



The simplest complete model of choice response time: Linear ballistic accumulation

Scott D. Brown*, Andrew Heathcote

School of Psychology, University of Newcastle, Callaghan, NSW 2308, Australia

Accepted 7 December 2007

Abstract

We propose a linear ballistic accumulator (LBA) model of decision making and reaction time. The LBA is simpler than other models of choice response time, with independent accumulators that race towards a common response threshold. Activity in the accumulators increases in a linear and deterministic manner. The simplicity of the model allows complete analytic solutions for choices between any number of alternatives. These solutions (and freely-available computer code) make the model easy to apply to both binary and multiple choice situations. Using data from five previously published experiments, we demonstrate that the LBA model successfully accommodates empirical phenomena from binary and multiple choice tasks that have proven difficult for other theoretical accounts. Our results are encouraging in a field beset by the tradeoff between complexity and completeness.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Reaction time; Response time; Lexical decision; Decision; Choice; Mathematical models

1. Introduction

The psychological processes that govern decision making have been the focus of detailed study for over half a century. The choices under study are usually quite simple: for example, a perceptual decision about brightness; a choice about one's memory for a

* Corresponding author.

E-mail address: scott.brown@newcastle.edu.au (S.D. Brown).

URL: <http://science-it.newcastle.edu.au/~sdb231/> (S.D. Brown).

probe stimulus; or even a simple judgment of the legality of word vs. non-word letter strings. Occasionally, choices between more than two alternatives are considered (e.g., Busemeyer & Townsend, 1992, 1993; Lacouture & Marley, 1991, 1995, 2004; Usher, Olami, & McClelland, 2002). Theories of choice are tested on their ability to accommodate empirical patterns of choice probability and latency (response time, RT, although some models have also considered confidence: Vickers & Lee, 1998, 2000). Even when we restrict our focus to binary choices between simple alternatives in a well-controlled environment, the richness of empirical observations is overwhelming. Complicated effects are observed on the shape of RT distributions, the relative speed of correct and incorrect responses, and the interaction of all these with error rates. This richness has resulted in models of choice RT becoming increasingly complicated over the past 50 years (amongst others, see: Laming, 1968; Link & Heath, 1975; McClelland, 1979; McMillen & Holmes, 2006; Ratcliff, 1978, 1988; Stone, 1960; Usher & McClelland, 2001; Usher et al., 2002; Van Zandt, Colonus, & Proctor, 2000; for reviews see Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Luce, 1986; Ratcliff & Smith, 2004).

We hope to advance this research effort with a new theory of choice RT—the linear ballistic accumulator (LBA). The LBA is simpler than the leading models in the field, and yet it accommodates all the important empirical phenomena. The LBA also comes with detailed but simple analytic solutions, making it easy to apply. The model's success in accounting for empirical phenomena is surprising, given it shares essential properties with older models that have proven inadequate. The LBA uses linear, independent response accumulators, as in (Vickers, 1970, 1978, 1979; Smith & Vickers, 1988, 1989) and some others (e.g., Audley & Pike, 1965; Laberge, 1962; Townsend & Ashby, 1983; Van Zandt et al., 2000). This arrangement has proven inadequate for modeling very fast errors that are sometimes observed (e.g., Ratcliff & Smith, 2004), as well as the non-normality of RT distributions in long-RT tasks. The LBA also uses linear deterministic accumulation, similar to the models of Reddi and Carpenter (2000), Grice (1972) and Reeves et al. (2005); and yet each of these models also provides an inadequate account of incorrect responses.

The LBA is a greatly simplified instance of the dominant theoretical framework for models of choice RT for the past 50 years: *sequential sampling*. Beginning with Stone (1960), and continuing through to Ratcliff and Smith (2004), theoretical accounts for the variability in responses and RTs have assumed that a decision is made by the accumulation of “evidence” that varies randomly from moment to moment. A canonical example is illustrated in the left panel of Fig. 1. Suppose a choice is to be made between two competing alternatives—perhaps to classify the letter string “SIRF” as either a valid word, or as a non-word. The two possible responses (either “word” or “non-word”) are identified with separate evidence accumulators. These accumulators gather evidence for each response, increasing their amount of evidence with time. When the amount of evidence in either accumulator reaches a response threshold, the decision corresponding to that accumulator is produced, and the decision time is the amount of time taken to reach the response threshold.¹ Most importantly, the evidence accumulation process is stochastic—there are random moment-to-moment fluctuations in the amount of evidence supporting each response alternative. This randomness explains the variability in RTs (e.g.,

¹ We assume that RT is the sum of decision time plus a constant extra time, representing all the processing that does not involve decision making, such as time for perception and response production.

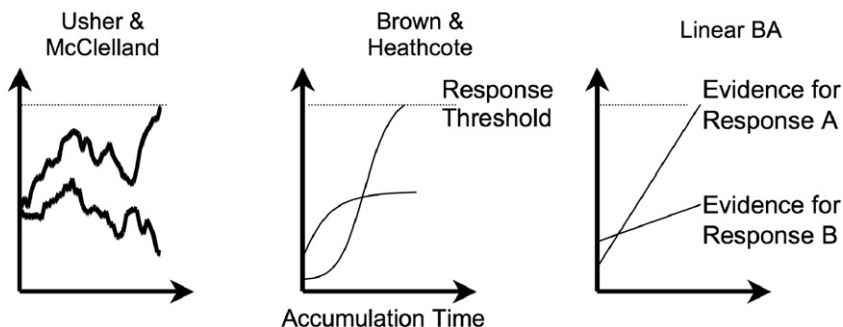


Fig. 1. A schematic illustration of Usher and McClelland’s (2001), left panel stochastic accumulator panel. Brown and Heathcote’s (2005a), center panel ballistic accumulator was simplified by omitting within-trial stochastic variation. The linear ballistic accumulator (right panel) is simpler still—omitting nonlinear passive decay and response competition processes.

different response times are observed even if an identical stimuli is repeated) and variability in the responses made (e.g., incorrect responses are also observed).

The model we have described is Usher and McClelland’s (2001) “leaky competing accumulator” (also known as the “mutual inhibition” model). As with all successful theories of choice RT, the model includes several extra components in addition to the evidence accumulation process. There have been four particularly important additions: two extra sources of random variability, and two nonlinear processes. The sources of extra random variability allow that the initial amounts of evidence in favor of each response (“start points”) and the average speed of evidence accumulation for each response (“drift rates”) to fluctuate from trial to trial. The nonlinear processes used by Usher and McClelland were passive decay of accumulated evidence, and response competition—evidence accumulating in favor of one response decreases the evidence for the other response.

Brown and Heathcote (2005a) proposed a simplification of Usher and McClelland’s (2001) model. In a break from 50 years of stochastic sequential sampling models, our “ballistic accumulator” omitted the within-trial randomness from the evidence accumulation process, as shown in the center panel of Fig. 1. We demonstrated that the ballistic accumulator accommodated all of the important empirical phenomena, using *only* the four processes described above (two trial-to-trial variabilities, and two nonlinearities) without any variability in the evidence accumulation itself. Here, we propose a further simplification: omitting the nonlinearities from the ballistic accumulator. This new model, which we call the linear ballistic accumulator (LBA), is illustrated in the right panel of Fig. 1. Evidence accumulates linearly for both responses, without moment-to-moment variability, continuing until the response threshold is reached for one response. Evidence accumulation for each response is also independent of evidence accumulating for other responses.

Linear and independent evidence accumulation is a rare assumption amongst models of choice RT, which usually include response competition explicitly (as in Usher & McClelland, 2001) or implicitly (as in single accumulator models, such as Ratcliff’s, 1978, diffusion model), or assume passive decay of accumulated evidence (Smith & Ratcliff, 2004). Even with its very basic architecture, the LBA model accounts for all the most important empirical phenomena, including RT distribution shape, speed–accuracy tradeoffs, and the relative speed of correct vs. incorrect responses. The simplicity of the LBA provides an

important advantage relative to all other models of choice RT: analytic solutions for the predicted distributions and probabilities. Similar solutions are available for some other models of choice RT, although these are not always easy to use. The unique advantage of the LBA model is that, unlike other choice RT models, the analytic solutions extend to choices between *any number* of response alternatives. In the following sections, we describe how the LBA model relates to previous attempts at theoretical simplification, and then describe its most important mathematical properties (with details in the [Appendix A](#)). Finally, we show that the LBA provides a good description of data from five previously published experiments.

2. Simpler models of RT

The increasing complexity of theories for choice RT has inspired several previous attempts at simplification, with reduced assumptions about variability and nonlinearity in evidence accumulation. We briefly describe five attempts, ending with the ballistic accumulator of [Brown and Heathcote \(2005a\)](#). We highlight the strengths and weakness of each model, and compare them to our proposed LBA model.

The most recent attempt at theoretical simplification is the EZ-diffusion model of [Wagenmakers, van der Maas, and Grasman \(2007\)](#). This model is extremely simple, with just one source of variability in evidence accumulation—within-trial randomness—and simple linear accumulation (although evidence for one response does count against the other). The EZ-diffusion model is even simpler than the LBA, but it is incomplete. Wagenmakers et al. proposed the EZ-diffusion as a *descriptive* rather than *process* model, with the aim of adequately describing data as simply as possible. The tradeoff in developing such a simple model was that it could not account for some of the empirical phenomena in choice RT, such as the relative speed of correct vs. incorrect responses. Since the EZ-diffusion does not purport to be a complete model of choice RT, these inadequacies may be overlooked so that the benefits of its simplicity and usability may be enjoyed. The LBA model, on the other hand, is proposed as a complete model of choice RT—we maintain that the model, while simple, can account for all of the important empirical phenomena. A further difference is that the EZ-diffusion model, like all random walk and diffusion models, is naturally restricted to binary choices, whereas the LBA model extends easily to choices between any number of alternatives.

Three other simplified theories of choice RT have been proposed: [Reddi and Carpenter's \(2000\)](#) LATER; [Grice's \(1972\)](#) variable criterion model; and [Reeves, Santhi, and DeCaro's \(2005\)](#) random-ray model. These models share a common structure, based on a standard random walk, but without within-trial variability in evidence accumulation process. The left panel of [Fig. 2](#) is a schematic illustration of a standard random-walk based model, such as [Ratcliff's \(1978, 1988, 2002\)](#) diffusion model. In random walk models, the choice under consideration is always binary, with the two possible responses represented by two response thresholds (one upper, one lower). There is just one evidence accumulation process, and it begins somewhere between these two thresholds. As (variable) evidence is accumulated, the process wanders about until it eventually reaches one of the two thresholds, triggering a decision. The simplification proposed by Grice, the LATER model, and the random-ray model was to omit the variability from the evidence accumulation process, as illustrated in the right panel of [Fig. 2](#). These models still predict

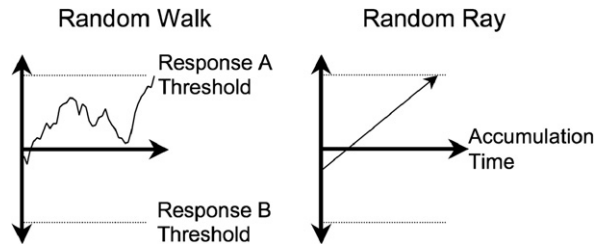


Fig. 2. Schematic illustration of a standard random walk model (left) and the random ray model (right) of Reeves et al. (2005). The random ray model is similar to other simplified models, such as Grice's (1972) and Reddi and Carpenter's (2000).

response variability because other sources of variability are included (in the drift rates, start points, or thresholds).

The similarity between the LBA and the models of Grice (1972), Reddi and Carpenter (2000), and Reeves et al. (2005) is striking: compare the right-hand panels of Figs. 1 and 2. This begs the question of why our approach should succeed where these models have failed. The answer lies in the particular failures of the earlier models to predict incorrect responses. Taking the random ray model as a canonical example, one must consider how incorrect responses arise. Like almost all diffusion models, the random ray model assumes a normal distribution for the speed of evidence accumulation (the “drift rate” or the slope of the arrow in Fig. 2). The normal distribution means that on some trials the drift rate will be negative, and the arrow will point *downwards*. This results in an incorrect response, when the evidence accumulation process reaches the lower threshold. The problem is that the predicted response times for these incorrect responses are *negatively* skewed, which is never observed in data. The negative skew is caused by projecting the tail of a normal distribution (i.e., those negative drift rates) onto the lower boundary. Grice's model makes equally problematic predictions about incorrect responses (see, e.g., Ratcliff, 2001). The original version of the LATER model is similarly problematic, as it fails to address incorrect responses at all. More recent developments of LATER (e.g., Leach & Carpenter, 2001) include the possibility of response variability, although these variants still fail to account for the detailed patterns found in incorrect response time distributions.²

The LBA model, on the other hand, makes accurate predictions for both correct *and* incorrect responses. The key difference is that the LBA model assumes *separate* evidence accumulation processes for each response, rather than using a single random walk. Separate accumulators allow the properties of the two responses (e.g., RT distributions) to be quite similar, as observed in data. The LBA model is a simplification of Brown and Heathcote's (2005a) ballistic accumulator, which was in turn a simplification of Usher and McClelland's (2001) model. Brown and Heathcote advanced the argument that their ballistic accumulator was an improvement over Usher and McClelland's model using Occam's Razor—they showed that the ballistic accumulator was simpler, but still accommodated

² Dzhafarov (1993) provides a related analysis. In particular, his Theorem 2.3.2 demonstrates that any reasonable family of response time distributions can be produced by an independent accumulator model, if one allows arbitrary distributions of drift rates and/or arbitrary nonlinear evidence accumulation functions. Dzhafarov's work aims at a different level to ours, providing a re-description of the data rather than positing a falsifiable model.

the data. Here, we extend this argument one step further, and demonstrate that the (even simpler) LBA model accommodates all the requisite empirical phenomena. Defining the “simplicity” of accumulator models is not easy, but a coarse approximation may be obtained by counting parameters. We note that, to fit RT distributions from k experimental conditions (for correct and incorrect responses) the LBA model typically requires $k + 4$ parameters, while competing models require at least this many, and occasionally more: Brown and Heathcote’s ballistic accumulator requires $k + 6$; Ratcliff’s diffusion requires $k + 5$ (or $+6$ depending on some extra components).

The LBA model enjoys one final advantage over other models of choice RT, as it has simple analytic solutions for choices between any number of different alternatives. The importance of such solutions should not be underestimated. One of the reasons for the success of (Ratcliff’s, 1978, 2002; Ratcliff & Smith, 2004) diffusion model was its well-developed analytic solutions for binary choices. The LBA model extends this advantage to choices between more than two alternatives, opening up new modeling possibilities.

3. The linear ballistic accumulator

The LBA model represents a choice between N alternatives ($N = 2, 3, \dots$) using N different evidence accumulators, one for each response. This is illustrated for the binary ($N = 2$) case in Fig. 3, which shows two evidence accumulators, one for “Response A” and one for “Response B”. Each evidence accumulator begins the decision trial with a starting amount of evidence (k) that increases at a speed given by the “drift rate” (d). Accumulation continues until a response threshold (b) is reached. The first accumulator to reach the threshold decides the overt response, and the time taken to reach the threshold decides the RT (plus some extra constant time for non-decision processes, t_0). Simple geometry dictates that the time taken for an accumulator to reach the threshold is $(b - k)/d$.

The LBA model explains the observed variability in data using two sources of between-trial randomness. The starting points for evidence accumulators are random values drawn from a uniform distribution on the interval $[0, A]$, and the drift rates are drawn from normal distributions with means v_1, v_2, \dots, v_N for the different response accumulators, and a common standard deviation s . All random values are drawn independently for each accu-

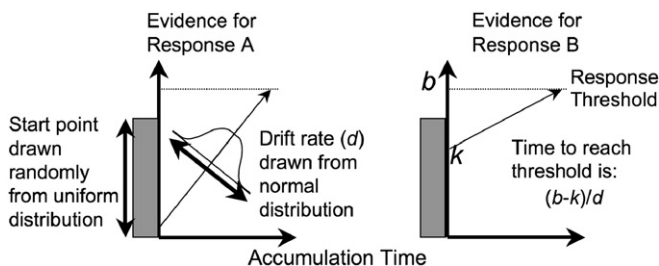


Fig. 3. A two-choice version of the LBA. Evidence for Response A is gathered in the left hand accumulator, and for Response B in the right hand accumulator. Starting values for the evidence accumulation processes (e.g., k) are drawn randomly and independently from identical uniform distributions on the interval $[0, A]$. The drift rate (the speed of evidence accumulation, e.g., d) is drawn independently for each response from normal distributions with standard deviation s . A response is triggered when the first accumulator reaches the threshold b .

mulator, and are independent across decision trials. We will refer to a uniform distribution on the interval (x,y) as $U[x,y]$ and a normal distribution with mean μ and standard deviation σ as $N(\mu,\sigma)$. The functions $\phi(\cdot|\mu,\sigma)$ and $\Phi(\cdot|\mu,\sigma)$ refer to the normal distribution's density and cumulative density functions, respectively. We will use lower case letters to refer to probability density functions (PDFs) and upper case letters for cumulative distribution functions (CDFs).

Consider for a moment just the i th accumulator. Since the starting value for the evidence accumulator is a random sample from a uniform distribution on $[0,A]$, the amount of evidence that needs to be accumulated to reach the threshold b is a sample from the uniform distribution $U[b-A,b]$ (we assume $b \geq A$). The drift rate is a random draw from $N(v_i,s)$, so the distribution function for the time taken for the i th accumulator to reach threshold is the given by the ratio of these two, which has CDF (at time $t > 0$):

$$F_i(t) = 1 + \frac{b-A-tv_i}{A} \Phi\left(\frac{b-A-tv_i}{ts}\right) - \frac{b-tv_i}{A} \Phi\left(\frac{b-tv_i}{ts}\right) + \frac{ts}{A} \phi\left(\frac{b-A-tv_i}{ts}\right) - \frac{ts}{A} \phi\left(\frac{b-tv_i}{ts}\right) \quad (1)$$

The associated PDF is:

$$f_i(t) = \frac{1}{A} \left[-v_i \Phi\left(\frac{b-A-tv_i}{ts}\right) + s \phi\left(\frac{b-A-tv_i}{ts}\right) + v_i \Phi\left(\frac{b-tv_i}{ts}\right) - s \phi\left(\frac{b-tv_i}{ts}\right) \right] \quad (2)$$

Most often, we are not interested in the time take for one particular evidence accumulator to reach threshold, but rather in the time taken for the first accumulator from all N to reach threshold. In particular, we want the *defective* distribution of response times for the i th accumulator. By a “defective distribution”, we mean a distribution that is not normalized to one, but instead to the probability of the response with which it's associated. Eqs. (1) and (2) allow the direct calculation of the defective PDF for the i th accumulator by:

$$\text{PDF}_i(t) = f_i(t) \prod_{j \neq i} (1 - F_j(t)) \quad (3)$$

The associated CDF can be evaluated by numerical integration of Eq. (3). The probability of making response i is given by the integral over the positive real line of Eq. (3) or equivalently by the associated CDF evaluated at $t \rightarrow \infty$.

Eqs. (1)–(3) are derived in detail in [Appendix A](#). An assumption of our derivations is that at least one of the accumulators has a finite and positive (first passage) time to achieve threshold. Since drift rates are sampled from normal distributions, there is some probability of a negative drift rate, namely $\Phi(-\frac{v_i}{s})$ for the i th accumulator. The probability of an undefined response time is the probability of *all* accumulators having negative drift rates: $\prod_{i=1}^N \Phi(-\frac{v_i}{s})$. With the parameters we estimate from data below, undefined response times are never predicted for more than 0.5% of trials in any of the five experiments.

It is conceivable that human subjects give “undefined responses” in at least that many cases: for example, [Ratcliff and Tuerlinckx \(2002\)](#) directly estimate a parameter (p_0) describing the proportion of “contaminants” at between 1% and 5%. Undefined response times could also be eliminated altogether by assuming a strictly positive distribution for

drift rates, such as the lognormal distribution (see, e.g., Hayes, Sutton, & Mewhort, submitted for publication). We chose the normal distribution for practical reasons, because it is both tractable and conventional. That is, the normal distribution allowed us to complete many of the differential and integral calculus operations in the Appendix A, which would have been troublesome with other distributions, and almost all other choice RT models have assumed normal distributions for drift rates.

There are two ways in which one may interpret the idea of “ballistic” accumulation. By analogy with Newtonian mechanics, one may imagine the growth of activation in each accumulator as a pre-defined process that cannot be stopped or changed once it has begun. We do not intend for our model to adhere to this strongest definition, but rather we use the term “ballistic” to refer to two critical elements. First, evidence accumulation is non-stochastic, which is the key difference between the LBA and most other models of choice RT. Second, evidence accumulation is ballistic in the sense that the future trajectory of each evidence accumulator is completely determined once its initial conditions are known—corresponding to the notion of ballistics in Newtonian mechanics. This second constraint may require relaxation under certain unusual experiments, for example when the information content of a stimulus is dramatically altered during the course of a decision. Brown and Heathcote (2005b) used stimuli that changed from favoring one response to another, part-way during a trial (see also Usher & McClelland, 2001 & Vickers, 1995). Under these conditions, it is possible that the assumptions of the LBA may need to be relaxed so that the drift rate changes when the stimulus category is changed. This assumption contrasts with our usual assumption of an unchanging rate of evidence accumulation, even for stimuli that include random noise that varies from moment to moment. More experimentation with dynamically changing stimuli is required to identify those situations in which our simplest assumptions will not hold.

4. Predicting fast and slow errors

The relative speeds of correct and incorrect response times have become important for theories of choice RT (see, e.g., Luce, 1986; Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004). When choices are easy, and participants are told to respond quickly, incorrect responses are faster than correct responses, but when choices are difficult, and accuracy is emphasized, incorrect responses are slower than correct responses. As in other complete models of RT, the LBA accommodates this pattern via the interplay between start point variability (uniform samples from $[0, A]$) and drift rate variability (normally distributed samples with standard deviation s). When response speed is emphasized, the response threshold (b) is fixed at or near the top of the start point distribution (A). This means that a very fast response will occur whenever a start point is sampled close to the top of its distribution. With such a high starting point there is little integration time taken to reach the threshold, so responses are about equally divided between correct and incorrect responses, depending on which response accumulator sampled the high start point. Conversely, when start points are sampled from lower in the distribution, responses are slower, and integration time is longer. The longer integration time allows the correct response (with the higher drift rate) to overtake the incorrect response, so these slow responses are mostly correct. All of this means that when response speed is emphasized, incorrect responses are mostly due to start point samples from the upper end of the distribution, and are relatively fast.

When instructions stress response accuracy rather than response speed, the response threshold is set well above the top of the start point distribution. This makes integration time quite long for all decisions, so the effect of variability in start point is integrated out, and errors are caused mostly by drift rate variability. The larger mean drift rate for correct responses ensures that threshold-crossing times are faster for correct responses than incorrect responses. The smaller drift rate for the incorrect response accumulator means it reaches the response threshold more slowly, causing incorrect responses to be slower than correct responses.

5. Benchmark tests

We test the LBA model against several sets of benchmarks, using data from five previously published experiments. We first analyze three lexical decision experiments, which illustrate the model's basic predictions and show that the LBA accounts for response accuracy and for the shape and speed of RT distributions for both correct and incorrect choices. We then extend the model analyses to a brightness discrimination experiment, using unaveraged (individual participant) data. Those data exhibit a complex interaction of fast and slow errors, which has proven challenging for other models of choice RT. We then examine the LBA model's account of the shape of speed–accuracy tradeoff (SAT) curves. Finally, we extend our analyses to multiple choice data, using an absolute identification task with ten alternatives. These data allow the LBA's predictions for RT distributions and response probabilities to be tested in a multiple choice paradigm, and using individual participant data.

For a choice between two alternatives, the LBA model appears to have six parameters: the response threshold (b), the upper end of the start point distribution (A), drift rates for each response (v_1 and v_2), a standard deviation for drift rate samples (s), and the amount of time taken for non-decision components of response time (t_0). However, as with other accumulator models, one parameter must be eliminated as a “scaling parameter”. The scaling property of the model is such that, if all parameters (except t_0) were doubled, the model's predictions would be unchanged. To remove this redundancy, one parameter must be fixed arbitrarily, but there are many ways to do this. For example, [Usher and McClelland \(2001\)](#) chose to fix the sum of the drift rates arbitrarily at one ($v_1 + v_2 = 1$), and we follow this convention in all our two-choice analyses. When considering simultaneous fits to more than one drift-rate condition, other ways of fixing the scaling problem are possible. For example, rather than assuming that the sum of the drift rates is constant, one might assume that the drift rate for the incorrect response is a constant value (say, 1), or that the ratio of the drift rates for the correct and incorrect responses is a constant value. As with all accumulator models, these assumptions will lead to different predictions about error responses, and imply different psychological interpretations of the model.

5.1. Lexical decision experiments

Lexical decision is the task of classifying letter strings as either valid words (“SURF”) or non-words (“SIRF”). Data from lexical decisions exhibit reliable speed–accuracy tradeoffs and effects of word type (e.g., [Brown & Steyvers, 2005](#); [Ratcliff, Gomez, & McKoon, 2004](#)). Ratcliff et al. conducted a series of nine experiments manipulating stimulus features, such as the ease-of-pronunciation for non-words (e.g., “XJQF” is hard to pronounce, but

“SULF” is not) and the natural occurrence frequency for words (e.g., “THEN” appears often in written texts, but “PANE” does not). They observed reliable changes in response accuracy, and the shapes of RT distributions for correct and incorrect responses. We analyze all experiments for which they reported complete results for incorrect and correct RT distributions: Experiments 3–5. Ratcliff et al. modeled the data using their diffusion model, which required $k + 6$ free parameters when applied to k different experimental conditions. The LBA accounts for the same data at least as well using $k + 4$ parameters (compare our model fits with Tables 5 and 12 from Ratcliff et al.).

In their Experiment 3, Ratcliff et al. (2004) used three classes of letter strings: pronounceable pseudowords (“CLARP”), high-frequency words (“CHAIR”) and low-frequency words (“CINCH”). Fig. 4 shows the defective RT distributions from these three conditions, using quantile estimates. Each panel of Fig. 4 shows five quantile estimates from each of the correct response and incorrect response time distributions. The quantile estimates (open diamond symbols) give the RT below which 10%, 30%, 50%, 70% and 90% of the data occur. The quantile estimates can be interpreted as five points on the defective cumulative distribution function for the data. The quantile estimates also carry information on response accuracy because we use *defective* cumulative distributions. This means that the relative heights of the correct and incorrect distributions show the proportion of correct vs. incorrect responses. For example, the 90% quantile for correct responses is graphed at a probability of .84, indicating that 84% of the data fall below the corresponding RT (0.865 s). This is the 90% quantile estimate because 90% of the *correct responses*, which are 93% of the data (i.e., 7% of all responses were errors), fall below 0.865 s.

The solid lines with small symbols show the predicted quantile estimates from the LBA model. The model’s predictions match the data very well; the worst error is a 0.053-s overprediction for the 90% quantile estimate from incorrect responses to low frequency words. This misfit does not overly trouble us because the data in question (the slowest 10% of incorrect responses) represent only 0.7% of data from the experiment.

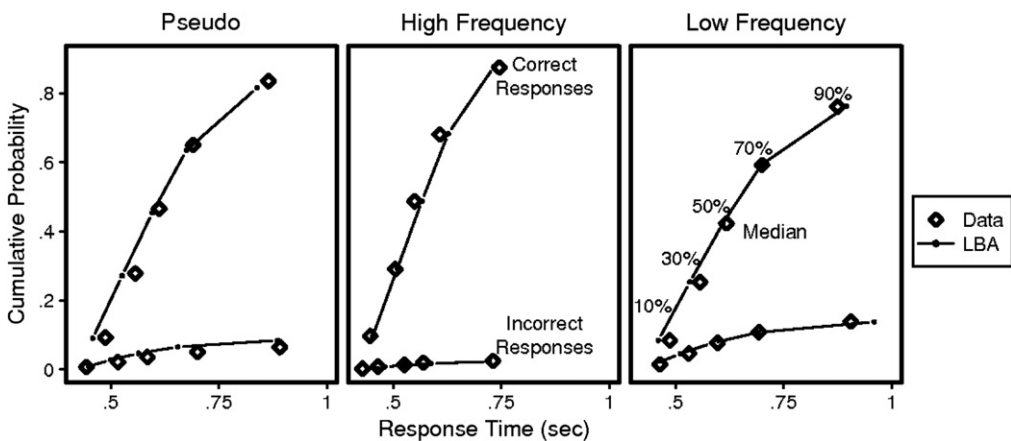


Fig. 4. Participant-average quantile estimates (symbols) from Ratcliff et al.’s (2004) Experiment 3, and predicted quantiles from the LBA model (lines). The three panels represent data from the three stimulus classes: pseudowords, high-frequency words and low frequency words. In each panel, the two lines show .1-, .3-, .5-, .7- and .9-quantile estimates for correct responses (upper lines) and incorrect responses (lower lines).

Table 1
Parameters for the LBA model for fits to Experiments 3, 4 and 5 from Ratcliff et al. (2004)

	t_0	A	b	s	v_R	v_P	v_H	v_L	v_V
Expt. 3	0.425	380	380	0.428		1.04	1.32	0.916	
Expt. 4	0.392	377	377	0.511	1.34		1.37	1.22	
Expt. 5	0.438	460	460	0.422		1.07	1.36	1.00	0.802

Note: v indicates mean drift rate for: random letter strings (R); pseudowords (P); high-frequency words (H); low-frequency words (L) and very low-frequency words (V).

We fit the model to the data assuming that mean drift rate (v) was different for the three conditions, but that all other parameters (t_0 , A , b , and s) were equal across the three conditions. The mean drift rate for the response accumulator corresponding to incorrect responses was (arbitrarily) set at one minus that for correct responses. Taking the quantile estimates for correct and error RT distributions given in Ratcliff et al.'s (2004) tables, we compared the goodness of fit between CDFs predicted by the LBA model and the data using the quantile maximum product method of (Heathcote, Brown, & Mewhort, 2002, see also Heathcote & Brown, 2004). We then used a standard simplex algorithm to maximize this statistic over the parameters.

The free parameter values are reported in Table 1—note that the estimate of b (the decision threshold) is equal to A (the upper limit of the start point distribution) in all three experiments. This reflects the faster incorrect than correct responses for these data, which occurs when starting point variability is relatively large. There is very good agreement between common parameter estimates for the three experiments (e.g., the three drift rate estimates for high frequency words, v_H , agree to within 4%) supporting the notion that the LBA parameters reflect consistent underlying aspects of decision-making. Finally, we note that the problem of undetermined response times, due to negative drift rates, is very small: the probability is less than 0.4%.

Figs. 5 and 6 show data from Ratcliff et al.'s (2004) Experiments 4 and 5, along with LBA model fits. The format of the figures is the same as in Fig. 4, except for the stimulus condition that was contrasted with low- and high-frequency words. Experiment 3 (above) used pronounceable pseudowords, while Experiment 4 (in Fig. 5) used random letter strings. Experiment 5 (in Fig. 6) returned to pronounceable pseudowords, and included a new category of very low frequency words. In each case, the LBA model fits the data allowing only drift rate to change between the stimulus conditions.

The LBA model fits the data very well in all three experiments (compare our fits with diffusion model fits from Tables 3 and 12 of Ratcliff et al., 2004). In each data set, there were several important phenomena that the model captured. First, the incorrect RT quantiles (lower line in each plot) were further to the left than the corresponding correct RT quantiles, indicating that incorrect responses were faster than correct responses. The LBA model accommodates this pattern, predicting “fast errors” by virtue of setting the response threshold at the upper limit of the starting point distribution ($b = A$), rather than far above it ($b \gg A$). All three experiments also showed changes in accuracy across the stimulus conditions, as exemplified by changes in the overall heights of the lines across panels in Fig. 6. The LBA model fits this pattern via changes in the drift rate, with larger drift rates leading to higher accuracy, and hence higher defective CDFs for correct responses, and lower defective CDFs for incorrect responses. Finally, the LBA model

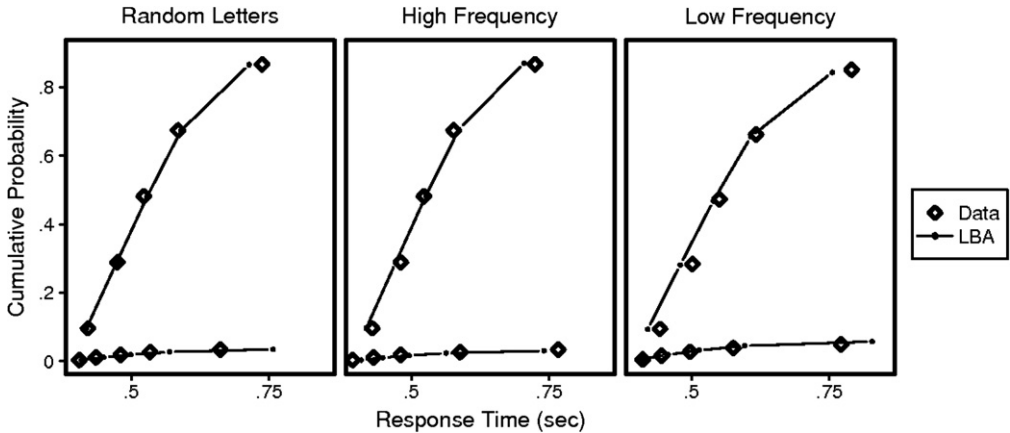


Fig. 5. Quantile estimates (symbols) from Ratcliff et al.'s (2004) Experiment 4, and predicted quantiles from the LBA model (lines). The three panels represent data from the three stimulus classes: random letter strings, high-frequency words and low-frequency words.

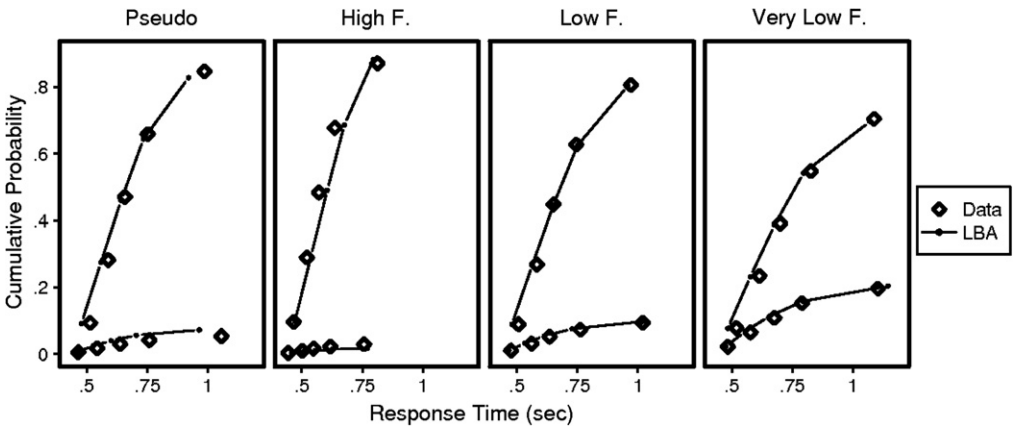


Fig. 6. Quantile estimates (symbols) from Ratcliff et al.'s (2004) Experiment 5, and predicted quantiles from the LBA model (lines). The four panels represent data from the four stimulus classes: pronounceable pseudowords, high-frequency words, low-frequency words, and very low-frequency words.

captures the shape of the RT distributions, with a sharp leading edge and long right tail (e.g., the CDF flattens out between the .7- and .9-quantile estimates for every distribution). In all three experiments, the worst misfit between LBA predictions and data occurs for the slowest quantile estimate for incorrect responses. The LBA over-predicts this value by as much as 0.09 s (random letter strings in Experiment 4) and under-predicts by just as much (for pronounceable pseudowords in Experiment 5). Two things suggest that these misfits are due to random sampling variability in data: they are not systematic, being sometimes over-predictions and sometimes under-predictions; and the quantile estimates in question are always the most variable in the data set, arising from the slowest 10% of incorrect responses, less than 1% of data in each stimulus condition.

5.2. Speed and accuracy emphasis in brightness discrimination

Ratcliff and Rouder (1998) performed a simple brightness discrimination experiment that has since proven seminal in the development of two-choice RT models. They asked three participants (J.F., K.R. and N.H.) to classify patches of pixels as either “bright” or “dark”, about 8000 times each. The brightness of the pixel patch was manipulated in 33 levels, from very dark (which was almost always classified as “dark”) through to midway (which was classified “dark” and “bright” about equally often) to very bright (almost always classified as “bright”). They also manipulated the urgency of responding, by changing the relative emphasis placed on accuracy vs. speed in the instructions and feedback given to the participants. When participants were given “speed-emphasis” instructions, they were told not to be overly concerned about errors, and whenever a response took longer than 0.55 s, the feedback “TOO SLOW” was given. Under “accuracy-emphasis”, participants were told not to make mistakes, if they could be avoided, and when an error was made to a relatively easy-to-classify stimulus, the feedback “BAD ERROR” was given.

Ratcliff and Rouder’s (1998) data have proven very testing for models of two-choice RT. They display complex interactions of mean RT with response probability and speed-vs. accuracy-emphasis. These interactions are impossible for some models of choice RT to accommodate, leading some to suggest those models be rejected (e.g., Ratcliff & Smith, 2004). Fig. 7 displays Ratcliff and Rouder’s data, using a separate panel for each of the three subjects (compare with Ratcliff and Rouder’s Fig. 4).

Fig. 7 graphs mean RT for each of the 33 stimulus conditions (symbols) separately for accuracy emphasis and speed emphasis instructions (upper and lower parts of each plot). For each condition, data are separated into “correct” and “incorrect” on the basis of the modal response, and mean RT is graphed separately for correct and incorrect responses, against the probability of the response type (i.e., a “latency-probability” or LP plot, see

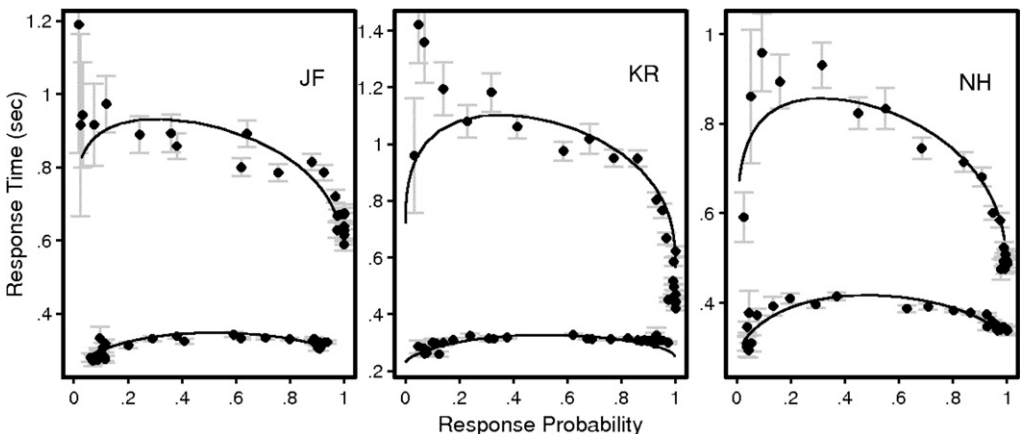


Fig. 7. Mean RT (symbols) and predicted mean RT from the LBA model (lines) for three subjects from Ratcliff and Rouder (1998). The upper and lower lines are for accuracy and speed emphasis conditions, respectively. Within each condition, there are 33 separate points—one for each level of stimulus brightness. Following Ratcliff and Rouder, we collapsed “bright” and “dark” responses, so that the left side of each plot represents correct responses to very easy-to-classify stimuli, and the right side of each plot represents (very rare) incorrect responses to the same stimuli. The center of each plot shows data from difficult stimuli, which were nearly equally often classified as “dark” or “bright”. Bars indicate ± 1 standard error.

Audley & Pike, 1965 or Ratcliff & Rouder, 1998). An example may make this clearer; consider just the data from the very easiest stimulus class (either very dark, or very bright), under speed emphasis instructions, for participant JF. These stimuli were presented 44 times in total, and JF classified them correctly 40 times, hence the probability of a correct response was 91% and the probability of an error was 9%. The mean RT for JF's correct responses was 0.331 s, and for the incorrect responses the mean RT was 0.265 s. These data are represented in Fig. 7 as two points, in JF's panel, on the lower line—for speed emphasis data. One point corresponds to correct responses and is placed at $\{x = .91, y = 0.331\}$ and the other represents incorrect responses, at $\{x = .09, y = 0.265\}$. All other conditions are similarly represented, so that each stimulus condition corresponds to two points on the LP plot: one for correct responses, at $x = p$ and one for incorrect responses, at $x = 1 - p$.

The data show some phenomena that have proven very challenging for models of two-choice RT. First, there is a very large difference in RT between the speed- and accuracy emphasis conditions: mean RT more than doubles in the accuracy condition. This large speed difference is accompanied by only a small difference in accuracy between the two conditions, with only about 6% between the most accurate condition in speed emphasis vs. accuracy emphasis. Even more challenging is the relative speed of correct vs. incorrect responses. Under accuracy emphasis, for all stimulus conditions, error responses are slower than correct responses. This is evident when comparing correct responses from any point on the right-hand half of the graph (say, $x = p$) with the corresponding incorrect responses, at $x = 1 - p$. However, when participants were put under speed emphasis, the situation reversed; mean RT for incorrect responses was *faster* than mean RT for correct responses. This crossover pattern has also been observed by others, see Luce (1986) for a summary, and has proven difficult to model.

The fits of the LBA model (solid lines in Fig. 7) show that it predicts the change from slow errors under accuracy emphasis to fast errors under speed emphasis, for all three subjects. We constrained LBA parameter estimates following Ratcliff and Rouder's (1998) diffusion model analysis of their data, to keep the results comparable. We allowed a separate drift-rate estimate for each of the 33 stimulus conditions, but constrained these estimates to be equal across speed-emphasis and accuracy-emphasis conditions.³ We kept the non-decision component of RT (t_0), and the variability in drift rates (s) constant across all conditions. To model the difference between speed-emphasis and accuracy-emphasis instructions, Ratcliff and Rouder allowed changes in boundary separation (equivalent to the response threshold, b) and start point variability (A). Ratcliff and Rouder constrained another free parameter by assuming that start point variability was a constant proportion of boundary separation (i.e., they fixed the ratio $b:A$) resulting in five free parameters for their model, in addition to a drift rate estimate for each condition. We modeled the speed-emphasis condition by fixing the response threshold at its lowest value ($b = A$), consistent with our analysis of the lexical decision data above. We estimated a separate (larger) response threshold for the accuracy emphasis condition. Like Ratcliff and Rouder, we allowed different amounts of start point variability (A) for speed and accuracy emphasis conditions, but unlike Ratcliff and Rouder, we did not constrain these to be constant pro-

³ We further simplified the fitting by estimating only four drift-rate parameters, and interpolating the full set of 33 from these four, using spline functions.

Table 2
Parameters for Ratcliff and Rouder (1998)

Participant	Instructions	t_0	s	A	b
J.F.	Speed	.208	.285		181
	Accuracy			527	673
K.R.	Speed	.188	.290		177
	Accuracy			750	900
N.H.	Speed	.233	.338		219
	Accuracy			475	572

portions of the response threshold.⁴ These considerations resulted in the LBA model using exactly the same number of parameters as the diffusion model—five free parameters, plus one drift rate per stimulus condition. The estimated parameters for the LBA model fits are shown in Table 2. Once again, the problem of non-terminating RTs caused by negative drift rate samples is negligible. The predicted proportion of non-terminating responses for the three subjects was: J.F. = 0.05%, K.R. = 0.04% and N.H. = 0.23%. Once again, we estimated parameter values using the QMP method of Heathcote et al. (2002).

5.3. Response deadlines and speed–accuracy tradeoff

An important aspect of speeded choice is the ability to trade speed for accuracy. We have already addressed this in our fits to Ratcliff and Rouder's (1998) data, where two different speed–accuracy settings were implemented using feedback and instructions for participants. In *response deadline* paradigms, the speed–accuracy tradeoff is investigated quantitatively, via the speed–accuracy tradeoff (SAT) function relating the accuracy of decisions to the time taken to make them. Previous research (e.g., Luce, 1986; McElree & Doshier, 1989; Ratcliff & Rouder, 2000; Usher & McClelland, 2001) has found that the shape of the SAT curve is approximately shifted-exponential, so it is important to determine whether the predictions from the LBA match this benchmark. To test LBA we first require the distribution (across repeated trials) of activation values at any time $t > 0$ for the i th response accumulator. Given a start point k drawn from $U[0, A]$ and a drift rate d drawn from $N(v_i, s)$ the activation at time t is $k + td$, and hence the distribution function, W_i , and probability density function, w_i , for this quantity are (note that z represents activation, and t represents integration time):

$$W_i(z, t) = \frac{1}{A} \left(\frac{ts}{\sqrt{2\pi}} \left(e^{-\frac{(\frac{z}{t} - v_i)^2}{2s^2}} - e^{-\frac{(\frac{z-A}{t} - v_i)^2}{2s^2}} \right) + (A - z + v_it) \Phi \left(\frac{z - A}{t} \middle| v_i, s \right) \right. \\ \left. + (z - v_it) \Phi \left(\frac{z}{t} \middle| v_i, s \right) \right) \quad (4)$$

$$w_i(z, t) = \frac{1}{A} \left(\Phi \left(\frac{z}{t} \middle| v_i, s \right) - \Phi \left(\frac{z - A}{t} \middle| v_i, s \right) \right) \quad (5)$$

⁴ Some readers may feel that start point variability *should* be constrained to be a constant proportion of the response threshold, as in Ratcliff and Rouder's (1998) analysis. We assessed the fit of the LBA model under that constraint, and found it was marginally poorer (but also one parameter simpler—contact the authors for a graph and details). We chose to show the unconstrained version, because we could see no a priori reason why start point variability should be allowed a different *value* between the speed and accuracy conditions, but should not be allowed a different *proportion*.

To make predictions about accuracy in a deadline (signal-to-respond) paradigm, we assume the model operates without response thresholds. When a deadline arises and a response is required, the model returns whichever response corresponds to the accumulator with the greatest activation. The probability that the i th accumulator has the greatest activation at time $t > 0$ is given by combining Eqs. (4) and (5):

$$p_i(t) = \int_{-\infty}^{\infty} w_i(x, t) \prod_{j \neq i} W_j(x, t) dx \quad (6)$$

The speed–accuracy tradeoff function is given by Eq. (6) and evaluated for whichever response (i) is correct. To compare with benchmark results from binary choice paradigms, the response probability must be transformed to sensitivity. Assuming unbiased responding, this transformation is:

$$d'(z) = 2 \cdot \Phi^{-1}(p_i(z)) \quad (7)$$

Using Eqs. (6) and (7), and parameter settings taken from the fits to Ratcliff and Rouder's (1998) data, we calculated predicted SAT functions for three participants, for response deadlines from 0 to 1.2 s, in steps of 20 ms (shown by the open circles in Fig. 8). The solid lines in Fig. 8 show the best fitting shifted-exponential functions for each participant's SAT curve. The agreement between the predicted SAT functions from the LBA model and the shifted exponential form is very good. There is a difference at the highest performance levels, where the LBA model predicts a slower approach to asymptote than does the shifted exponential function, but the discrepancy is so small it is imperceptible in the graph.

5.4. Multiple choices

A strength of the LBA model is its ability to account for choices between more than two alternatives, but so far we have used only binary choice tasks and illustrations. The pre-

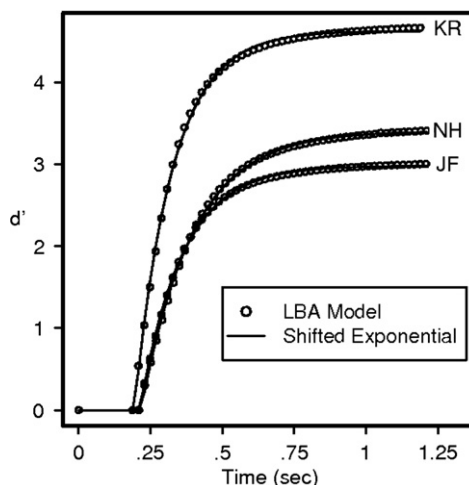


Fig. 8. Predicted SAT functions from the LBA model (circles) and best-fitting shifted exponential functions (lines). Parameters were taken from fits to the three participants of Ratcliff and Rouder (1998); see Table 2.

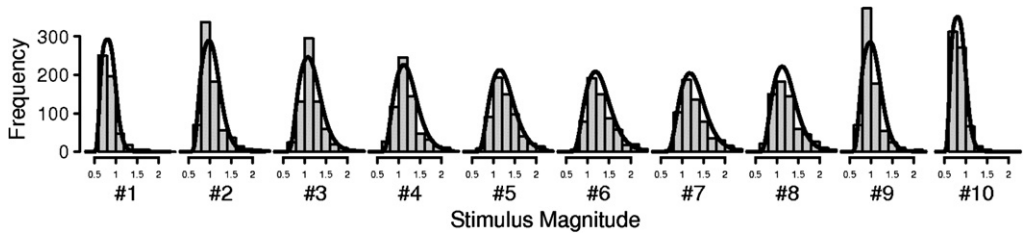


Fig. 9. Correct response time histograms (gray bars) and predicted distributions from the LBA model (black lines) for data from Lacouture and Marley's (2004) Experiment 2. The 10 plots correspond to the 10 stimuli, from the smallest line length (#1) to the largest (#10). All panels use the same axes (the abscissa units are seconds).

dictions of the LBA model naturally extend to more than two choices without any increase in the difficulty of the mathematics. Eq. (3) simply incorporates all the response choices: one as a PDF (f_i) and the others as CDFs (the products of $1 - F_i$ over $i \neq j$). We test the multiple-choice predictions of the model using data from Lacouture and Marley (2004), who ran a multiple choice experiment with 10 stimuli, and 10 responses. They used lines as stimuli, with the smallest line labeled “#1” and the longest labeled “#10”. In their Experiment 2, they had a single participant perform around 7500 absolute identification trials: on each trial, one of the 10 lines was shown to the participant, who tried to produce the corresponding response label.⁵ The data showed all of the standard absolute identification phenomena in both choice probabilities and response times. For example, responses to end stimuli (#1 and #10) were faster and more accurate than responses to middle stimuli.

When modeling data with ten stimuli and ten responses, the LBA model could use 100 free parameters as drift rates, one for each stimulus–response combination. However, some simplifying assumptions cut this number down to just 11 free drift-rate parameters:

1. The values of most parameters were fixed across all 10 stimuli ($t_0 = .27$ s, $A = 705$, $b = 1335$, and $s = .251$).
2. When stimulus i is presented, the drift rate for the correct response is a free parameter, v_i , $i = 1 \dots 10$.
3. The drift rate for all other (nine) incorrect responses is a fixed fraction of the drift rate for the correct response, say $v_i \times g$. (We estimated $g = .225$.)

These assumptions were made to fit the data in the simplest way possible. They are essentially descriptive (as were our parameters for the binary choice data above). A deeper approach, but one that is outside the scope of this article, would articulate a theory of absolute identification that *produced* drift-rate parameters for the 10 stimuli, rather than simply estimating them. For such an approach, see Lacouture and Marley (2004) or Brown, Marley, Donkin, and Heathcote (in press). Fig. 9 shows the response time histograms for correct responses along with corresponding predictions from the LBA model.

The LBA model successfully captures the shape of the RT distributions, and even the way this shape changes across the stimuli. For example, stimuli on the ends of the range

⁵ Half of the absolute identification trials used manual (keyboard) responses, and half used a voice key. There was almost no difference in the two data sets, so we combined them.

have more peaked distributions, and faster RTs on average, than those in the center. The histograms and density functions used in Fig. 9 are unconventional in the RT literature, which has historically preferred quantiles and cumulative distributions. There is a good reason for this preference—it is quite difficult to gauge model misfit in density functions, especially in the tails of the distributions. To provide a more stringent test, Fig. 10 graphs the same data and model fits using RT quantiles (left panel) and response accuracy (right panel). The quantile analysis is the same as used for the lexical decision data—for each stimulus (#1...#10) we calculated the RT below which 10%, 30%, 50%, 70% and 90% of the data fell. These points are plotted vertically above each of the 1...10 stimulus labels on the x -axis, using different symbols (see legends).

The LBA model accommodates aspects of the data that have proven very challenging for other models. For example, the fastest RTs (the leading edge of the RT distribution, shown by the 10% quantiles in the left panel) changes by about 0.2 s across the stimulus range. Other models have had difficulty with related changes, in binary choice situations (e.g., Ratcliff et al., 2004). The LBA model simultaneously captures the changes in RT distribution and response accuracy, with only one free parameter (drift rate) changing across the stimulus magnitudes. The most notable model failures are for stimuli #1 and #10, where the data are less accurate than the model predicts, and for stimulus #8, where the data include many very slow responses.

In our analyses so far, we have demonstrated that the LBA model has complete theoretical solutions for choices between any number of response alternatives, which is an advantage of the LBA over other models of choice RT. We have separately evaluated the performance of the LBA model against binary ($N = 2$) choice data, and against multiple choice data ($N = 10$). This leaves unanswered the question of whether the LBA can capture empirical regularities observed when data are collected across many different choice set sizes (N). The most notable aspect of such data is Hick's Law (1952), which,

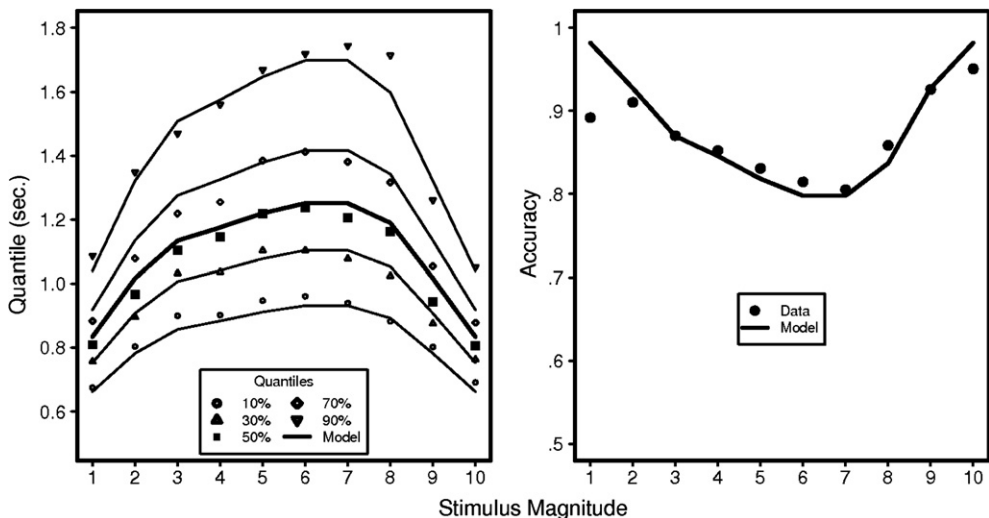


Fig. 10. Response time quantiles (left panel) and response accuracy (right panel) vs. stimulus magnitude. Symbols show data from Lacouture and Marley's (2004) Experiment 2, and solid lines show predictions from the LBA model.

in its simplest form, states that the average time taken to choose between N choices is proportional to $\log_2 N$. Naively, the LBA model fails to accommodate this law because, if no parameters are changed, increasing the number of accumulators leads to larger error rates and *faster* RTs. This prediction arises because predicted RT is the minimum finishing time from a set of N independent random variables. Naturally, the minimum sample from N independent distributions decreases as N increases.

Parameter adjustments with set size will allow most accumulator models, including the LBA, to accurately fit Hick's Law, as demonstrated by Usher et al. (2002). They found that increasing response thresholds in an approximately exponential manner with choice set size allowed the model to predict Hick's Law, and keep error rates approximately constant; a similar assumption in the LBA model has similar effects. Other assumptions about changes in parameter with set size are also plausible, and any of these may accommodate Hick's Law. For example, it may be imagined that the decision mechanism has a fixed "total power", hence the sum of the drift rates to all accumulators should be fixed. Alternatively, a decision mechanism may have a fixed sensitivity, so that the ratio of the maximum drift rate to the sum of the other drift rates should be fixed as N grows. The conclusion we draw is that predicting Hick's Law is a complicated task, too large to be tackled here. Instead, we note that the LBA model is certainly not in conflict with Hick's Law, and will accommodate it under appropriate parameter settings.

6. Limitations of the model

6.1. Neurophysiological considerations

The LBA is a simplification of Usher and McClelland's (2001) leaky competitive accumulator. Usher and McClelland's model, along with other sequential sampling models, has received support from arguments of neural plausibility. The noisy accumulation of evidence in these models can be likened to changes in the firing rates or membrane potentials of neurons. Indeed, direct recordings from neurons involved in perceptual decisions have been modeled using noisy sequential sampling (see, e.g., Gold & Shadlen, 2001; Mazurek, Roitman, Ditterich, & Shadlen, 2003; Ratcliff, Cherian, & Segraves, 2003; Roitman & Shadlen, 2002; Schall, 2001; Smith & Ratcliff, 2004). The LBA may be criticized on the basis that it omits the within-trial variability that is a basic feature of neural activity, and hence is "neurally implausible". We believe that there are a number of reasons not to weigh such arguments too heavily; we outline just two here.

First, it is a basic truth that all models (of choice RT and other phenomena) are approximations at some level. All models include simplifications for tractability. These simplifications are included for convenience, and are not intended as statements about the underlying nature of the process at hand. All models of choice RT interpose many layers of approximation between neurophysiological reality and theoretical description: ion flows, membrane potentials, action potentials; all of these are approximated away, and the LBA model adds just one further step. These approximations should not be perceived as inherently harmful to scientific progress. In physics, for example, the ideal gas law and Newtonian motion are both extremely useful models of reality. Their power and utility persist even though both include well-known and demonstrably false simplifications. Indeed, the utility of these laws is increased by these approximations, because the resulting simplicity allows wider application. More modestly, we hope that the simplicity and

tractability of the LBA model will allow its application to a wider range of choice RT paradigms than may otherwise have been possible.

Second, some of the leading models of choice phenomena from the neurophysiological literature omit within-trial variability, just as in the LBA. Most strikingly, Reddi and Carpenter's (2000) LATER model is like the LBA model, except that it has only one response accumulator and does not consider between-trial variability in the starting points of evidence accumulation. These limitations mean that the LATER model provides an incomplete account of behavioral data, particularly of error responses. More recent developments of the LATER model have improved its account of error responses somewhat, by adding a second accumulator (see Carpenter, 2004; Hanes & Carpenter, 1999; Leach & Carpenter, 2001; Reddi, 2001; Reddi, Asrress, & Carpenter, 2003). It seems to us that if the neurophysiological community accept models that omit within-trial variability, the psychological community should be careful about ruling out the LBA on grounds of "neural plausibility".

6.2. Specialised experimental paradigms

There are some experimental paradigms that have been used to investigate those processes that we have specifically omitted in the LBA. For example, neurophysiologists often employ random dot kinematograms, which include moment-to-moment random fluctuations. These stimuli naturally suggest moment-to-moment variability the evidence accumulation process, which is one of those processes included in many other decision making models (e.g., Ratcliff, 1978; Usher & McClelland, 2001) but not in the LBA. While it is possible that random dot stimuli will implicate moment-to-moment variability in decision models, we note that this is not a foregone conclusion. It is quite plausible that early visual processing smooths moment-to-moment fluctuations and provides a downstream decision process with a constant input, just as assumed in the LBA. Indeed, this is exactly how Reddi and Carpenter (2000, see also Reddi, 2001) model the neural processes underlying decisions using random dot kinematograms.

The other processes we have omitted from the LBA model that are often included in other decision-making models are the nonlinear aspects of passive decay (leakage) and response competition (lateral inhibition). Experimental paradigms have been developed to directly investigate these processes, by manipulating the timing of information for and against each response. For example, (Vickers, 1995, see also Pietsch & Vickers, 1997 & Brown & Heathcote, 2005b) presented a rapid sequence of two stimuli, in which one stimulus occurred more often than the other, and found that some participants were most influenced by stimuli occurring near the end of the sequence, while others were most influenced by those at the beginning of the sequence. These results can be interpreted in terms of passive decay and lateral inhibition: if accumulated evidence tends to decay with time, information presented earlier in a sequence will be forgotten; conversely, if response competition is strong, early evidence favoring one response will not be overcome by later evidence. Usher and McClelland (2001) used an experiment of exactly this sort to support the inclusion of leakage and competition components in their model. The LBA, on the other hand, assumes linear integration without leakage or competition, and so it may not provide a complete account of experiments that manipulate the timing of evidence. However, it is possible to provide an account of these experiments within an accumulator framework *without* appealing to leakage and competition, so the paradigm may not prove

as diagnostic as it seems. The key finding is that some participants are relatively uninfluenced by the start of the stimulus sequence, and others are relatively uninfluenced by the end of the stimulus sequence. A simple explanation of these effects can be given in terms of attention: some participants fail to attend to the start of the sequence, while others fail to attend to the end. There are plausible and well-documented mechanisms by which accumulator models can exhibit both kinds of attention failures. Laming (1968) discusses situations in which evidence accumulation does not always begin exactly at the stimulus onset, and Ratcliff (1988) developed models in which evidence accumulation stops before an overt response is given, because an internal response threshold is reached before the stimulus is extinguished.

7. Conclusions

We have developed the linear ballistic accumulator (LBA) model as a complete model of choice RT data, for decisions with any number of alternatives, $N = 1, 2, 3, \dots$. The model is the only one of its kind to have simple analytic solutions for RT distributions in the $N > 2$ cases. We have demonstrated that the model accounts for the distribution of RTs for correct and incorrect responses using data from five previously published experiments. The LBA model accommodates complex patterns in the relative speed of correct vs. incorrect responses, and the shape of the speed–accuracy tradeoff function.

Currently, decision making theorists are most concerned with the $N = 2$ case (binary choices). In this paradigm, there are many theoretical accounts that an end-user could apply—Ratcliff and Smith (2004) recommend Ratcliff's diffusion model as the most complete, but users might also consider Brown and Heathcote's (2005a) ballistic accumulator, or Busemeyer and Townsend's (1992, 1993) decision field theory. If a much simpler, but less complete description of the data is required, one might employ Wagenmakers et al.'s (2007) EZ diffusion. For multiple choices between $N > 2$ alternatives, only the LBA model has simple-to-use analytic solutions, making it the preferred choice.

To aid application of the LBA model, we have provided a complete mathematical implementation in the free statistical computer language R (<http://www.r-project.org/>), including functions for response time distributions (PDFs and CDFs), response probability, and the speed–accuracy tradeoff. The code is quite high-level and easy to read (around 100 lines in total), so it is simple to translate into other languages. The code works efficiently for any number of response alternatives, comes with an example program that illustrates its use, and is available for free download from <http://science-it.newcastle.edu.au/~sdb231/>.

Appendix A.

In this appendix, without loss of generality, we consider analyses for the first node ($i = 1$) only. Identical results hold for $i = 2, 3, \dots$, except with v_1 replaced by v_i . The derivation of Eq. (1) is as follows. Suppose p is a random uniform deviate on $[b - A, b]$, and q is a random normal deviate with mean v and standard deviation s (note that q is almost certainly not zero). Let $F_1(t)$ be the cumulative distribution function for the ratio p/q , then:

$$\begin{aligned}
 F_1(t) &= \text{prob}\left(\frac{p}{q} < t\right) \\
 &= \text{prob}(p < qt) \quad \# \text{ assumes } q > 0. \\
 &= \int_{-\infty}^{\infty} U(u|b-A, b)\phi(u|tv_1, ts) \, du \\
 &= \int_{b-A}^b \frac{u-b+A}{A} \phi(u|tv_1, ts) \, du + 1 - \Phi(b|tv_1, ts)
 \end{aligned}$$

The algebra from this point is cumbersome, but simple:

$$\begin{aligned}
 F_1(t) &= \int_{b-A}^b \frac{u-b+A}{A} \phi(u|tv_1, ts) \, du + 1 - \Phi(b|tv_1, ts) \\
 &= \frac{1}{A} \int_{b-A}^b u \cdot \phi(u|tv_1, ts) \, dt + \frac{A-b}{A} \int_{b-A}^b \phi(u|tv_1, ts) \, du + 1 - \Phi(b|tv_1, ts) \\
 &= \frac{1}{A} \int_{b-A}^b u \cdot \frac{1}{ts\sqrt{2\pi}} e^{-\frac{(u-tv_1)^2}{2ts^2}} \, du + \frac{b-A}{A} \Phi(b-A|tv_1, ts) - \frac{b}{v_1} \Phi(b|tv_1, ts) + 1 \\
 &= -\frac{ts}{A\sqrt{2\pi}} \int_{b-A}^b \left(-\frac{(u-tv_1)}{ts^2}\right) e^{-\frac{(u-tv_1)^2}{2ts^2}} \, du + \frac{b-A-tv_1}{A} \Phi(b-A|tv_1, ts) \\
 &\quad + \frac{tv_1-b}{A} \Phi(b|tv_1, ts) + 1 \\
 &= -\frac{ts}{A\sqrt{2\pi}} e^{-\frac{(u-tv_1)^2}{2ts^2}} \Big|_{b-A}^b + \frac{b-A-tv_1}{A} \Phi(b-A|tv_1, ts) + \frac{tv_1-b}{A} \Phi(b|tv_1, ts) + 1 \\
 &= -\frac{ts}{A\sqrt{2\pi}} \left(e^{-\frac{(b-tv_1)^2}{2ts^2}} - e^{-\frac{(b-A-tv_1)^2}{2ts^2}} \right) + \frac{b-A-tv_1}{A} \Phi(b-A|tv_1, ts) \\
 &\quad + \frac{tv_1-b}{A} \Phi(b|tv_1, ts) + 1
 \end{aligned}$$

Re-expressing the exponential terms using the normal density equation, and re-arranging:

$$\begin{aligned}
 F_i(t) &= 1 + \frac{b-A-tv_i}{A} \Phi\left(\frac{b-A-tv_i}{ts}\right) + \frac{tv_i-b}{A} \Phi\left(\frac{b-tv_i}{ts}\right) + \frac{ts}{A} \phi\left(\frac{b-A-tv_i}{ts}\right) \\
 &\quad - \frac{ts}{A} \phi\left(\frac{b-tv_i}{ts}\right)
 \end{aligned}$$

To derive the PDF (Eq. (2)) from this CDF (Eq. (1)) we differentiate with respect to t . First note that $\frac{d\Phi}{dt}(b|tv_1, ts) = \frac{d\Phi}{dt}\left(\frac{b}{t}|v_1, s\right) = -\frac{b}{t^2} \phi\left(\frac{b}{t}|v_1, s\right)$. Then:

$$\begin{aligned}
 f_1(t) &= \frac{d}{dt} \left[-\frac{ts}{A\sqrt{2\pi}} \left(e^{-\frac{(b-tv_1)^2}{2ts^2}} - e^{-\frac{(b-A-tv_1)^2}{2ts^2}} \right) + \frac{b-A-tv_1}{A} \Phi(b-A|tv_1, ts) \right. \\
 &\quad \left. + \frac{tv_1-b}{A} \Phi(b|tv_1, ts) + 1 \right] \\
 f_1(t) &= \frac{s}{A\sqrt{2\pi}} \left(-e^{-\frac{(b-tv_1)^2}{2ts^2}} + e^{-\frac{(b-A-tv_1)^2}{2ts^2}} \right) + \frac{s}{A\sqrt{2\pi}} \left(-b \frac{b-tv_1}{(ts)^2} e^{-\frac{(b-tv_1)^2}{2ts^2}} \right. \\
 &\quad \left. + (b-A) \frac{b-A-tv_1}{(ts)^2} e^{-\frac{(b-A-tv_1)^2}{2ts^2}} \right) \\
 &\quad + \frac{1}{A} \left[-v_1 \Phi(b-A|tv_1, ts) + (b-A-tv_1) \frac{-b+A}{t^2} \phi\left(\frac{b-A}{t} \middle| v_1, s\right) \right. \\
 &\quad \left. + v_1 \Phi(b|tv_1, ts) + (tv_1-b) \frac{-b}{t^2} \phi\left(\frac{b}{t} \middle| v_1, s\right) \right] \\
 f_1(t) &= \frac{s}{A\sqrt{2\pi}} \left[-\left(1 + \frac{b(b-tv_1)}{(ts)^2} \right) e^{-\frac{(b-tv_1)^2}{2(ts)^2}} + \left(1 + \frac{(b-A)(b-A-tv_1)}{(ts)^2} \right) e^{-\frac{(b-A-tv_1)^2}{2(ts)^2}} \right] \\
 &\quad + \frac{1}{A} \left[-v_1 \Phi(b-A|tv_1, ts) + (b-A-tv_1) \frac{-b+A}{t^2} \phi\left(\frac{b-A}{t} \middle| v_1, s\right) \right. \\
 &\quad \left. + v_1 \Phi(b|tv_1, ts) + (tv_1-b) \frac{-b}{t^2} \phi\left(\frac{b}{t} \middle| v_1, s\right) \right]
 \end{aligned}$$

Again we re-express the exponential terms using the normal density function and get:

$$f_i(t) = \frac{1}{A} \left[-v_i \Phi\left(\frac{b-A-tv_i}{ts}\right) + s \phi\left(\frac{b-A-tv_i}{ts}\right) + v_i \Phi\left(\frac{b-tv_i}{ts}\right) - s \phi\left(\frac{b-tv_i}{ts}\right) \right]$$

The CDF for the distribution of activation at any node i at time t and for activation z (Eq. 4) is derived via integration:

$$\begin{aligned}
 W_1(z) &= \text{prob}(k + td < z) \\
 &= \int_{-\infty}^{\infty} \phi(x|v_1, s) \cdot \text{prob}(k < z - tx) dx \\
 &= \int_{\frac{z-A}{t}}^{\frac{z}{t}} \frac{z-tx}{A} \phi(x|v_1, s) dx + \int_{-\infty}^{\frac{z-A}{t}} \phi(x|v_1, s) dx \\
 &= -\frac{t}{A} \int_{\frac{z-A}{t}}^{\frac{z}{t}} x \phi(x|v_1, s) dx + \left(1 - \frac{z}{A} \right) \Phi\left(\frac{z-A}{t} \middle| v_1, s\right) + \frac{z}{A} \Phi\left(\frac{z}{t} \middle| v_1, s\right) \\
 &= -\frac{t}{A} \int_{\frac{z-A}{t}}^{\frac{z}{t}} s^2 \frac{(x-v_1)}{s^2} \frac{1}{s\sqrt{2\pi}} e^{-\frac{(x-v_1)^2}{2s^2}} dx + \left(1 - \frac{z}{A} + \frac{v_1 t}{A} \right) \Phi\left(\frac{z-A}{t} \middle| v_1, s\right) \\
 &\quad + \left(\frac{z}{A} - \frac{v_1 t}{A} \right) \Phi\left(\frac{z}{t} \middle| v_1, s\right) \\
 &= \frac{1}{A} \left(\frac{ts}{\sqrt{2\pi}} \left(e^{-\frac{(\frac{z}{t}-v_1)^2}{2s^2}} - e^{-\frac{(\frac{z-A}{t}-v_1)^2}{2s^2}} \right) + (A-z+v_1 t) \Phi\left(\frac{z-A}{t} \middle| v_1, s\right) \right. \\
 &\quad \left. + (z-v_1 t) \Phi\left(\frac{z}{t} \middle| v_1, s\right) \right)
 \end{aligned}$$

The associated PDF (Eq. (5)) is again obtained by differentiation:

$$\begin{aligned}
 w_1(z) &= \frac{d}{dz} W_1(z) \\
 &= \frac{d}{dz} \left[\frac{ts}{A\sqrt{2\pi}} \left(e^{-\frac{(\bar{z}-v_1)^2}{2s^2}} - e^{-\frac{(\bar{z}-A-v_1)^2}{2s^2}} \right) + \frac{A-z+v_1t}{A} \Phi\left(\frac{z-A}{t} \middle| v_1, s\right) + \frac{z-v_1t}{A} \Phi\left(\frac{z}{t} \middle| v_1, s\right) \right] \\
 &= \frac{1}{A} \left(-\frac{(z-v_1t)}{t} \phi\left(\frac{z}{t} \middle| v_1, s\right) + \frac{(z-A-v_1t)}{t} \phi\left(\frac{z-A}{t} \middle| v_1, s\right) \right) \\
 &\quad - \frac{1}{A} \Phi\left(\frac{z-A}{t} \middle| v_1, s\right) + \frac{A-z+v_1t}{tA} \phi\left(\frac{z-A}{t} \middle| v_1, s\right) + \frac{1}{A} \Phi\left(\frac{z}{t} \middle| v_1, s\right) + \frac{z-v_1t}{tA} \phi\left(\frac{z}{t} \middle| v_1, s\right) \\
 &= \frac{1}{A} \left(\Phi\left(\frac{z}{t} \middle| v_1, s\right) - \Phi\left(\frac{z-A}{t} \middle| v_1, s\right) \right)
 \end{aligned}$$

References

- Audley, R. J., & Pike, A. R. (1965). Some alternative stochastic models of choice. *British Journal of Mathematical and Statistical Psychology*, *18*, 207–225.
- Bogacz, R., Brown, E., Moehlis, E., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review*, *113*, 700–765.
- Brown, S., & Heathcote, A. (2005a). A ballistic model of choice response time. *Psychological Review*, *112*(1), 117–128.
- Brown, S. D., & Heathcote, A. (2005b). Practice increases the efficiency of evidence accumulation in perceptual choice. *Journal of Experimental Psychology: Human Performance & Perception*, *31*, 289–298.
- Brown, S. D., & Steyvers, M. (2005). The dynamics of experimentally induced criterion shifts. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *31*, 587–599.
- Brown, S. D., Marley, A. A. J., Donkin, C., & Heathcote, A. (in press). An integrated architecture for absolute identification. *Psychological Review*.
- Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences*, *23*, 255–282.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459.
- Carpenter, R. H. S. (2004). Contrast, probability, and saccadic latency: Evidence for independence of detection and decision. *Current Biology*, *14*, 1576–1580.
- Dzhafarov, E. N. (1993). Grice-representability of response time distribution families. *Psychometrika*, *58*(2), 281–314.
- Gold, J., & Shadlen, M. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Science*, *5*(1), 10–16.
- Grice, G. R. (1972). Application of a variable criterion mode to auditory reaction time as a function of the type of catch trial. *Perception and Psychophysics*, *12*, 103–107.
- Hanes, D. P., & Carpenter, R. H. S. (1999). Countermanding saccades in humans. *Vision Research*, *39*, 2777–2791.
- Heathcote, A., & Brown, S. D. (2004). Reply to Speckman and Roudier: A theoretical basis for QML. *Psychonomic Bulletin & Review*, *11*, 577–578.
- Heathcote, A., Brown, S. D., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, *9*(2), 394–401.
- Heathcote, A., Hayes, B., Sutton, K. & Mewhort, D.J.K. (submitted for publication). A theory of fluent choice.
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11–26.
- Laberge, D. (1962). A recruitment theory of simple behavior. *Psychometrika*, *27*(4), 375–396.

- Lacouture, Y., & Marley, A. A. J. (1991). A connectionist model of choice and reaction time in absolute identification. *Connection Science*, 3, 401–433.
- Lacouture, Y., & Marley, A. A. J. (1995). A mapping model of bow effects in absolute identification. *Journal of Mathematical Psychology*, 39, 383–395.
- Lacouture, Y., & Marley, A. A. J. (2004). Choice and response time processes in the identification and categorization of unidimensional stimuli. *Perception & Psychophysics*, 66(7), 1206–1226.
- Laming, D. R. (1968). *Information theory of choice reaction times*. New York: Academic press.
- Leach, J. C. D., & Carpenter, R. H. S. (2001). Saccadic choice with asynchronous targets: Evidence for independent randomisation. *Vision Research*, 41, 3437–3445.
- 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.
- Mazurek, M., Roitman, J., Ditterich, J., & Shadlen, M. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex*, 13(11), 891–898.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330.
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, 118, 346–373.
- McMillen, T., & Holmes, P. (2006). The dynamics of choice among multiple alternatives. *Journal of Mathematical Psychology*, 50(1), 30–57.
- Pietsch, A., & Vickers, D. (1997). Memory capacity and intelligence: Novel techniques for evaluating rival models of a fundamental information-processing mechanism. *The Journal of General Psychology*, 124(3), 231–339.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling the accumulation of partial information. *Psychological Review*, 95, 238–255.
- Ratcliff, R. (2001). Putting noise into neurophysiological models of simple decision making. *Nature Neuroscience*, 4, 336–337.
- Ratcliff, R. (2002). A diffusion model account of reaction time and accuracy in a brightness discrimination task. *Psychonomic Bulletin & Review*, 9, 278–291.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9, 347–356.
- Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, 26(1), 127–140.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating the parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9, 438–481.
- Ratcliff, R., Cherian, A., & Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two choice decisions. *Journal of Neurophysiology*, 90, 1392–1407.
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, 111, 159–182.
- Reddi, B. A. J. (2001). Decision making: The two stages of neuronal judgement. *Current Biology*, 11, R603–R606.
- Reddi, B. A. J., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature Neuroscience*, 3, 827–830.
- Reddi, B. A. J., Asrress, K. N., & Carpenter, R. H. S. (2003). Accuracy, information, and response time in a saccadic decision task. *Journal of Neurophysiology*, 90, 3538–3546.
- Reeves, A., Santhi, N., & DeCaro, S. (2005). A random-ray model for visual search and object recognition. *Spatial Vision*, 18, 73–83.
- Roitman, J., & Shadlen, M. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, 22(21), 9475–9489.
- Schall, J. (2001). Neural basis of deciding, choosing and acting. *Nature Reviews in Neuroscience*, 2, 33–42.
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27, 161–168.
- Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, 32, 135–168.

- Smith, P. L., & Vickers, D. (1989). Modeling evidence accumulation with partial loss in expanded judgment. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 797–815.
- Stone, M. (1960). Models for choice–reaction time. *Psychometrika*, *25*, 251–260.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. New York: Cambridge University Press.
- Usher, M., & McClelland, J. L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review*, *108*, 550–592.
- Usher, M., Olami, Z., & McClelland, J. L. (2002). Hick’s law in a stochastic race model with speed–accuracy tradeoff. *Journal of Mathematical Psychology*, *46*, 704–715.
- Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*, 208–256.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, *13*, 37–58.
- Vickers, D. (1978). An adaptive module of simple judgements. In J. Requin (Ed.), *Attention and performance, VII*. Hillsdale, NJ: Erlbaum.
- Vickers, D. (1979). *Decision processes in visual perception*. New York: Academic Press.
- Vickers, D. (1995). The Frequency Accrual Speed Test (FAST) index: A new measure of “mental speed?” *Personality and Individual Differences*, *19*, 863–879.
- Vickers, D., & Lee, M. D. (1998). Dynamic models of simple judgments. I: Properties of a self-regulating accumulator module. *Nonlinear Dynamics, Psychology, and Life Sciences*, *2*(2), 169–194.
- Vickers, D., & Lee, M. D. (2000). Dynamic models of simple judgments. II: Properties of a self-organizing PAGAN (Parallel, Adaptive, Generalized Accumulator Network) model for multi-choice tasks. *Nonlinear Dynamics, Psychology, and Life Sciences*, *4*(3), 1–31.
- Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *4*, 3–22.