latencies to stimuli with dominant strength functions. The following computer simulation demonstrates that there exists a case in which the mean correct decision time is *slower* to a dominating stimulus than to a dominated one. Let $\boldsymbol{\varepsilon}_\mathbf{i}$ be independent, exponentially distributed random variables with density $f(z) = e^{-(z+1)}$, $z > -1$. Each $\boldsymbol{\varepsilon}_\mathbf{i}$ "starts" at -1 and has a rate of 1 and a mean of zero. Although this distribution of evidence has not been used in a psychological model, it is sufficient to demonstrate that dominance does not necessarily imply quicker correct mean decision times. Let the stimulus intensity function g induce evidence $\mathbf{G}_i = g(i) + \boldsymbol{\varepsilon}_\mathbf{i}$. Consider an accumulator model in which negative realizations of $\mathbf{G_i}$ are summed in an accumulator with criteria $c_- = 3$ and positive realizations are summed in an accumulator with criteria $c_+ = 7$. The dominating stimulus (g) was a ramped stimulus and the dominated stimulus (h) was a stepped stimulus.

$$g = \begin{cases} .01{*}i & 0 \leqslant i < 8, \\ .08, & i \geqslant 8. \end{cases}$$

$$h = \begin{cases} 0, & i < 8, \\ .08, & i \geqslant 8. \end{cases}$$

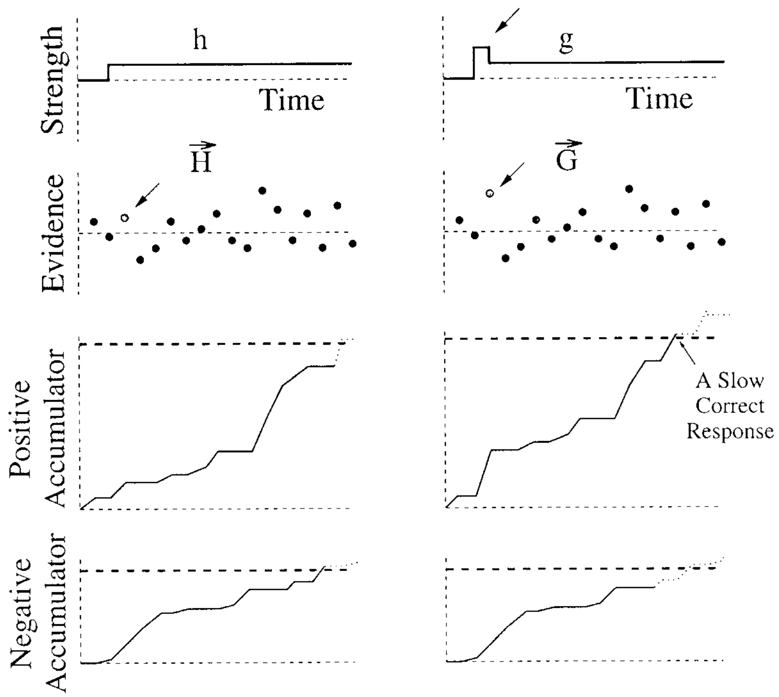

FIG. 4. A graphical explanation of how evidence accrual and monotonic induction fail to predict quicker latency for ramped stimuli. See the text for details.

Ten simulation runs of 500,000 trials each were performed. In all 10 runs the correct response latencies were longer for ramped stimuli than for stepped stimuli. The probability of this happening by chance is less than .001. The above demonstration made use of an unusual distribution of evidence, a translated exponential. The exponential was used because it provides for a large degree of skewness and, as will be shown, this skewness is an important factor in determining the relative speeds of correct responses to ramped and stepped stimuli.

Figure 4 provides a graphical explanation of why evidence accrual and monotonic induction are insufficient for quicker correct responses to dominating stimuli. There are two stimulus strength functions displayed in the first row of panels: g and h, with $g \geqslant h$. There is only a small interval in time where $g > h$, and this interval is indicated by the top arrow. The evidence stochastic processes induced are $\mathbf{G}_i = g_i + \varepsilon_i$ and $\mathbf{H}_i = h_i + \varepsilon_i$, where ε_i are independent and identically distributed. Realizations $\vec{G}$ and $\vec{H}$ are shown in the second row of panels. $\vec{G}$ and $\vec{H}$ are both generated from the same realization $\vec{\varepsilon}$, and hence both have the same probability. The strength difference between g and h is reflected in the third time component of $\vec{G}$ and $\vec{H}$. These two differing components are indicated by the arrows in the second row of panels. Positive and negative accumulator values (the rectified running sums) are shown in the bottom two panels. The negative accumulator of $\vec{H}$ becomes absorbed first, producing an error. The positive accumulator of $\vec{G}$ becomes absorbed first, producing a correct response. For this realization, the absorption took a long time; longer than the average absorption time (both counters have high values). However, the decision time of this trial only affects the mean correct response time for stimulus g and not for stimulus h. If there are several realizations like $\vec{G}$ and $\vec{H}$, then the correct latency may be longer for the stimulus strength g than for h. Unfortunately, it is difficult to derive expressions that assess the contribution of paths like $\vec{G}$ and $\vec{H}$ to correct response time predictions for a given model and induction.

SIMULATIONS

The computer-simulated counterexample showed that there are accumulator models in which correct responses to ramped stimuli are slower than those to stepped stimuli. The example is idiosyncratic in that the accumulator boundaries are asymmetric and favor the incorrect response. For this case, correct response probabilities are extremely low in the simulations (.12 for stepped stimuli and .21 for ramped stimuli). To explore whether the phenomenon of slower correct responses to ramped stimuli (than to stepped stimuli) occurs generally or just in cases with extreme response bias, accumulator and random walk models were extensively simulated as follows.

1. The first step in simulating both models was to specify an evidence accrual stochastic process. At each time interval a distribution with a mean of zero was sampled. There were three different such distributions: a normal distribution, a

TABLE 1

Simulations of the Random Walk and Accumulator Models

Factors	Levels
Factors for both random walks and accumulators	
Sample mean	.6, .1, .01
Sample distribution	normal
	positively skewed exponential
	negatively skewed exponential
Boundary decay	1.0, .98 (1.0 = no decay)
Boundary height with no decay	5, 10, 15
Boundary height with decay of.98	20, 40, 60
Boundary asymmetry	1, 1.4, .6
Notch depth	80% of strength over baseline
Factors for random walks	
Ramp width	60 samples, 30 samples
Number of premature samples (ramp width of 60)	15, 30, 45, 60
Number of premature samples (ramp width of 30)	7, 15, 22, 30
Notch duration	40
Notch SOA	5
Factors for accumulators	
Ramp width	8 samples
Number of premature samples	2, 4, 6, 8
Notch duration	20
Notch SOA	2

positively skewed exponential (this was the distribution used in the above example), and a negatively skewed exponential. The normal distribution was chosen because it is typically used in evidence accrual models and is a plausible evidence distribution. The exponentials were chosen because they have the property of extreme skewness. If skew is an important variable in determining the order of correct mean latencies for stepped and ramped stimuli, then simulating the processes with the exponential offers a powerful means of observing this relationship. To incorporate the stimulus strength the zero-mean samples were translated by the stimulus strength. The sample mean entries in Table 1 show the amount of translation when the stimulus was at full strength (the extremum of the step or ramp).

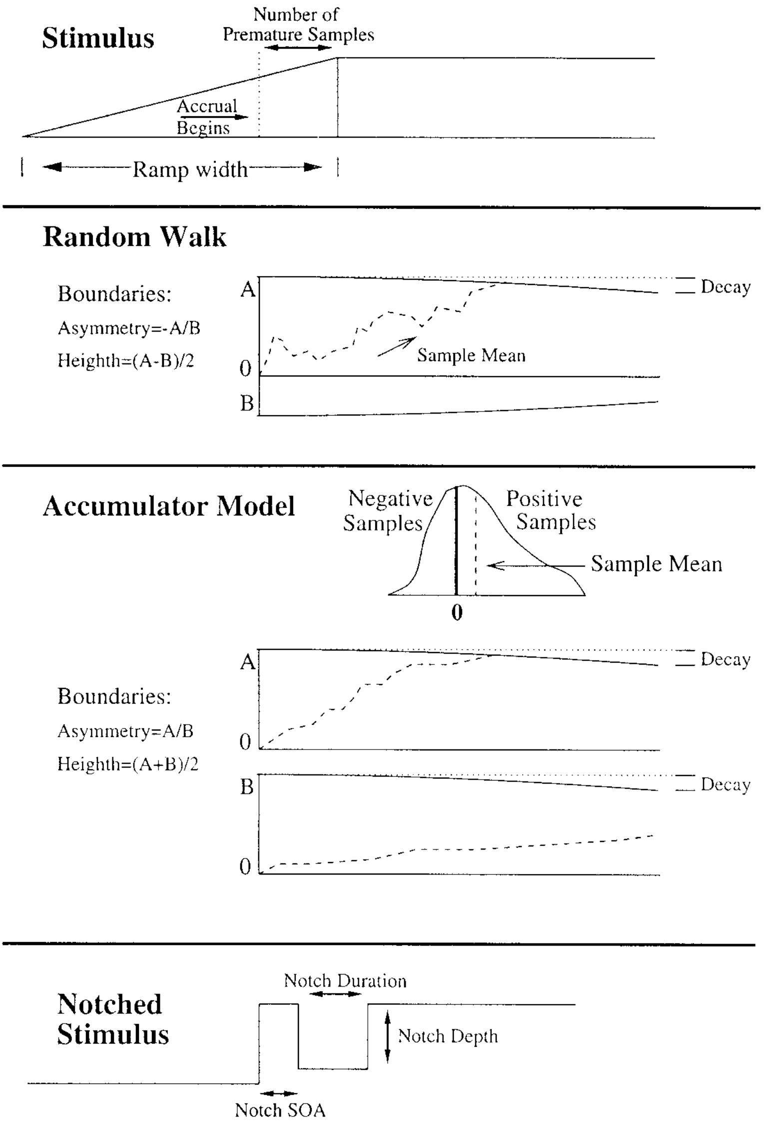

FIG. 5. A schematic of the simulation parameters. The values of parameters in the simulations are given in Table 1.

2. Random walks and accumulators were implemented as deterministic algorithms on the evidence stochastic process. For the random walk models, the (translated) samples were added until they crossed one of the two boundaries. For the accumulator model, if the translated sample was greater than zero its value was accumulated in the positive accumulator. If the translated sample was less than or equal to zero, its value was accumulated in the negative accumulator. A graphical sketch of these algorithms, as well as the role of some of the manipulated factors, are depicted in Fig. 5.

3. The boundaries were also manipulated in several ways; in some conditions they were quite large or, alternatively, quite small. The asymmetry of the boundaries and boundary decay were also manipulated (see Table 1 and Fig. 5). Boundary decay was implemented by allowing the boundaries in successive time intervals to be scaled by .98. Boundary decay and boundary asymmetry are also depicted in Fig. 5.

4. There were two other factors manipulated in the simulation: the width of the ramp and the time point at which sampling began. The width of the ramp was the number of time intervals in which the strength was slowly changing for ramped stimuli. Accrual always began prematurely, that is, before the stepped onset of the stepped stimuli, but the number of premature samples was manipulated through four levels. A schematic view of these two factors is shown in Fig. 5 and the levels of these factors are indicated in Table 1.

These manipulations were crossed to produce 1,296 random walk simulation conditions and 648 accumulator model simulation conditions. Each condition was simulated for 10,000 trials.

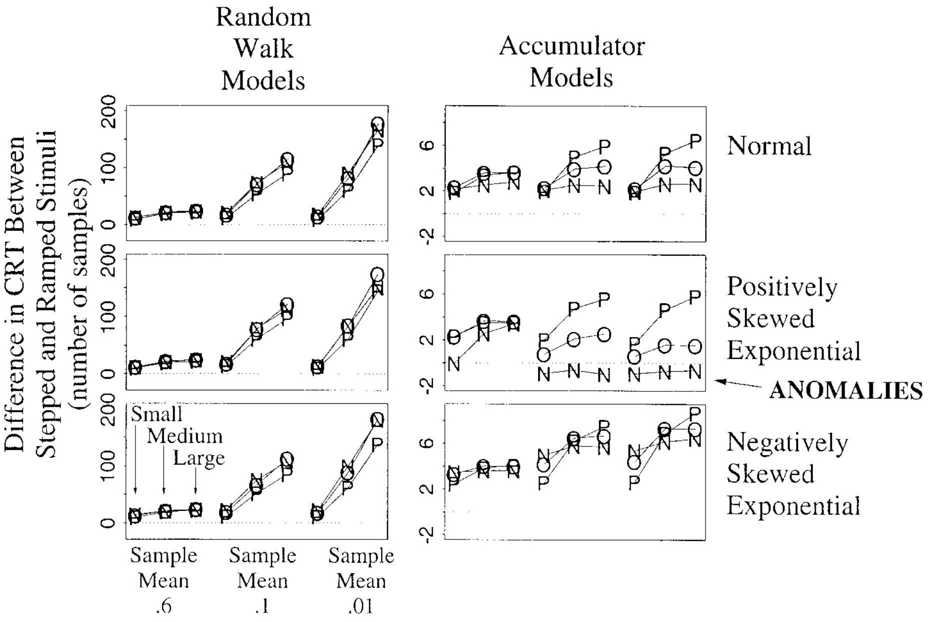

FIG. 6. Simulation results for random walk and accumulator models processing stepped and ramped stimuli. Anomalous cases occurred only when the boundaries favored an error response and the evidence was skewed toward the positive response. See the text for details.

Figure 6 shows the results of the simulations for both random walk and accumulator models. Each of the 6 plots has 9 lines connecting 27 points labeled with "O," "P," or "N." There is one point for each condition and the y-coordinate of a point denotes the difference in correct response time between stepped and ramped stimuli for any one of the simulation conditions. Positive differences indicate that correct responses to stepped stimuli took longer than those to ramped stimuli. There are 3 points within each line, and each point corresponds to a different boundary height. The left-most point on each line corresponds to small boundaries (5), the center point corresponds to medium boundaries (10), and the right-most point corresponds to large boundaries (20). The points labeled "O" are for symmetric bounds, the points labeled "P" are for asymmetric bounds that favor the positive response, and the points labeled "N" are for asymmetric bounds that favor the negative response. The 9 lines in each plot are clustered into three groups: the left-most, middle, and right-most groups of curves are for sample means of .6, .1, and .01, respectively. There are 6 different plots; the left-hand column of plots displays results for the random walk model and the right-hand column of plots displays results for the accumulator model. The rows of plots correspond to different forms of the evidence distribution as indicated on the right-hand side of the figure. The figure shows only the results for ramps of widths of 30 samples (random walk) and 8 samples (accumulator) with no boundary decay and accrual starting at the beginning of the ramps. The results are qualitatively similar for the other conditions with boundary decay and less premature sampling.

None of the 1,296 random walk simulations produced slower mean correct latencies to ramped stimuli than to stepped stimuli. This can be seen in Fig. 6 by observing that none of the curves for the random walk model are negative. In most cases the speed-up for ramped stimuli was massive. Figure 7 displays the RT advantage. Correct response latencies to stepped and ramped stimuli are both plotted for

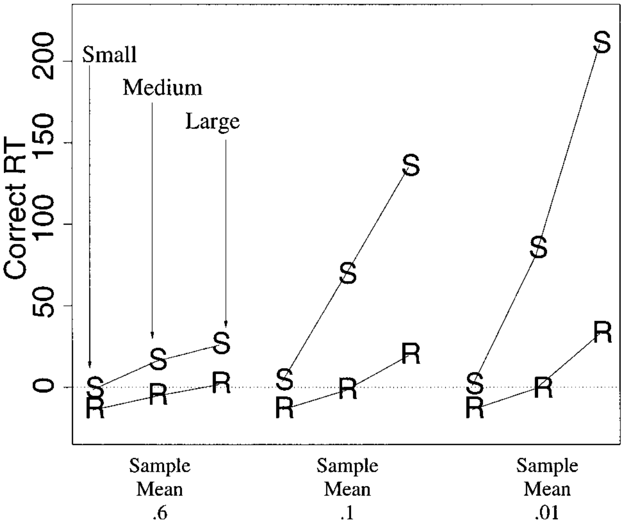

FIG. 7. Simulation results for the random walk with symmetrical bounds and normally distributed evidence samples. The letter "S" denotes responses to stepped stimuli. The letter "R" denotes responses to ramped stimuli.

a random-walk model processing normally distributed evidence with symmetric bounds.

The same conclusion is not true for the accumulator: in 83 of the 648 conditions, the mean correct latency to ramped stimuli was slower than that to stepped stimuli. The anomalous cases are highlighted by the "ANOMALIES" indicator on the right-hand side of Fig. 6. All of the anomalous cases occur under the condition that the boundaries are skewed favoring a negative response and the distribution of the evidence samples is skewed toward the positive response. The boundary asymmetry in these cases yields several errors: the probability of a correct response ranges between .05 and .35. The response bias is so severe that the responses are below chance. There were no simulated cases in which anomalies occurred and accuracy is above chance. Overall, the simulations show that, for practical purposes, correct responses to stepped stimuli are slower than those to ramped stimuli.

To further investigate violation of the conjecture that evidence accrual models predict quicker decision times to dominating stimuli, the random walk and accumulator models were simulated with stepped and notched stimuli. The format of these simulations was similar to that of the previous ones. The factors related to the boundaries and the skew of the evidence samples were the same as before, but there were additional factors related to the time course of the notch manipulation (see Fig. 5 for a schematic of these factors). For the random walk, the notch began five samples after the stepped onset (this is called the notch SOA) and lasted 40 samples in duration. For the accumulator, the notch began 2 samples after onset and lasted 20 samples in duration. The notch corresponded to an 80% reduction

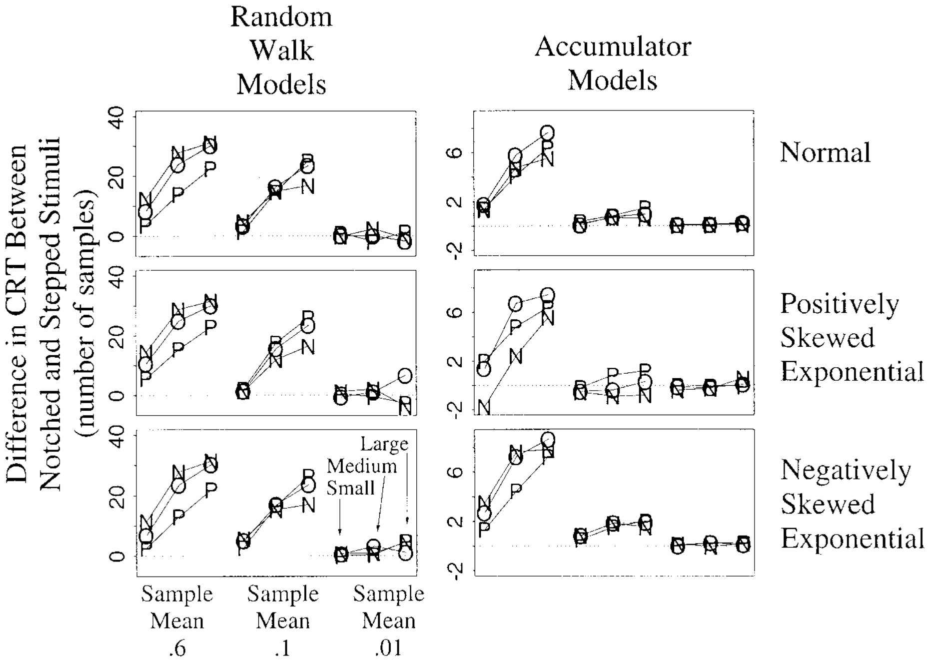

FIG. 8. Simulation results for random walk and accumulator models processing stepped and notched stimuli. See the text for details.

in strength (this factor is termed the notch depth). Accumulation began at the stepped onset for both stimulus types. There were 162 simulated conditions for each model.

The results of the simulations are shown in Fig. 8. The figure follows the same format as those for previous simulations. The plotted results are for the cases with no boundary decay, but the results for the cases with boundary decay are qualitatively similar. For most conditions, there is a response time advantage for stepped stimuli relative to notched stimuli, but there are exceptions. For the random walk, the exceptions are most likely due to unsystematic variability in the (simulated) mean response time (RT).[5] The exceptions for the accumulator model, which occur for negatively skewed boundaries and positively skewed evidence samples, are larger than the (two-tailed) 95% confidence interval around the no-difference point and hence most likely represent true anomalies. Overall, the results parallel the previous results for stepped and ramped stimuli. Except for cases with inordinate asymmetry and a corresponding extreme degree of response bias, correct responses to stepped stimuli are quicker than those to notched stimuli. There is an important difference between these results and the previous ones. The largest differences between correct response times for stepped and ramped stimuli occur for less intense stimuli (small sample means) while the largest differences for stepped and notched stimuli occur for more intense stimuli (large sample means). Experimenters who wish to test evidence accrual models may make use of this fact to enhance the power of their designs.

CONCLUSION

The main result of this paper is that evidence accrual models *do* predict more accurate responses to dominant strength functions, but correct responses are *not* necessarily quicker. There are conditions under which correct responses to ramped stimuli of Fig. 1 are slower than those to stepped stimuli. However, as shown by simulation, these anomalous cases occur only in accumulator models in which the response boundaries are skewed such that the correct responses are infrequent. It is fairly trivial to design experiments without such extreme response biases and hence to avoid these anomalous cases. As a rule of thumb, in the absence of extreme response bias, evidence accrual models predict quicker mean correct response times as well as greater accuracy to dominating stimuli.

When Accrual Begins

There is a limitation to the results presented here—they are conditioned on evidence accrual beginning at the same time point regardless of the time course of the stimulus. This assumption may not hold in some experimental paradigms. For example, it is plausible that evidence accrual to the ramped stimulus could begin

[5] To approximate the standard error of the simulated mean RT, the apparently anomalous conditions were simulated 100 times. The small negative differences between notched and stepped RT for the random walk models are well within the 95% confidence interval and do not likely reflect a true negative difference.

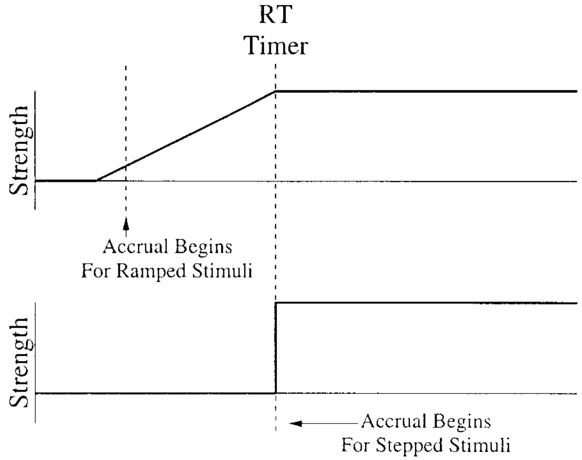

FIG. 9. An illustration of differing starting points of accruals for stepped and ramped stimuli. It is plausible that accrual to ramped stimuli begins before accrual to stepped stimuli. In this case, the ramped stimuli do not dominate stepped stimuli. Even so, processing of ramped stimuli occurs earlier than processing of stepped stimuli relative to the RT timer.

toward the start of the ramp[6] but evidence accrual to the corresponding stepped stimulus could begin at the stepped onset. This arrangement is shown in Fig. 9. Relative to the points where accrual begins, ramped stimuli would no longer dominate stepped stimuli, but instead would be dominated by them. Hence the point of accrual is a critical issue, and in the preceding analyses it has been assumed that the starting point of accrual is the same for all stimulus types. Unfortunately, it is difficult to find empirical or theoretical guidance about the time point at which accrual begins. The conventional approach is to use stepped stimuli and to assume sampling begins at the stepped onset. This convention may be plausible in standard paradigms because the time between the warning signal and the stimulus is short and constant and hence easily learned. There have been critiques of this assumption and researchers have considered the case where observers begin accruing samples before onset, e.g., premature sampling (Laming, 1968; Rouder, 1996).

Researchers concerned about the assumption that the time point at which accrual begins is constant across different stimulus types have two alternatives. The first is to limit empirical tests to the comparison of stepped and notched stimuli. Because notched and stepped stimuli share a common abrupt onset, it is far less plausible that accrual should begin at different time points. Second, researchers who use stepped and ramped stimuli can still test the conjunction of evidence accrual and monotonic induction if they are willing to assume that accrual begins earlier for ramped stimuli than for stepped stimuli, such as in Fig. 9. This assumption is plausible (after all, the ramp stimulus does appear earlier). The effect of accruing earlier for ramped stimuli than for stepped stimuli is to speed the decision latency to ramped stimuli relative to that to stepped stimuli. To show this, all 1944 simulations

[6] It is unlikely that accumulation would begin at the exact start of the ramp. Presumably, the small strength changes at the beginning of the ramp portion would be undetectable against a background of perceptual noise.

of the random walk and accumulator model were repeated under the condition that accrual to ramped stimuli starts at the beginning of the ramp while accrual to stepped stimuli starts at the beginning of the stepped onset. All of the 1944 simulations resulted in quicker mean correct decision times to ramped stimuli than to stepped stimuli. There were no anomalous cases.

Empirical Applications

Rouder (1995, 2000) used the results presented here to test the conjunction of a monotonic induction and evidence accrual in a two-choice luminance discrimination task. He presented observers with stepped and ramped stimuli (with the ramped portion of the ramped stimuli preceding the corresponding stepped stimuli) in a mixed-list design. The foreperiod for both the stepped and ramped stimuli was variable, and hence it was impossible for the observers to accurately time the onset of the stimuli. As shown in Fig. 10, Rouder found an unexpected interaction of foreperiod and onset on performance. When the foreperiod was short in duration, the responses to stepped stimuli were more accurate and quicker than responses to ramped stimuli. This result indicates that for these short-foreperiod stimuli the conjunction of monotonic induction and evidence accrual cannot hold. The reverse pattern held for stimuli that were presented after a long-duration foreperiod; responses were more accurate and quicker to ramped stimuli. To explain the interaction, Rouder rejected the monotonic induction assumption and developed a model in which evidence reflects contributions from the overall luminance level (see Panel A of Fig. 3) as well as the rate of change of the luminance level (see Panel B of Fig. 3). Early in the trial, the observers monitor for abrupt changes in luminance and then, if this strategy fails, they switch and monitor the steady-state luminance level.

Rouder's (1995, 2000) empirical application demonstrates the usefulness of the results presented here. By using stimuli with differing time courses, Rouder concluded that the link between the evidence and the luminance stimuli is complex; it is mediated by channels sensitive to quick temporal change as well as by overall luminance level. It is an empirical question whether the use of stimuli with differing

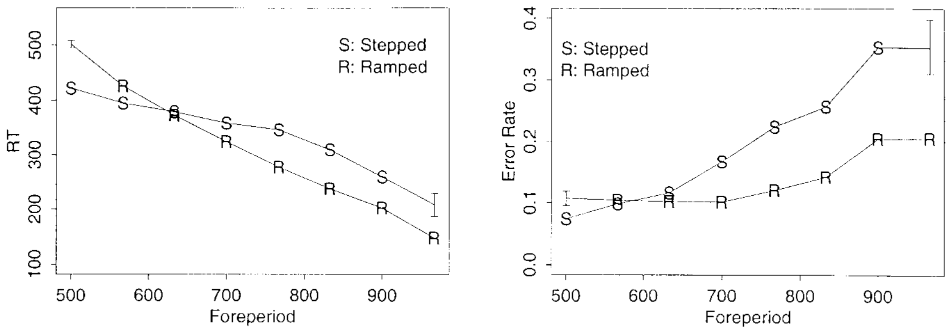

FIG. 10. Correct RT and error rate as a function of foreperiod for stepped and ramped stimuli (from Rouder, 2000). Performance is better to stepped stimuli than to ramped stimuli for stimuli with short-duration foreperiods, but the reverse holds for stimuli with long-duration foreperiods. Error bars represent 95% confidence intervals.

time courses such as stepped, ramped, and notched stimuli can yield similarly rich results in other domains.

APPENDIX

We present proof that inductions through linear filters with positive impulse response functions are monotonic.

Statement. Let $\mathbf{G_i} = l_g(i) + \boldsymbol{\varepsilon}_\mathbf{i}$, where the $\boldsymbol{\varepsilon}_\mathbf{i}$ are independent and identically distributed and l is a linear filter with a nonnegative impulse-response function. The induction from g to $\vec{\mathbf{G}}$ is monotonic.

Proof. Let g and h be two strength functions such that $g \geqslant h$. The response l_g of a linear filter with impulse response function I to stimulus strength g is

$$l_g(i) = \int_{-\infty}^{i} I(x)\, g(x-i)\, dx.$$

For a fixed value of i, $g(x-i) \geqslant h(x-i)$ for all x. When the impulse response function is nonnegative, $I(x)\, g(x-i) \geqslant I(x)\, h(x-i)$, for all x. Likewise,

$$l_g(i) = \int_{-\infty}^{i} I(x)\, g(x-i)\, dx \geqslant \int_{-\infty}^{i} I(x)\, h(x-i)\, dx = l_h(i).$$

Let $\mathbf{G_i} = l_g(i) + \boldsymbol{\varepsilon}_\mathbf{i}$ and $\mathbf{H_i} = l_h(i) + \boldsymbol{\varepsilon}_\mathbf{i}$. $l_g(i) \geqslant l_h(i)$ implies $F_{\mathbf{G_i}}(x) \leqslant F_{\mathbf{H_i}}(x)$. ∎

REFERENCES

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

Graham, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford.

Heath, R. A. (1984). Random-walk and accumulator models of psychological discrimination: A critical evaluation. *Perception*, **13**, 57–65.

LaBerge, D. L. (1963). A recruitment theory of simple behavior. *Psychometrika*, **27**, 375–396.

LaBerge, D. (1994). Quantitative models of attention and response processes in shape identification tasks. *Journal of Mathematical Psychology*, **38**, 198–243.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.

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

Link, S. W. (1992). *Wave theory of difference and similarity*. Hillsdale. N.J.: Erlbaum.

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

Luce, R. D. (1986). *Response times*. New York: Oxford University Press.

Pickett, R. M. (1967). Response latency in a pattern perception situation. *Acta Psychologica*, **27**, 160–169.

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

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

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

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

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

Rouder, J. N. (2000). Assessing the roles of change discrimination and luminance integration: Evidence for a hybrid race model of perceptual decision making in luminance discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, **26**.

Sanders, A. F., & Ter Linden, W. (1967). Decision making during paced arrival of probabilistic information. *Acta Psychologica*, **27**, 170–177.

Smith, P. L. (1990). A note on the distribution of response time for a random walk with gaussian increments. *Journal of Mathematical Psychology*, **34**, 445–459.

Smith, P. L. (1995). Multiple detector models of visual simple reaction time. *Psychological Review*, **102**, 567–593.

Sperling, G. (1964). Linear theory and the psychophysics of flicker. *Documenta Opthamologica*, **18**, 3–15.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, **13**, 37–58.

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

Vickers, D., & Smith, P. (1985). Accumulator and random-walk models of psychophysical discrimination: A counter-evaluation. *Perception*, **14**, 471–497.

Received September 30, 1998; revised November 19, 1999