Journal of Experimental Psychology:
Human Perception and Performance
2009, Vol. 35, No. 6, 1865–1897

0096-1523/09/$12.00 DOI: 10.1037/a0016926

Reward Rate Optimization in Two-Alternative Decision Making: Empirical Tests of Theoretical Predictions

Patrick Simen
Princeton University

David Contreras
University of Granada, Granada, Spain

Cara Buck
University of California, San Diego

Peter Hu
University of Chicago

Philip Holmes and Jonathan D. Cohen
Princeton University

The drift-diffusion model (DDM) implements an optimal decision procedure for stationary, 2-alternative forced-choice tasks. The height of a decision threshold applied to accumulating information on each trial determines a speed–accuracy tradeoff (SAT) for the DDM, thereby accounting for a ubiquitous feature of human performance in speeded response tasks. However, little is known about how participants settle on particular tradeoffs. One possibility is that they select SATs that maximize a subjective rate of reward earned for performance. For the DDM, there exist unique, reward-rate-maximizing values for its threshold and starting point parameters in free-response tasks that reward correct responses (R. Bogacz, E. Brown, J. Moehlis, P. Holmes, & J. D. Cohen, 2006). These optimal values vary as a function of response–stimulus interval, prior stimulus probability, and relative reward magnitude for correct responses. We tested the resulting quantitative predictions regarding response time, accuracy, and response bias under these task manipulations and found that grouped data conformed well to the predictions of an optimally parameterized DDM.

Keywords: diffusion, response time, decision making, optimization, reward rate

When an organism extracts signals out of noisy inputs from the environment, it faces a fundamental tradeoff: Should it spend more time observing a stimulus to increase certainty about its identity and the appropriate response to it, or should it act more quickly at the cost of greater inaccuracy? Such a tradeoff between speed and accuracy has long been recognized as a ubiquitous feature of human behavior in speeded response tasks (Fitts, 1966; Garrett, 1922; Pachella & Pew, 1968; Schouten & Bekker, 1967; Wickelgren, 1977). However, the factors that lead to a particular tradeoff are still not well understood.

Patrick Simen, Princeton Neuroscience Institute, Princeton University; David Contreras, Psychology Department, University of Granada, Granada, Spain; Cara Buck, Psychology Department, University of California, San Diego; Peter Hu, Psychology Department, University of Chicago; Philip Holmes, Program in Applied and Computational Mathematics and Mechanical and Aerospace Engineering Department, Princeton University; Jonathan D. Cohen, Princeton Neuroscience Institute and Psychology Department, Princeton University.

This work was supported by Public Health Service Grants MH062196 (to Patrick Simen, Cara Buck, Jonathan D. Cohen, and Philip Holmes) and F32MH080524 (to Patrick Simen) and by Department of Education Grant DE-FG02-95ER25238 (to Philip Holmes).

We thank Josh Gold for providing critical software and generous technical assistance and our reviewers for many valuable comments and suggestions.

Correspondence concerning this article should be addressed to Patrick Simen, Green Hall, Washington Road, Princeton, NJ 08544. E-mail: psimen@princeton.edu

Clues about the nature of speed–accuracy tradeoff (SAT) selection have emerged from theoretical and behavioral research on decision making in simple, two-alternative forced-choice (2AFC) tasks, which require participants to choose one or the other alternative on every trial (e.g., Audley & Pike, 1965; Busemeyer & Townsend, 1993; LaBerge, 1962; Laming, 1968; Link, 1975; Link & Heath, 1975; Ratcliff, 1978; Smith & Vickers, 1989; Stone, 1960; Usher & McClelland, 2001; Vickers, 1970). Other clues come from physiological research on the neural mechanisms that may underlie this type of decision making (e.g., Carpenter & Williams, 1995; Gold & Shadlen, 2002; Hanes & Schall, 1996; Ratcliff, Cherian, & Segraves, 2003; Roitman & Shadlen, 2002; Schall, 2001; Shadlen & Newsome, 2001; Smith & Ratcliff, 2004). In particular, a large body of evidence (e.g., Palmer, Huk, & Shadlen, 2005; Ratcliff & Rouder, 2000; Ratcliff, Thapar, Gomez, & McKoon, 2004; Voss, Rothermund, & Voss, 2004) now strongly suggests that decision making in 2AFC tasks can be accurately described by the drift-diffusion model (DDM; Ratcliff, 1978), for which the SAT can be controlled by adjusting a single parameter (the decision threshold parameter, described later).

In its simplest form, the DDM is simply an application of the sequential probability ratio test (SPRT) to a decision-making task (Stone, 1960). The SPRT (Wald, 1945) is the optimal algorithm for two-alternative hypothesis testing when the likelihoods of data samples under each hypothesis are known and stationary (constant from trial to trial): on average, the SPRT will be fastest to reach a decision for a given level of accuracy and most accurate for a

given response time (RT), relative to any other procedure (Wald & Wolfowitz, 1948).

In sequential sampling models (which include the SPRT as a special case), evidence favoring each of the two alternatives is added to any prior expectations by repeatedly sampling a given stimulus. When the sampling happens continuously, the iterative log-likelihood ratio computation of the SPRT is equivalent to a drift-diffusion (DD) process (Feller, 1968), which we discuss in the following section. If the task involves free responding, in which participants can respond at any time after stimulus onset, then the corresponding response is made when the evidence favoring one alternative crosses a decision threshold. The choice of threshold determines the SAT: Lower thresholds permit faster responses but at the expense of less accumulation of information and therefore less accurate performance; higher thresholds support greater accuracy but at the expense of slower responding. The choice of starting point determines the response bias: If the starting point of evidence accumulation is closer to one response's threshold, then the probability of that response increases.

The drift parameter of the DDM is equivalent to the average rate at which information accumulates. If drift is determined by the logarithm of the stimulus likelihood ratios (and not modulated, e.g., by strategic control processes), then the conditions of the SPRT–optimality theorem apply, and no other model can make decisions faster on average than the DDM for a given level of accuracy. However, what level of accuracy—and therefore, which point along the model's SAT function—should be preferred? How should prior beliefs be incorporated into the decision process?

The SPRT does not specify how to select a particular SAT (by specifying a threshold value) or a particular response bias (by specifying a starting point), and little is known about how human participants do so. One possibility is that they seek to maximize the number of correct responses per unit time, especially in fixed-duration tasks in which faster responding leads to a greater total number of trials. This is equivalent to maximizing the rate of reward when correct responses earn rewards. Reward maximizing behavior has long been used in signal detection theory to construct receiver-operating characteristic curves (Tanner & Swets, 1954), and the effectiveness and logical consistency of payoffs as feedback in human behavioral research in general have been recognized at least since the 1960s (Edwards, 1961). Recent theoretical work has demonstrated that, for any given set of task parameters, there is a unique, optimal combination of threshold and starting point for the DDM that will maximize the expected reward rate (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006). This result can be used to make quantitative predictions about the way in which task factors should influence SAT and response bias. In this study, we sought to test these predictions and determine whether human participants adjust SATs and response biases to maximize reward rate.

Our study focuses on three factors in particular: the average response–stimulus interval (RSI), which determines the pace of the task, the prior probability of each of the two stimuli, and the relative reward associated with correct responses to each stimulus. Bogacz et al. (2006) examined the influence of these variables on the optimal threshold and the optimal starting point of evidence accumulation for the DDM. In the section that follows, we briefly review this theoretical work and the behavioral predictions it entails. We then describe three experiments conducted to test these predictions. Their results provide new support for the DDM as a model of human decision-making performance in 2AFC tasks; additionally, they support the hypothesis that human participants adapt response thresholds and starting points to maximize rewards, as predicted by a reward-rate-optimized DDM.

The DDM

We now briefly describe the DDM and the quantitative RT and accuracy predictions that we test in our experiments.

A DD process is the limiting case of a random walk in which the time between steps becomes vanishingly small (Feller, 1968). Technically, it is defined by the simple stochastic differential equation:

$$dx = A\ dt + c\ dW. \tag{1}$$

Here x represents the net evidence accumulated in favor of one of the two alternatives (and $-x$ evidence in favor of the other); the drift A represents the discriminability of the stimulus favoring one alternative (with $-A$ favoring the other, assuming equal discriminabilities); and c weights the influence of a Wiener (Brownian motion) process W, which represents the cumulative effect on x of white noise in the stimulus[1] (see Figure 1). A sample path of the process (i.e., a particular random walk trajectory) begins with x at a specified starting point x_0, which can be taken to represent the decision maker's prior belief about the relative likelihood of each stimulus type. It ends when the value of x exceeds a threshold $\pm z$ in the positive or negative direction. This "first-passage" across a threshold defines the decision time (DT) of the process. In fitting to empirical data, an additional residual latency component T_0 (reflecting sensory and motor processes unrelated to the decision itself) is added to DT to derive the predicted RT: $RT = DT + T_0$.

For the DDM with starting point equidistant from both thresholds, expected DT (denoted $\overline{DT}$) and the expected proportion of errors (denoted $\overline{ER}$) depend only on the signal-to-noise ratio (SNR) A/c and the threshold-to-signal ratio z/A, as described by the following analytic expressions (Busemeyer & Townsend, 1992; cf. Bogacz et al., 2006 and Gardiner, 2004):[2]

$$\overline{ER} = \frac{1}{1 + e^{(2Az/c^2)}}, \tag{2}$$

$$\overline{DT} = \frac{z}{A}\tanh\frac{Az}{c^2}. \tag{3}$$

[1] Specifically, a Wiener process describes the idealized Brownian motion of a point particle moving in one dimension whose position (plotted on the vertical axis in Figure 1) becomes more and more uncertain over time as a result of continuous bombardment by upward and downward impulses constituting a Gaussian white-noise process. Such a particle's vertical position is distributed normally with standard deviation $c\sqrt{t}$, where t is the amount of time elapsed since the start of the process (Gardiner, 2004). This distribution therefore describes the process of diffusion in liquid of a substance consisting of many such particles. As a description of evidence integration, it also follows directly from the assumption of sequential sampling from one of two Gaussian distributions with equal variance c^2 and means equal to $-A$ and A, respectively (Ratcliff, 1978); in the terminology of signal detection theory, $\beta = 0$, and $d' = 2Ac$ (Green & Swets, 1966). Under this interpretation, x represents the logarithm of the odds ratio that the signal comes from one or the other distribution.

[2] When the starting point is not equidistant, as is optimal when one stimulus is more likely than the other, $\overline{ER}$ and $\overline{DT}$ have more complicated expressions that are given in Appendix B.

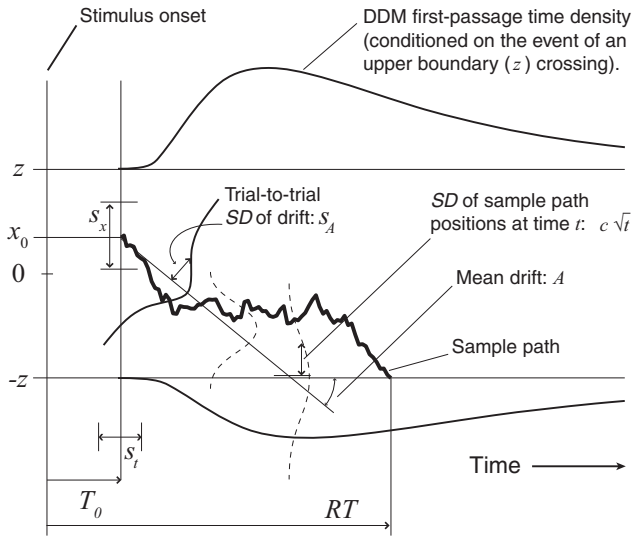

Figure 1. Parameters, first-passage density, and sample path for the extended drift-diffusion model (DDM). Parameters of the DDM are labeled according to the terminology of Bogacz et al. (2006); see Appendix A for a translation into the terminology of Ratcliff and colleagues (e.g., Ratcliff & Rouder, 1998). RT = response time.

This allows quantitative predictions to be made about $\overline{DT}$ and $\overline{ER}$ for a given drift A, noise level c, and threshold z. Drift and noise reflect the influence of two primary factors: the intrinsic discriminability of the stimulus in the environment and the signal-to-noise properties of the internal processes responsible for transducing, encoding, and attending to the stimulus. The former can be experimentally manipulated, and the latter is frequently assumed to be relatively stable for motivated performance within a given task condition. Accordingly, A and c can be estimated for a particular stimulus and individual. What is less clear is the basis on which decision makers choose the threshold z and the starting point x_0—that is, how they choose to trade off speed against accuracy and how they choose a response bias, if any. We tested the hypothesis that participants make decisions using a DD process and that they parameterize the process so as to maximize reward rate, under the assumption of a physically unavoidable upper bound on the SNR, A/c.

Reward-Rate Optimization of the DDM

Recent theoretical work (Bogacz et al., 2006) has shown that when drift, noise, mean RSI ($\overline{RSI}$), prior stimulus probability, and the relative reward for correct responses to each stimulus are held constant in a free-response 2AFC task, then there exist unique, optimal threshold, and starting point values for the DDM[3] that maximize expected reward rate ($\overline{RR}$), defined as follows (Gold & Shadlen, 2002):

$$\overline{RR} = \frac{1 - \overline{ER}}{\overline{DT} + T_0 + \overline{RSI}}. \quad (4)$$

Here we assume that errors are unrewarded. We now examine how optimal DDM parameterizations (those that maximize Equation 4) depend on the task conditions that we manipulate in our experiments.

RSI. By substituting Equations 2 and 3 into Equation 4 and solving for the maximum $\overline{RR}$, we can derive the following equation describing the optimal value of z as a function of A, c, T_0, and $\overline{RSI}$ (Bogacz et al., 2006):

$$e^{(2Az/c^2)} - 1 = \frac{2A^2}{c^2} \cdot \left(\overline{RSI} + T_0 - \frac{z}{A} \right). \quad (5)$$

(Here we assume that the starting point is equidistant between the two thresholds—$x_0 = 0$—because this maximizes expected reward rate in tasks with equally likely and equally rewarded stimuli.)

Because the left side of Equation 5 increases with z while the right side decreases, there is a unique value of z that solves Equation 5 for a given combination of A, c, T_0, and $\overline{RSI}$. This can be seen clearly in Panel A of Figure 2, in which expected reward rate ($\overline{RR}$, given by Equation 4) is plotted as a function of threshold for representative values of A, c, and T_0 (obtained by fitting the DDM to behavioral data), and for a variety of average RSI values. The figure shows that a unique, reward-rate-maximizing threshold exists for each RSI and that this optimal threshold value grows as RSI increases (the specific value can be determined by solving Equation 5 numerically, e.g., by Newton's method). Insofar as A, c, T_0, and z are stable for a given individual and task condition, their values can be estimated from behavioral performance and used to evaluate the goodness of fit of the DDM. Furthermore, if evidence suggests that A, c, and T_0 are stable for a given individual across manipulations of task variables such as $\overline{RSI}$, then changes in z can be estimated in response to such task manipulations and compared with the optimal values predicted by Equation 5. Panel B of Figure 2 plots optimal threshold values as a function of $\overline{RSI}$. Optimal threshold predictions in turn entail specific expected reward rates, RTs, and accuracies that can be compared with data (remaining panels of Figure 2). Experiment 1 was designed to test these predictions.

Stimulus probability. Thus far, we have focused on conditions in which each stimulus is presented equally often. If one stimulus appears more often than the other, then maximization of reward requires the starting point of evidence integration for the DDM (x_0) to be moved closer to the threshold corresponding to the more frequent response (Bogacz et al., 2006; cf. Edwards, 1965; & Laming, 1968). This produces faster RTs when the drift is in the direction of the closer threshold and more errors when the drift is in the opposite direction. However, the reduced frequency of trials for this case makes their increased inaccuracy worth the cost. Specifically, if Π denotes the probability of stimuli for which crossings of z are correct, then for optimal performance (Edwards, 1965), the initial condition of x should be set as

$$x_0 = \frac{c^2}{2A} \ln \left(\frac{\Pi}{1 - \Pi} \right). \quad (6)$$

Note that x_0 should equal 0 when $\Pi = 1/2$. In addition, a value of Π greater than 0.5 produces a reduction in the optimal threshold value,

[3] An analogous result holds in the case of equally likely and equally rewarded stimuli for any model that produces a concave speed–accuracy tradeoff function (SATF) relating accuracy to response time; see Appendix C. See also related derivations of thresholds and starting points that minimize a weighted sum of DT and accuracy rather than maximizing $\overline{RR}$ (Edwards, 1965; Rapoport & Burkheimer, 1971).

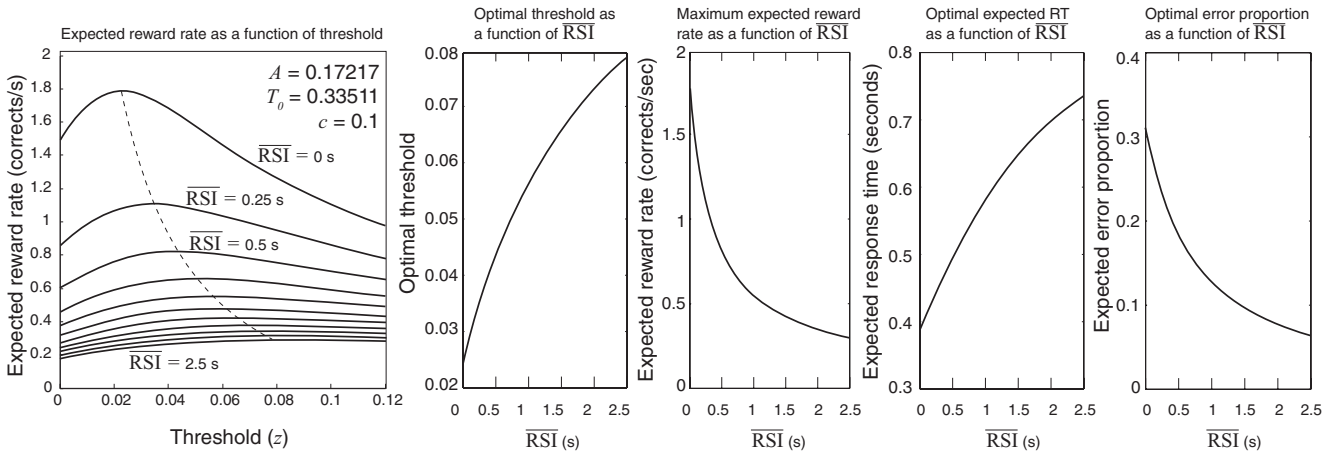

Figure 2. (A) Expected reward rate ($\overline{RR}$) plotted as a function of threshold z for a range of response–stimulus interval (RSI) values (dashed curve connects the peaks of each $\overline{RR}$ curve). (B) Optimal threshold as a function of RSI. (C) $\overline{RR}$ as a function of RSI, assuming optimal thresholds at each RSI. (D) Expected response time as a function of RSI, assuming optimal thresholds. (E) Expected proportion of errors ($\overline{ER}$) as a function of RSI, assuming optimal thresholds. RT = response time.

which in the case of unequal stimulus probability is obtained by numerically solving the following equation (Bogacz et al., 2006):

$$e^{(2Az/c^2)} - 1 = \frac{2A^2}{c^2} \cdot \left(\overline{\text{RSI}} + T_0 - \frac{z}{A} \right) + (1 - 2\Pi)\ln\left(\frac{\Pi}{1 - \Pi} \right). \tag{7}$$

(Equation 7 reduces to Equation 5 for $\Pi = 1/2$.) Expected accuracy and DT for the optimally parameterized DDM with $\Pi > 1/2$ are given in Appendix B.

Although the optimal value of x_0 does not depend on RSI, it interacts in interesting ways with the optimal threshold as the mean RSI is changed. Figure 2 (second panel from left) shows that, as RSI increases, the optimal threshold also increases. This relationship also holds in the case of unequal stimulus frequencies. Thus, decreasing the RSI while increasing the inequality in stimulus probabilities effectively exaggerates the shift of the starting point toward the threshold for the *favored* response (i.e., the response corresponding to the more likely stimulus). For RSIs that are sufficiently short and values of Π that are sufficiently close to 1, Equation 6 places the optimal starting point beyond the favored response threshold. In this case, the simplest interpretation of the theory predicts that the decision maker should forgo integration and choose the favored response on every trial. Assuming that there is a penalty for anticipatory responding (i.e., responding before stimulus onset), RT should simply reflect signal detection and therefore equal T_0, and the proportion of errors should equal the probability of the less likely stimulus. We refer to this behavior as *nonintegrative* responding to indicate that no integration of evidence is being carried out by the decision maker; nonintegrative responding is equivalent to making fast-guess responses (Ollman, 1966; Yellott, 1971), except that it involves always making the same guess that the favored response is correct.

Bogacz et al. (2006) described task conditions in which nonintegrative responding is expected by dividing the three-dimensional task parameter space into two regions separated by a curved, two-dimensional *critical probability surface*. This surface—on which the optimal starting point and threshold coincide—is defined by Equation 8, which describes it in terms of $\overline{\text{RSI}}$ as a function of Π, A, c, and T_0:

$$\overline{\text{RSI}} = \frac{\dfrac{\Pi - 1/2}{1 - \Pi} + \Pi \ln\left(\dfrac{\Pi}{1 - \Pi} \right)}{A^2/c^2} - T_0. \tag{8}$$

This surface is depicted in Figure 3. The parameters defining this space are the SNR (A/c), the average RSI, and the probability of the more likely stimulus, Π. The residual decision latency (T_0) determines the height of the surface. For asymmetries Π above this surface, nonintegrative responding is expected. For points below the surface, integrative responding is expected.

It seems reasonable to expect that a sufficiently strong asymmetry in stimulus ratios would lead participants to choose exclusively one alternative in speeded-response, 2AFC tasks regardless of other factors, such as RSI. However, Equation 8 prescribes a

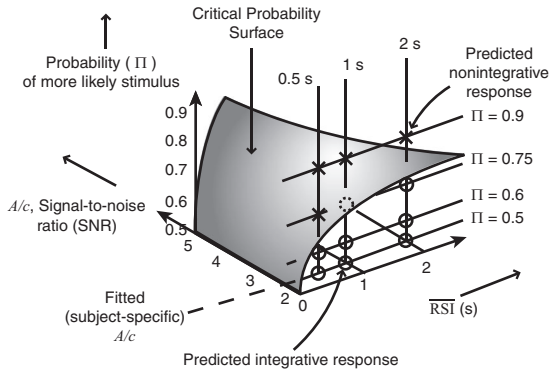

Figure 3. Critical probability surface, dividing parameter space into predicted integrative and nonintegrative conditions.

parametric and possibly counterintuitive relationship between DDM parameters and task parameters that should produce nonintegrative responding. In particular, this relationship implies that a given prior probability Π should produce nonintegrative responding for short RSIs but not for longer RSIs. Fits of the DDM parameters (particularly, T_0 and the ratio A/c) allow prediction of the values of Π and $\overline{\text{RSI}}$ at which this transition should occur if reward rate is being maximized. In Experiment 2, we covaried mean RSIs and stimulus ratios to determine whether such a surface exists and, if so, whether its shape conforms to the predictions of the DDM concerning reward rate maximization.

Relative reward. Because we assume that participants seek to maximize reward rate, direct manipulations of the reward associated with each response should also produce predictable effects on behavior. Bogacz et al. (2006) also investigated tasks in which a proportion r of some unit of reward is assigned to one response (when it is correct), and the remaining proportion $1 - r$ is assigned to the other response when correct. In contrast to the case of unequal stimulus probabilities, analytical expressions for optimal starting points were not obtainable in the case of reward asymmetries. However, numerical results indicated that differences in reward should produce effects similar to those of unequal stimulus probabilities, except that values of r were predicted to produce stronger response biases than those produced by equivalent values of Π (in contrast to relative reward, the absolute magnitude of the rewards was predicted to be irrelevant).

Specifically, two expressions were obtained that define an interval within which the optimal starting point should lie. As the sum of $\overline{\text{RSI}}$ and T_0 grows small, Equation 9 defines the upper boundary of this interval, which is the same as the optimal starting point for unequal stimulus probabilities if Π is replaced by r:

$$x_0 \rightarrow \frac{c^2}{2A}\ln\left(\frac{r}{1-r}\right). \quad (9)$$

As the sum of $\overline{\text{RSI}}$ and T_0 grows large, Equation 10 defines the lower boundary of this interval:

$$x0 \rightarrow \frac{c^2}{4A}\ln\left(\frac{r}{1-r}\right). \quad (10)$$

The optimal starting point shift is thus smaller in the case of reward asymmetry than in the case of an equivalent stimulus probability ($r = \Pi$). Optimal thresholds, in contrast, are dramatically reduced in response to reward asymmetry relative to stimulus proportion asymmetry. The net effect is that the optimal separation between the starting point and the favored response threshold is smaller in the case of reward asymmetry.

Thus, we should expect unequal rewards to bias decision makers toward one response over the other in a manner qualitatively like that predicted for unequal stimulus probabilities. Bogacz et al. (2006) numerically computed a critical reward surface that is analogous to the critical probability surface of Figure 3 but that predicts a transition to nonintegrative responding at larger values of $\overline{\text{RSI}}$. Experiment 3 was designed to test this prediction.

Extended DDM and Data Fitting

The theoretical work described earlier has focused on the simplest version of the DDM, in which the absolute value of the drift, the starting point, and the residual latency are all assumed to be constant for a given participant and a given task condition. Hereinafter, we refer to this version of the DDM as the pure DDM. The pure DDM, like the SPRT itself, predicts equal mean RTs for correct and error responses, but this prediction is frequently violated in practice and has led some to reject the SPRT as a decision-making model (e.g., Luce, 1986). However, assuming random variability across trials in A, x_0, and T_0 corrects this deficiency (Ratcliff & Rouder, 1998; Ratcliff, Van Zandt, & McKoon, 1999). We refer to this form of the model (the pure DDM with three additional parameters, s_A, s_x, and s_t, respectively, as well as a fourth parameter, p_0, specifying the proportion of contaminant RTs uniformly distributed between the minimum and maximum RT, as in Ratcliff & Tuerlinckx, 2002) as the extended DDM (depicted in Figure 1). (Assuming that thresholds are set optimally, the pure-DDM/SPRT equivalence and the theorem of Wald and Wolfowitz [1948] imply that rewards are maximized when s_A, s_x, s_t, and p_0 are all 0.)

The extended DDM fits a broader range of empirical data sets (especially those with differences in average RT between correct and error responses), but it has not yet been found to be amenable to formal analysis (although see Bogacz et al., 2006, for analytical approximations and numerical approaches). Thus, the extended DDM does not yield explicit relationships such as those of Equations 5–7. Furthermore, although adding more parameters gives the DDM enough flexibility to fit data, it also exacerbates a problem that occurs during fitting: This is that fitted values of DDM parameters are correlated with each other (Ratcliff & Tuerlinckx, 2002). For example, when fitting data, a minimum fit error parameter set can be modified by simultaneously increasing both drift and threshold; this leads to a parameter set with larger values that may nevertheless have a fit error nearly as low as the original; reducing multiple parameters simultaneously can similarly result in good fits. Thus, there is a tendency for parameter values to rise and fall together during fitting. However, because variability parameters are equal to 0 in the pure DDM and cannot be less than 0 in the extended DDM (indeed, fitted values of these are almost always greater than 0), these correlations among parameters appear to explain why, in fits to our empirical data, the extended DDM always results in larger drift, threshold, and T_0 parameter values than in fits of the pure DDM.[4]

The values of these parameters are critical for the numerical accuracy of the predictions of Equations 5–7, but no widely accepted method exists for controlling parameter inflation as parameters are added to the simpler, pure DDM. If fit error is the only criterion on which parameter values are judged, then larger values are acceptable. If a source of bias toward larger values

[4] To understand why this might happen, we have generated simulated data sets using the DDM and then contaminated them by a small proportion of RTs from other distributions. Although parameters can be recovered accurately by extended-DDM fits to uncontaminated data, this is not always the case when unmodeled contaminants are included (e.g., contaminants that are narrowly distributed rather than uniformly distributed between the minimum and maximum RT). In such cases, extended-DDM fits tend to inflate the variability parameter estimates (making them greater than 0) as well as the theory-critical drift, threshold, and residual latency parameters. Thus, parameter inflation in fits to our empirical data may result from the inclusion of unmodeled contaminants.

exists, however, then techniques should be considered for limiting the growth of parameters during fitting.

Our approach to the parameter-inflation phenomenon was to use the extended DDM to fit data but to constrain its variability parameters by applying upper bounds on their allowable values. This approach left the pure-DDM parameters free to take on any values (including those that would disconfirm our hypotheses) while demonstrably reducing parameter inflation. Figure 4, for example, demonstrates that drift and T_0 increased as upper bounds on drift variability, starting point variability, and residual latency variability were relaxed in fits to the data from Experiment 1 (standard error bars were generated in a cross-validation procedure that involved fitting 150 subsets of half the data at each upper bound value). Threshold values in the three RSI conditions, in contrast, remained flat across bound values. Starting points (not plotted), showed the same constancy. At the same time, fit error naturally decreased as constraints were relaxed. Validation error, computed by applying the fitted parameters in each fit to the unfitted half of the data, showed no signs of overfitting; that is, it never increased as bounds were relaxed. However, failure to find evidence of overfitting does not imply the absence of possible bias in the fitting procedure.

Because we currently have no method for selecting an optimal tradeoff between parameter inflation and fit error, we relied on simulations to determine the best bound values. We set the bounds in our data analyses (listed in Table 1) roughly equal to the variability parameter values recovered in the most accurate fits of A and T_0 to the simulated data sets of Ratcliff and Tuerlinckx (2002) (see Figure 6 in that article). We relied on these extended DDM simulations because they used parameter values that were relatively close to those obtained by fits to our data and because these values are representative of fits to data from a wide range of experiments (e.g., Ratcliff & Rouder, 1998, 2000; Ratcliff & Smith, 2004; Ratcliff et al., 1999). Also, because the simulated data in the correlation analyses of Ratcliff and Tuerlinckx (2002) assumed a constant value of T_0, the bound on its corresponding

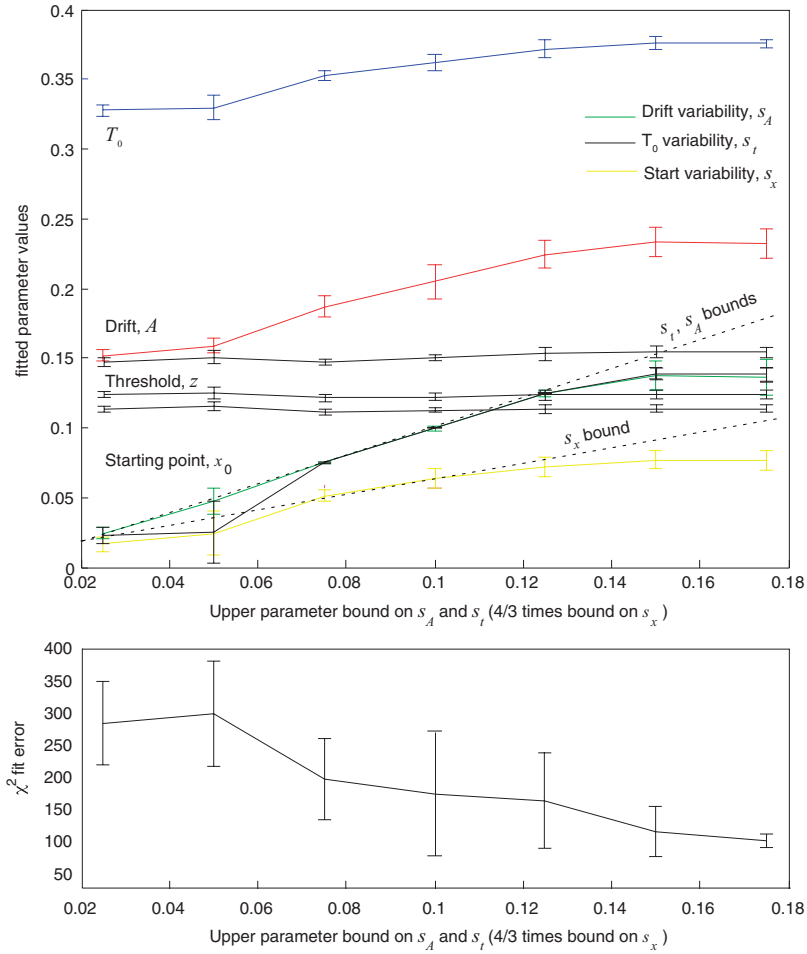

Figure 4. Top panel: Average extended drift-diffusion model parameter values from fits to 150 subsets of half the data (sampled with replacement) in each condition of Experiment 1, plotted as a function of the upper bound applied to the s_t and s_A parameters during fitting (error bars represent the standard error of the mean). Drift A and residual latency T_0 inflate as upper bounds on s_A, s_z, and s_t increase, indicating a possible source of bias in parameter estimation. Bottom panel: chi-square fit error as a function of upper bound values. Average fit error for the bound value closest to the bounds used in our analyses was approximately 200.

Table 1
Fitted Parameter Values (Pooled Participant Data) for Group Data From Response–Stimulus Interval (RSI) Conditions, With Equally Likely Stimuli

Parameter	Value		Bound
Drift (A)	0.17348		
Noise coefficient (c)	0.1		
Drift SD (s_A)	0.068683		0.08
Starting point range (s_x)	0.03		0.03
T_0 range (s_t)	0.1		0.10
Contaminant proportion (p_0)	0.029182		0.05
Residual latency (T_0)	345.47 ms[a]		
Total χ^2 fit error	194.5232		
RSI condition	500 ms	1 s	2 s
Threshold (z)	0.0558	0.0606	0.0734
Optimal threshold	0.0438	0.0566	0.0730
Starting point (x_0)	0.0029	0.0016	0.0016
Optimal starting point	0	0	0
χ^2 fit error	84.58	46.44	62.47

Note. Comparisons to empirical histograms for this fit appear in Figure 7, and comparisons to empirical quantile-probability plots appear in Figure 6.
[a] Compare with average signal detection RT of 301 ms.

variability parameter, s_t, came from our cross-validation procedure. The bounds occur roughly halfway between an asymptotic chi-square fit error of approximately 100 for completely unconstrained fits at the right edge of the graphs and a fit error of approximately 300 for the maximally constrained model (which better approximates the pure DDM) at the left edge. (The exact placement of these bounds does not drastically affect the numerical accuracy of our analytical predictions of optimal parameter values until it results in drift values well above 0.2 and T_0 values well above 370 ms, at which point predictions and fitted values match only qualitatively.)

The result was a model that could be fit much faster than the pure DDM. Resulting fit errors were small enough for the model to pass an Akaike information criterion test for model selection (Akaike, 1974) over the pure DDM, but fitted values of the theory-critical A and T_0 parameters were nevertheless close to those obtained by fitting the pure DDM. We used the resulting estimates of A and T_0 to make predictions about the effects of $\overline{\text{RSI}}$ on threshold setting in Experiment 1, the interaction of RSI with stimulus probabilities in Experiment 2, and the effect of unequal rewards for left and right responses in Experiment 3.

Experiment 1

In this experiment, we held the SNR of the stimulus constant and manipulated the mean RSI across blocks of trials in a free-response, 2AFC motion discrimination task with equally likely stimuli (i.e., $\Pi = 0.5$). We sought to test the hypothesis that participants' SATs shift across conditions in the absence of explicit instructions. We also sought to determine whether the extended DDM could account for RT distributions and accuracy in all conditions and whether fitting the model to data would produce parameter estimates that conform to the following predictions of the pure DDM,[5] parameterized to maximize reward rate (Bogacz et al., 2006):

Prediction 1a: Estimates of drift (A) should be constant across all RSI conditions, reflecting the assumption that participants are motivated and allocate maximum attention to the task and further reflecting the fact that the optimal strategy is to extract as much information as possible from the stimulus (which has a fixed SNR) in all task conditions.

Prediction 1b: Estimates of residual latency (T_0) should be constant across conditions and commensurate with an independently observed signal detection RT (in a signal detection task with easily detectable signals).

Prediction 1c: Estimates of the starting point x_0 should be 0 in all conditions, reflecting no predisposition toward either response.

Prediction 1d: Estimates of the threshold parameter (z) should increase as $\overline{\text{RSI}}$ increases, reflecting a shift toward accuracy (see Figure 2, Panel B).

Prediction 1e: Estimates of the threshold parameter should equal the function z (A, c, T_0, $\overline{\text{RSI}}$) defined implicitly by Equation 5, evaluated at the current RSI, and with the fitted values of A/c and T_0.

Method

Participants. Twelve participants, ranging in age from 19 to 64 years ($M = 26$), were recruited from the Princeton University campus area to participate in ten 1-hr task sessions. Experiment 1 consisted of the first five sessions; the second five sessions constituted Experiment 2. For their performance, participants were paid the greater of $10 or their total earnings in the task. Participants earned 1 cent for each correct response given, and no explicit penalties were imposed for errors. Average earnings were around $15 per session.

One participant performed at chance in all sessions, and these data were discarded. One participant dropped out after a single session. Data from two sessions were corrupted by power failures for a third participant, and this participant's remaining data were excluded from analysis. Another participant did not comply with instructions and did not wear vision-correcting glasses during some sessions, so these data were excluded as well. Finally, an older participant's data were excluded (reducing the average age to 23 and the maximum age to 27) so that age-related performance changes would not affect our findings. Data were therefore analyzed for 7 participants who completed the 10 sessions. Data for each participant were analyzed only for the last 7 of 10 sessions to reduce the impact of practice effects on the analysis.

Apparatus and stimuli. Stimuli were presented on a standard computer monitor; button press responses were entered on a standard keyboard. Stimulus display and response collection were conducted with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) extensions to MATLAB running on an Apple G4 Power

[5] These predictions are approximately the same for the extended DDM, but the approximation is worse for larger values of the variability parameters (s_A, s_x, and s_t) in the extended DDM.

Mac with the OS 9 operating system. Stimulus generation software was created for use with the Psychophysics Toolbox by J. I. Gold.

Stimuli were random-dot kinematograms, similar to those used in a series of psychophysical and decision-making experiments involving monkeys as participants (e.g., Britten, Shadlen, Newsome, & Movshon, 1992; Gold & Shadlen, 2001; Shadlen & Newsome, 2001). Stimuli consisted of an aperture of approximately 3-in. (7.62-cm) diameter viewed from approximately 2 feet (~0.610 m; approximately 8° visual angle) in which white dots (2 × 2 pixels) moved on a black background. A subset of dots moved coherently either to the left or to the right on each trial, and the remainder of dots were distractors that jumped randomly from frame to frame of the display. Motion coherence was defined as the percentage of coherently moving dots. Dot density was 17 dots per square degree, selected so that individual dots could not easily be tracked.

Procedure. Motion coherence was adapted manually at the end of each of the first three experimental sessions to produce errors in at least 10% of responses. This was done to produce a substantial sample of error RTs, which is useful for constraining fits of the DDM (Ratcliff & Tuerlinckx, 2002). Some participants required no coherence adaptation, and average motion coherence ranged from the default value of 10% to a lower limit of 5%. No participants required an increase in motion coherence (except for the participant who performed consistently at chance and whose data was excluded from analysis).

Responses involved presses of the Z key on the lower left of the keyboard with the left index finger to signal perception of leftward motion and presses of the M key on the lower right with the right index finger to signal perception of rightward motion, as in the empirical work presented in Bogacz et al. (2006) and Bogacz, Hu, Cohen, and Holmes (in press). Correct responses were signaled by an auditory beep, and after every five trials, the current total of correct responses was displayed in the center of the screen in place of the motion aperture for a duration equal on average to the mean RSI duration in each block of trials. Errors were indicated by the absence of the auditory beep.

Two measures were taken to prevent anticipatory responses, in which participants do not integrate stimulus information but instead prepare a response before stimulus onset to reduce RT and thereby increase the total opportunity for reward.[6] First, the RSI on a given trial was selected from a normal distribution with a standard deviation of 100 ms to make stimulus onset unpredictable. Second, whenever responses were recorded before or within 100 ms after the stimulus onset, a penalty delay of 4 s was imposed to reduce the opportunity to earn rewards, and a buzzing error tone was presented.

In the first five 1-hr task sessions, the two stimulus types were equally likely. Each session consisted of one practice block of 4 min (practice was reduced to 2 min in Sessions 4 and 5), followed by twelve 4-min blocks within each of which $\overline{\text{RSI}}$ was held constant. $\overline{\text{RSI}}$ was 500 ms in three blocks, 1 s in three blocks, and 2 s in six blocks. There were twice the number of 2-s RSI blocks, as these produced a significantly smaller number of trials within a 4-min block. The order of blocks and conditions was counterbalanced across sessions and across participants with a Latin square design. Self-paced rest periods occurred between blocks.

Participants were informed that the RSI might be different in different blocks and that blocks would always last 4 min; therefore, faster responding would lead to more trials overall. They were encouraged to earn as much money as possible.

The 12 blocks in each session were followed by two 2-min blocks of a signal detection task with easily detectable stimuli. In each signal detection block, stimuli were the same as in previous blocks (with a mean RSI of 500 ms), but only a single response earned rewards (a left button press in one block, and a right button press in the other). In these blocks, participants were instructed to respond as quickly as possible with the designated button press as soon as the stimulus appeared, regardless of coherent motion direction. Although discriminating motion direction was relatively difficult in the preceding 2AFC blocks, simply detecting the presence of a high-luminance, moving-dots stimulus was not—signal detection RT was rapid and narrowly distributed; no misses occurred, and false alarms (anticipations) were rare. These blocks were used to establish a minimum signal detection RT for each response (left finger and right) that could be compared as a baseline with estimates of T_0 from the signal discrimination trials, as well as with the RTs of any potentially nonintegrative responses in Experiments 2 and 3.

Analysis. We directly examined RT distributions to assess the magnitude of SAT adjustment across RSI conditions and compared observations to speed and accuracy predictions based on model fits of the DDM and on the theory of optimal threshold parameterization in Equation 5.

To maximize statistical power and assess the generality of findings, we focused our analysis on group averaged data (while noting that similar results hold for almost all individual participants; individual performance for a selected participant is examined in Appendix E). Although pooling raw data from multiple participants presents potential dangers for interpretation (Estes & Maddox, 2005; Ratcliff, 1979), group RT distributions have been shown to be useful for analysis of RT data from multiple participants (Ratcliff, 1979), and they have been used successfully in practice (Spieler, Balota, & Faust, 1996; Ratcliff et al., 2004).

We assessed group performance by pooling together the data from all participants. Frequently, a Vincentizing procedure is used to construct group RT distributions from individual RT distributions (Ratcliff, 1979; Van Zandt, 2000). This involves averaging (or taking the median of) the quantiles of individual RT distributions to derive the quantiles of an estimated RT distribution for the "average" participant. One virtue of this approach is that a set of unimodal, individual distributions cannot lead to a multimodal "average" distribution (which clearly would not represent the typical participant), although some evidence suggests that this approach has drawbacks (Rouder & Speckman, 2004). In our case, however, the Vincentized distribution appeared nearly identical to the distribution of RTs obtained simply by pooling the raw data from multiple participants (possibly because our manipulations of

[6] In a pilot experiment in which RSIs were completely predictable, anticipatory responding was produced by most participants in all conditions, regardless of RSI, as evidenced by RTs of as little as 25 ms. This pattern of behavior is consistent with a strategy of maximizing reward rate by effectively reducing the residual latency T_0—indeed, it produced much higher reward rates than those observed in Experiment 1. However, it precludes any study of the effects of reward on integration processes in decision making, whereas the theory discussed in Bogacz et al. (2006) applies when overall rates of stimulus presentation are predictable but individual stimulus onsets are unpredictable and anticipatory responding is not beneficial.

motion coherence in the first three sessions to obtain at least 10% errors tended to equalize RT and accuracy among participants). We therefore conducted the analyses that follow by pooling untransformed RT data from multiple participants; the analysis of Vincentized data leads to nearly identical results.

Results

Although 5 of 7 participants displayed clear evidence of SAT adaptation across RSI conditions by the fifth session of Experiment 1, 2 participants did not. However, data from these participants were not excluded from the pooled data analysis, and these participants did show evidence of SAT adaptation in Experiment 2.

Differences in SAT across RSI conditions. A boxplot of RT data across three conditions in Figure 5 (left panel) shows that RTs for the average participant increased as RSI increased. All pairwise median RT differences were significant ($p < .05$, Wilcoxon rank-sum test). Notches in the boxes in Figure 5 represent nonparametric 95% confidence intervals around the median, which is denoted by the horizontal line in each box. The observed average RTs are indicated with circle markers and superimposed on these plots, and the corresponding predictions based on optimal threshold values and fitted values of A and T_0 are shown with Xs. (Note that the mismatch between predictions and observations in Figure 5 cannot derive entirely from suboptimal threshold selection, which would lead to longer RTs and greater accuracy, or shorter RTs and lower accuracy, than predicted. Instead, the estimates of A and T_0 must be somewhat noisy, as RTs are longer than predicted and accuracy is lower than predicted in the 1- and 2-s RSI conditions.)

Accuracy also increased as RSI increased, as shown in the center panel of Figure 5. Error bars indicating the standard error of the mean are barely visible; differences in error proportions were highly significant. These results are consistent with an increase in threshold as RSI increases, in accord with Prediction 1d.

A predicted SATF (see Footnote 3) is shown in the right panel of Figure 5, where accuracy is plotted as a function of RT. The solid SATF curve is generated by holding all DDM parameters constant while gradually increasing thresholds. Observed RT–accuracy pairs are marked with circles; predictions for SATs in corresponding RSI conditions are marked with Xs.

Quantile probability plots. Quantile probability plots (Ratcliff, 2001) provide a compact form of representation for RT and accuracy data across multiple conditions. In a quantile probability plot (e.g., Figure 6), quantiles of a distribution of RTs of a particular type (e.g., correct responses) are plotted as a function of the proportion of responses of that type: Thus, a vertical column of N markers would be centered above the position 0.8 if N quantiles were computed from the correct RTs in a task condition in which accuracy was 80%. (Following Ratcliff and Tuerlinckx [2002] in both plotting and model-fitting, we used five RT quantiles: 0.1, 0.3, 0.5, 0.7, and 0.9.) The ith quantile in each distribution is then connected by a line to the ith quantiles of other distributions.

Here we have further elaborated quantile probability plots to include a superimposed scatterplot of individual RTs in each condition. Each sample point is plotted at a vertical coordinate corresponding to its RT value, and at a horizontal coordinate corresponding to the response probability, plus a normally distributed, random offset (lat-

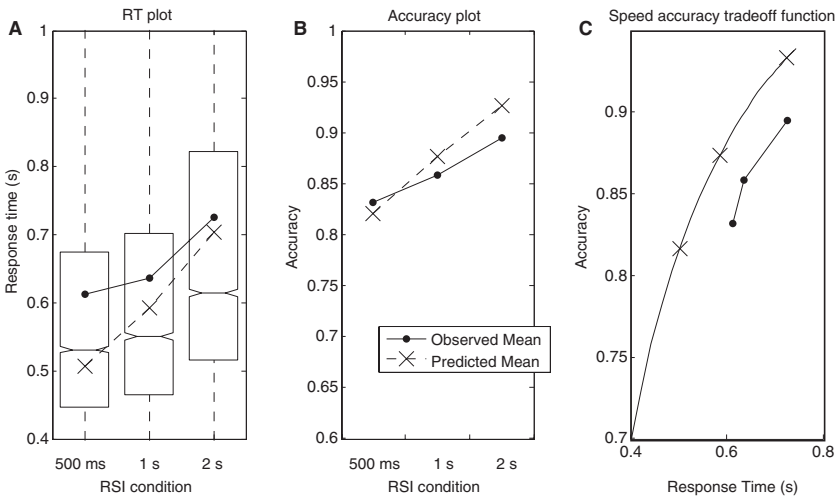

Figure 5. (A) Boxplot of response times (RTs) for pooled data for all participants. Boxes represent the interquartile range (difference between first and third quartiles), and lines bisecting the boxes represent medians. Notches represent nonparametric 95% confidence intervals around the median RTs. Dashed lines and X markers indicate the expected RT predicted by Equation 3 for a drift-diffusion model (DDM) with optimal thresholds, given values of A and T_0 obtained from the best fit of the model to the data. Solid lines and circle markers indicate observed RT averages (these are higher than the medians indicated in the boxplots because of the skew of the RT distributions). (B) Accuracy across conditions. Solid lines and circles indicate the observed proportions of correct responses. Dashed lines and Xs indicate the expected proportions, $1 - \overline{ER}$, where ER in each condition is obtained by substituting fitted A and T_0 values along with the optimal z value into Equation 2. (C) Predicted speed–accuracy tradeoff function (SATF; i.e., $1 - \overline{ER}$ as a function of $\overline{DT} + T_0$) based on a fit of the DDM. Circles indicate the observed tradeoffs in each condition; Xs indicate the optimal tradeoffs. RSI = response–stimulus interval.

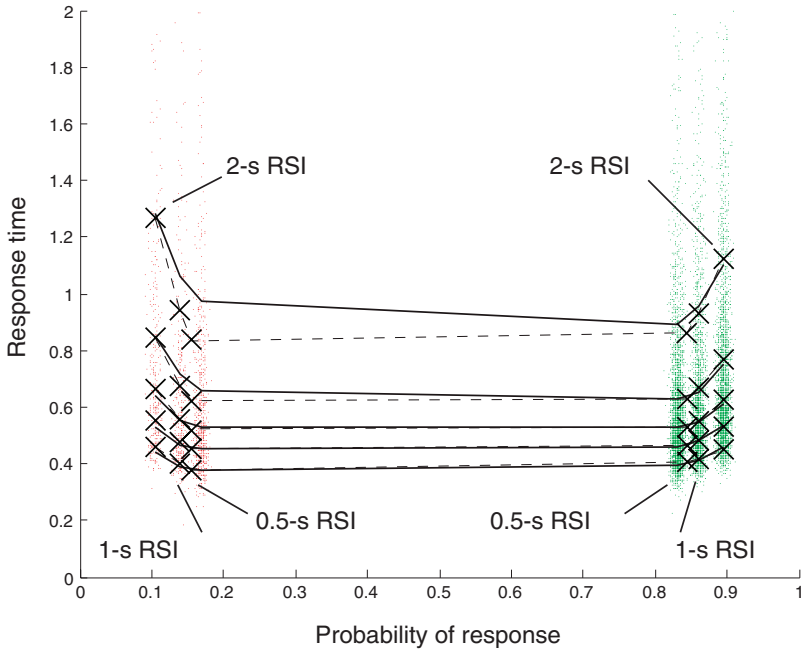

Figure 6. Quantile probability plot for pooled data from all participants in Experiment 1. Solid lines connect the *n*th quantile of the empirical data; Xs and dashed lines represent the predicted quantiles for the best fit (listed in Table 1). RSI = response–stimulus interval.

erally scattering individual RTs so that they can be discerned). This adds a visual representation of the number of responses in each condition to a quantile probability plot. Correct response RTs are plotted in green; error RTs are plotted in red.

In Figure 6, the quantile probability plot for the pooled participant data is shown for the fourth and fifth sessions together. The five lines correspond to the five RT quantiles that were computed. The six RT distributions depicted in the plot correspond to correct and error RTs in each of the three different RSI conditions: 500 ms, 1 s, and 2 s. Performance was much better than chance in all conditions, so the correct RT distributions appear on the right side of the plot.

For the average participant (represented by the pooled data), blocks with longer RSIs were associated with a higher likelihood of a correct response, as accuracy increased with increasing RSI. (This pattern also held for all but 2 of the participants individually.) Thus, the correct responses for the 2-s RSI condition appear as the rightmost column of quantiles. The error responses in this condition form the leftmost column.

Figure 6 clearly shows that the more likely correct responses coincided with longer RTs. Similarly, the corresponding error RTs were longer for the less likely errors. Thus, a tradeoff between speed and accuracy is depicted in the U-shaped plot. In fact, the data are consistent with the theory of optimal DDM parameterization and SAT adaptation: Blocks with longer RSIs were associated with more accurate but slower responses. In contrast, when changes in the drift parameter produce changes in accuracy, speed and accuracy do not trade off against each other; instead, RT and accuracy are negatively correlated. The resulting quantile probability plot in that case has an inverted-U shape, as in Ratcliff and Tuerlinckx (2002), where variations in drift, but not threshold, were simulated. This pattern of increasing RT as RSI increased was observed in all but 2 participants. The results for the average participant are thus, so far, consistent with threshold adaptation but not with drift adaptation.

There is also no significant difference between the median correct RT and the median error RT in the 500-ms ($p = .3239$, Wilcoxon rank-sum test) and 1-s ($p = .28$) RSI conditions, although there is a trend in which average error RT is slower than average correct RT by about 30 ms. Data from these conditions are therefore arguably consistent with the pure DDM. Average error RTs are significantly slower, by about 50 ms, in the 2-s RSI condition (medians are significantly different at $p = .0192$); data from this condition are therefore inconsistent with the pure DDM.

Model fits. We fit RT distributions using a constrained optimization algorithm implemented in MATLAB's fmincon.m function. Appendix D details the model-fitting procedure, but we note here that researchers often use an unconstrained Simplex algorithm (Nelder & Mead, 1965) to fit the DDM to data (e.g., Ratcliff & McKoon, 2008). In contrast, constrained optimization approaches allow a user to restrict parameter values with equality and inequality constraints, including bounding parameters above or below by a constant. As we noted previously, we restricted the extended DDM's additional variability parameters by bounding them above. Table 1 lists the bounds we used during fitting. We examined a range of upper bound values and found that fitted A and T_0 values bottomed out at values near those obtained from a pure DDM fit as the bounds were reduced to the following: 0.04 for s_z, 0.03 for s_A, and 0.08 for s_t. As previously discussed, we chose the bounds in Table 1 (0.03, 0.08, and 0.1, respectively, as well as 0.05 for contaminant proportion x_0) because they appeared to be the variability parameter values that were recovered in the most accurate fits of A and T_0 in the simulated data sets of Ratcliff and Tuer-

linckx (2002; see Figure 6 in that article) and because their simulations used pure DDM parameter values close to those obtained by fits to our data.

Table 1 lists parameter values from a fit to the group data. Following the practice of Ratcliff and colleagues, we set the value of the noise parameter c to 0.1 (c is a "scaling parameter," meaning that multiplying this term by any factor k will produce identical fits by multiplying the other DDM parameters by k—thus, the actual value of c is irrelevant; Ratcliff & Tuerlinckx, 2002). In these simultaneous fits to data from each RSI condition, all parameters other than threshold and starting point were constrained to be equal across RSI conditions. This is consistent with the notion that drift is constant when the DDM is parameterized optimally, and it maximizes the power of the analysis to see changes in threshold. At the same time, it leaves open the possibility that starting points will violate the prediction of being equidistant from the two thresholds. (Furthermore, a separate parametric bootstrap analysis with unconstrained fits showed no significant differences between any parameters other than threshold across conditions.)

Figure 7 shows a graphical comparison between histograms of the empirical data and the appropriately scaled RT densities corresponding to this model fit, separately for correct and error responses (top and bottom rows of plots, respectively). Visually, the match is close. However, model–data mismatches are more visible in quantile probability plots than in density plots, so we superimpose fitted quantile probability plots (X markers) on the empirical plots in Figure 6. Visually, the match in Figure 6 is also close, except in the case of the 500-ms RSI condition, where accuracy is slightly overestimated and the 0.9-quantile RT is significantly underestimated, and in the last two error quantiles for the 1-s RSI condition, where RT is underestimated. These shortcomings can be rectified by leaving the extended DDM's variability parameters completely unconstrained, but this comes at the cost of inflated drift and T_0 estimates. Fit error can be further reduced by allowing all parameters to vary, but this comes at the cost of weakening the power to detect threshold changes across conditions.

Quantitatively, the extended DDM's variability parameters contributed to a large reduction in fit error relative to pure DDM fits (pure DDM fits, which are not listed here, had chi-square fit errors on the order of 1,800, compared with 195 for the constrained, extended DDM). However, these variability parameters were not so obviously large as to rule out application of the optimality theory developed for the pure DDM. To confirm this, we simulated the extended DDM with the fitted parameter values and a range of threshold values to numerically estimate the expected reward rate as a function of threshold. This approximation (plotted as shown later in Figure 10) was close to the function predicted analytically by the pure DDM, with optimal thresholds appearing to be generally smaller than the optimal thresholds for the pure DDM (peaks of the extended DDM's simulated reward rate function are to the left of the peaks of the pure DDM's analytical reward rate function). We discuss this figure in more detail when we compare fitted thresholds to optimal values for the pure DDM later.

Confidence intervals for parameter estimates. To conduct hypothesis tests regarding the adaptation of model parameters across task conditions, we used the parametric bootstrap method (Efron &

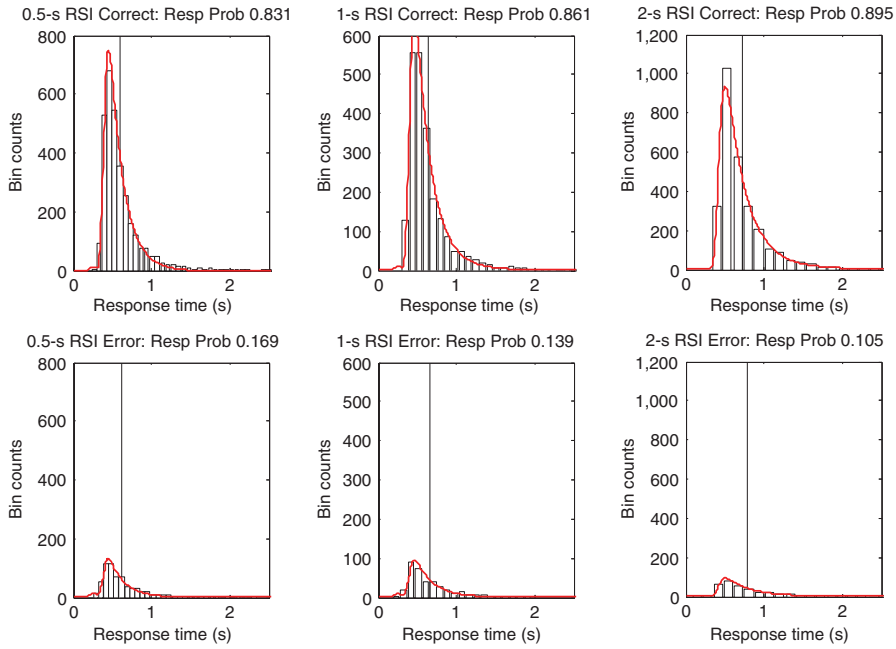

Figure 7. Group response time (RT) histograms and predicted RT densities from a fit of the drift-diffusion model (DDM), Sessions 4–5, Experiment 1. Columns correspond to distinct response–stimulus interval (RSI) conditions. The top row shows RT distributions for correct responses, and the bottom row shows the distributions of error RTs. Vertical lines indicate average RTs in each condition, computed separately for errors and corrects. Histogram bin widths were the same in both the correct and error plots for each RSI and were determined by the Freedman–Diaconis rule (described in Appendix D). Resp Prob = response probability.

Tibshirani, 1993) to construct confidence intervals around the fitted parameter values in each condition.

To test whether thresholds were adapted across conditions, and to determine whether other parameter adaptations were not the primary contributors to SAT adaptation, we generated 300 bootstrap samples of simulated RTs for the parameters obtained by fitting the extended DDM to the pooled RT data. Simulated RTs were generated with the probability integral transform method discussed in Tuerlinckx, Maris, Ratcliff, and De Boeck (2001) and computed in MATLAB with the cumulative RT distribution function CDFDif.m from the Tuerlinckx (2004) study. We then fit each simulated data set and computed nonparametric 95% confidence intervals around the median of the parameter estimates to test the statistical significance of parameter adaptations across RSI conditions.

Figure 8 shows superimposed histograms for the three different threshold estimates. The leftmost histogram corresponds to the 500-ms condition; the middle histogram to the 1-s RSI condition; and the rightmost histogram to the 2-s RSI condition. Whisker-bars plotted at the top of the tallest histogram bins denote 95% percentile confidence intervals for each parameter estimate. They indicate significant differences in the parameter estimates across conditions.

Thresholds and starting points were the only parameters that were allowed to range freely across RSI conditions in this bootstrap analysis. In other fits to group data that allowed all extended-DDM parameters to range freely, only the threshold parameters showed any significant differences across conditions. In contrast, fits to data from some individual participants did appear to show an increase in drift with increasing RSI. Such an increase in drift is inconsistent with prediction 1a. Whether this increase in estimated drift was simply due to correlations between drift, threshold, and residual latency (which showed an increasing trend as RSI increased in individual fits), or whether the SNR for the individual participants concerned actually increased when RSI was longer, is an open question. However, no participants displayed an inverted-U shape in their quantile probability plots and most clearly displayed a U shape. This suggests that, at minimum, thresholds were increasing simultaneously with adaptations in drift across RSI conditions. Fits to individual performance for some participants also suggested that T_0 may have increased as RSI increased. This increase violates Prediction 1b, but again, this may be an artifact of parameter correlations. There is no evidence of T_0 adaptation for the average participant.

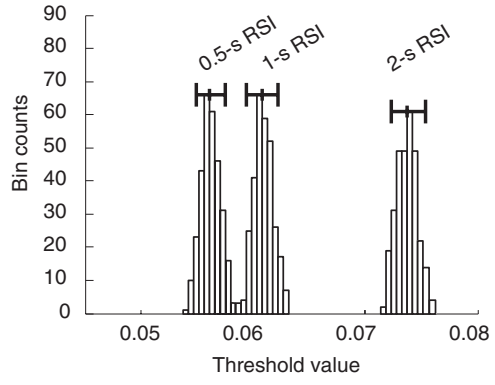

Figure 8. Parametric bootstrap estimates of threshold z, showing significant differences in threshold across conditions. Horizontal whisker lines denote 95% bootstrap confidence intervals around the median threshold value. RSI = response–stimulus interval.

Proximity of fitted thresholds to optimal values. Figure 9 shows fitted thresholds plotted as a function of the optimal thresholds for each condition. Optimal values were computed by numerically solving Equation 5 after substituting fitted drift and residual latency parameters. The best approximation to the optimal threshold occurred in the 2-s RSI condition (the optimal value was within the 95% confidence interval obtained by the parametric bootstrap analysis). The approximation was worse in the 1-s RSI condition and was quite far off in the 500-ms RSI condition. In the latter two cases, thresholds were suboptimally large. This is consistent with previous observations in the literature, which have been interpreted as reflecting an emphasis on accuracy over speed that results in a failure to maximize reward (Maddox & Bohil, 1998).

As we relaxed the upper bounds on the extended DDM's variability parameters during fitting, the fitted values of A and T_0 inflated. Substituting these inflated values into Equation 5 led to decreased values of the predicted optimal threshold, causing fitted thresholds to appear much larger than optimal. It is possible, however, that if participants implement the DDM but cannot control variability in starting point, drift, and T_0, then they may still be able to set thresholds to nearly optimal values for the extended DDM. These values might then only appear to be suboptimal according to an analysis based on Equation 5.

Analytical expressions for reward rate as a function of threshold do not exist for the extended DDM, so we tested this hypothesis by numerically simulating the extended DDM with the parameters from Table 1. The resulting reward rate curves are close to the analytical curves for the pure DDM but appear to have even smaller optimal thresholds (we also did this for a completely unconstrained fit of the extended DDM; results shown in Fig. F1 of Appendix F demonstrate a larger mismatch between fitted and optimal thresholds). The match between simulations of the constrained, extended DDM and analytical results suggests that predictions based on the pure DDM are likely to be useful in practice even if there is some variability in parameters that the pure DDM assumes to be constant.

Figure 10 shows these simulation-based curves along with the analytical reward rate curves for the pure DDM, and it illustrates the efficiency of reward gathering in the different RSI conditions of Experiment 1. Participants were able to achieve 97% of the maximum reward rate in the 500-ms RSI condition, 99% of the maximum in the 1-s RSI condition, and 99.9% in the 2-s RSI condition. Because relative reward harvesting efficiency increases as RSI increases, we speculate that performance might be even closer to optimal with longer RSIs (a 4-s RSI curve is plotted in Figure 10 for comparison).

Figure 10 also shows the effect of anticipations. The fitted threshold and the reward rate earned in each condition are marked with Xs. Blue Xs are based on summing up all rewards and dividing by the duration of blocks of trials. This duration may also include a number of 4-s penalty delays incurred for anticipatory responses. The DDM predictions of Equation 4 do not incorporate these delays, however, so we subtracted the total penalty duration from the block duration in each condition to get a corrected, earned reward rate estimate for comparison with the DDM predictions; these estimates are plotted with red Xs. The differences between the blue and red Xs in each

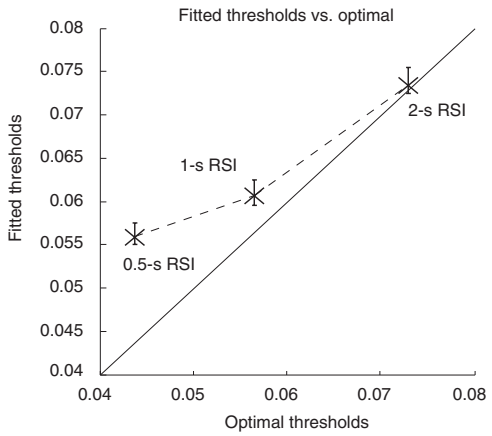

Figure 9. Plot of fitted thresholds versus optimal thresholds. Vertical crossbars indicate 95% confidence intervals around the fitted threshold values plotted as Xs. RSI = response–stimulus interval.

condition, therefore, indicate the proportion of anticipations in each condition, and they demonstrate that the frequency of anticipations decreased dramatically as RSI increased.

Discussion

Consistent with the predictions of an optimally tuned DDM, fits to pooled data from all participants (and to data from individual participants) suggest that threshold values increased with RSI across blocks (Prediction 1d) and that starting points remained equidistant from both thresholds (Prediction 1c). In the case of pooled data, no other parameters were seen to covary with mean RSI (Predictions 1a and 1b). An SAT function relating expected RT and accuracy is also determined by the drift parameter of the DDM, and this function was approximated by the observed SATs in the three RSI conditions. However, both individual participants and the average participant represented by pooled data appeared to set thresholds at values higher than optimal in two of the RSI conditions (violating Prediction 1e).

A possible explanation for suboptimally high thresholds and the suboptimally high accuracy that results is that participants may derive intrinsic value from accuracy itself (Maddox & Bohil, 1998). Another possible explanation for a propensity toward suboptimally high thresholds was proposed in Bogacz et al. (2006). There it was argued that if errors in threshold selection were to occur, then it would be better to err toward higher rather than lower thresholds. This argument derives from the skewed shape of the curve defining reward rate as a function of threshold (see Figures 2A and 10). This skew implies that reward rate decreases more rapidly as thresholds become suboptimally small than they do as thresholds become suboptimally large.

The proportion of anticipatory responses in each RSI condition suggests a third possibility: that participants may need to set thresholds higher than the optimum in conditions where anticipations are more likely. It may be that anticipation becomes a prepotent behavior at high response rates (which, e.g., are much higher in the 500-ms RSI condition than the 2-s RSI condition). If so, then setting thresholds artificially high may reduce the likelihood of anticipation by slowing the response rate, and the need for this slowing should decrease as RSI increases. Consistent with this explanation—or at least with a general impairment of strategic control at short RSIs—several researchers have found that RTs increase and accuracy decreases as RSI decreases below 500 ms (Jentzsch & Dudschig, 2009; Sommer, Leuthold, & Soetens, 1999).

Another curious aspect of the data is that the reward rate curves plotted as a function of threshold in Figure 10 flatten as the RSI increases. Under simple hill-climbing strategies for optimizing thresholds (e.g., Myung & Busemeyer, 1989), this flatness would suggest that deviations from optimal thresholds should be larger as RSIs increase. However, it may be that the amount of reward earned as a proportion of the total possible is the quantity that determines performance (e.g., such proportional judgments have often been proposed to underlie Weber's law for just-noticeable differences in perceptual judgments). If such ratios are what determine performance, then absolute amounts of reward (and flatter maxima of reward rate curves for longer RSIs) are irrelevant. These two factors together—proportional reward rate estimation and performance degradation with increasing task pace—constitute a possible explanation for improvements in performance as RSIs increase.

A fourth possibility is that reward simply does not have as strong an effect as predicted on behavior. It is important to note, however, that the theory of optimal DDM parameterization also predicts dramatic, qualitative changes in behavior in the case of unequally likely stimuli and unequally rewarded responses that result from optimal threshold and starting point shifts. Observing behavior consistent with these predictions would bolster the case for strategic threshold adaptation in Experiment 1. We assess these predictions in Experiments 2 and 3.

Experiment 2

In decision-making tasks involving multiple trials, stimulus ratios provide potentially useful information to the decision maker. When stimuli are unequally likely, a decision maker can exploit estimates of prior probability to improve earnings by favoring the response to the more frequent stimulus (we refer to this response as the favored response and the more likely stimulus as the favored stimulus). Optimizing the pure DDM produces precise, quantitative predictions about how the decision maker should respond to changes in stimulus probabilities (Π and $1 - \Pi$) when stimulus discriminability is held constant. The first two of these predictions are identical to those in Experiment 1, and the remainder are modified to account for unequal stimulus probabilities:

Prediction 2a: Estimates of drift (A) should be constant across all stimulus–probability and RSI conditions.

Prediction 2b: Estimates of residual latency (T_0) should be constant across conditions.

Prediction 2c: Estimates of the starting point x_0 should be shifted toward the favored response threshold as specified by Equation 6, reflecting a bias toward the favored response; the size of the optimal starting-point shift should be independent of the mean RSI.

Prediction 2d: As in Experiment 1, estimates of the threshold parameter (z) should increase as RSI increases, reflecting a shift of the SAT toward greater accuracy; threshold magnitudes in this case should equal the function $z(A, c, T_0, \text{RSI}, \Pi)$

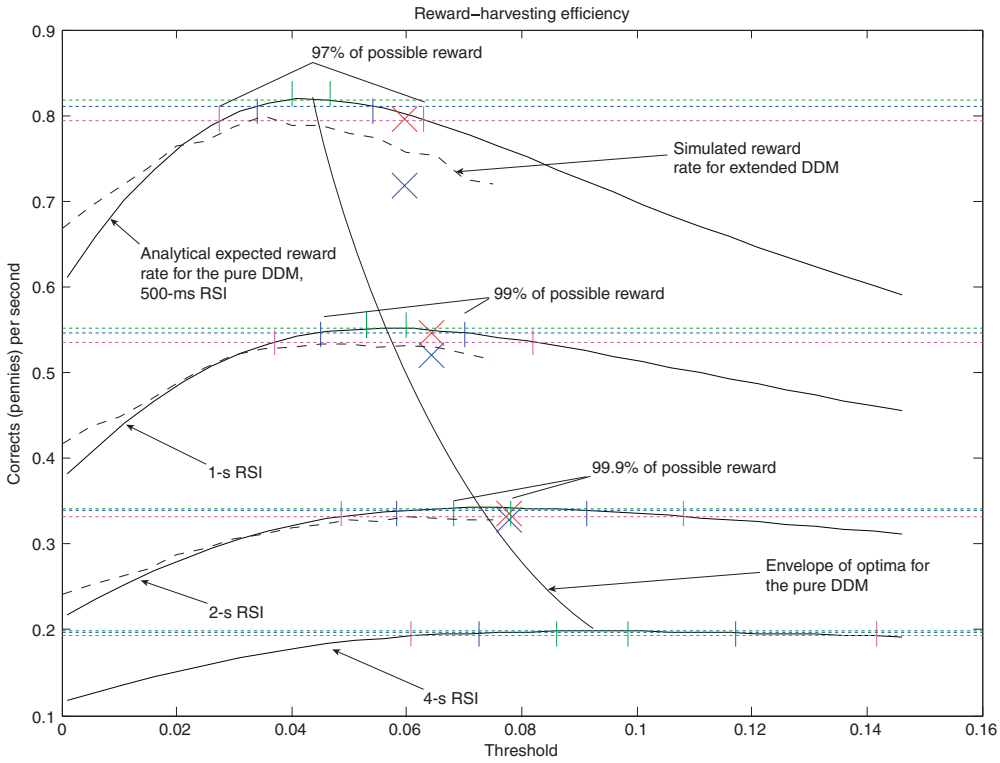

Figure 10. Reward harvesting efficiency of participants in three response–stimulus interval (RSI) conditions. One solid reward-rate curve per RSI condition represents the analytical expected reward rate for the pure drift-diffusion model (DDM) with the A and T_0 values listed in Table 1 and with extended-DDM variability parameters set to 0. Dashed reward-rate curves show the numerical average reward rate for the extended DDM with the nonzero variability parameters listed in Table 1, simulated 10,000 times at 16 different threshold values. Green vertical lines bound intervals within which a threshold setting is expected to produce 99.9% of the maximum reward; blue lines bound 99% intervals, and magenta lines bound 97% intervals. Superimposed on these plots are blue Xs denoting the fitted threshold in each condition and the observed rate of reward in each condition (total of rewards divided by total duration). Red Xs correct for the penalty delays incurred by anticipatory responding, illustrating the larger proportion of anticipations in conditions with a shorter mean RSI.

defined implicitly by Equation 7, evaluated at the current values of RSI and Π and the fitted values of A, c, and T_0.

Prediction 2e: Estimates of the threshold parameter (z) should decrease according to Equation 7 as Π increases; as shown numerically in Bogacz et al. (2006), the optimal threshold decrease should be smaller than the optimal starting-point shift.

Expected reward rate for the pure DDM in Experiment 2 is thus maximized by shifting the starting point of evidence integration (x_0) in the direction of the favored response threshold, by slightly reducing both thresholds, and by leaving drift to be determined entirely by the stimulus. (In contrast, for the extended DDM, it is possible that strategically adapting the mean drift value along with thresholds and starting points across conditions could maximize the expected reward rate.)

A particularly strong prediction of the optimally parameterized DDM is that, for particular combinations of a sufficiently short RSI and a sufficiently asymmetric stimulus ratio, the shift in starting point places it beyond the response threshold for the correct response. At this point, participants should exhibit nonintegrative responding. That is, on every trial, they should make the response corresponding to the more frequent stimulus, with average RT comparable with that observed in an easy signal detection task. Equation 8 expresses this prediction as a function of task conditions (RSI and stimulus probability) that defines the surface depicted in Figure 3. For conditions falling below the surface, participants should exhibit nonintegrative responding. Behavior conforming to these predictions would constitute strong support both for the DDM and for the hypothesis that participants adjust the parameters of their decision processes to maximize reward rate. To test these quantitative predictions, we conducted an experiment that was similar to Experiment 1 but that also involved manipulating the probabilities of the two stimuli in addition to the RSI.

Method

Participants. Participants were the same as in Experiment 1. They had completed the five sessions of Experiment 1 before the five sessions constituting this experiment.

Apparatus and stimuli. Apparatus and stimuli were identical to those in Experiment 1.

Procedure. Participants engaged in five 1-hr task sessions consisting of blocks of trials in which one stimulus (one direction of coherent motion) was more likely than the other. Within each block of trials, the direction of motion chosen to be more likely was selected randomly and with equal probability. Participants were informed that the stimulus probabilities, in addition to the RSI, might be different in different blocks. They were once again informed that blocks would always last 4 min; therefore, faster responding would produce more trials overall. They were encouraged to earn as much money as possible.

Each session consisted of one practice block of 2–4 min (practice was reduced in later sessions), followed by twelve 4-min blocks, within each of which a given set of task parameters was held constant. The task parameters were the RSI and the proportions of leftward and rightward stimuli (equivalently, the prior probability Π of the favored stimulus). For each participant, motion coherence was set to the same value as in Sessions 4 and 5 of Experiment 1. As in Experiment 1, the actual RSI on a given trial was jittered around the average value with a standard deviation of 100 ms to discourage anticipations. A 4-s penalty delay between trials was again enforced whenever responses occurred before 100 ms after stimulus presentation. RSI and Π were factorially covaried, with RSI taking values of 500 ms, 1 s, or 2 s, and Π taking values of 0.6, 0.75, or 0.9. The order of conditions was counterbalanced across sessions and across participants with a Latin square design. Two consecutive blocks of trials were allocated to each condition in which RSI was 2 s, because a 2-s RSI produced far fewer trials within a 4-min block than did an RSI of 1 s or 0.5 s. Finally, the 12 blocks in each session were followed by two 2-min blocks of a signal detection task identical to that of Experiment 1.

Analysis. To assess predictions, we examined in detail the performance of the average participant, represented by the pooled data from all participants. (Data for an individual participant are presented in Appendix E.) Because estimates of A and T_0 were all that were required to make behavioral predictions, we were able to base our predictions in Experiment 2 entirely on a fit to the data from Experiment 1. Estimates of A and T_0 were used to predict the optimal threshold z and starting point x_0 on the basis of Equations 5 and 6, respectively (c was assumed to be 0.1, as noted previously). These values of A, T_0, x_0, and z (and the values of the variability parameters s_t, s_x, and s_A derived from extended DDM fits) in turn predicted a specific RT, accuracy, and proportion of right versus left responses as a function of mean RSI and stimulus probability in the various conditions of Experiment 2.

We also fit the data of Experiment 2 itself simultaneously with the data from Experiment 1 (these simultaneous fits are the ones listed in Table 1 and Table 2), and the critical A and T_0 parameters were within 8% of the values found in fits to the data from Experiment 1 alone. However, fitting in this experiment was complicated by the stimulus–probability manipulation. Although the data conformed to our prediction of nonintegrative responding when RSI was small and Π was large, the resulting RT distributions (both for pooled data and for individual participants) were bimodal, or showed hints of bimodality, in most conditions. Bimodality appeared to result within participants from runs of nonintegrative trials interspersed with runs of integrative trials (see Figure E2), as well as in the pooled data from integration by some participants and nonintegration by others in some conditions. Because the DDM with a single set of parameters cannot predict a bimodal RT distribution, this made fitting the DDM to data from Experiment 2 effectively impossible when stimulus ratios were greater than 60:40 and RSI was less than 1 s.[7]

Table 2
Fitted Parameter Values for the Average Participant (Pooled Data, All Participants)

Parameter	Value		Bound
Drift (A)	0.17348		
Noise coefficient (c)	0.1		
Residual latency (T_0)	345.47 ms[a]		
Drift standard deviation (s_A)	0.068683		0.08
Starting point range (s_x)	0.03		0.03
T_0 range (s_t)	0.1		0.10
Contaminant proportion (p_0)	0.029182		0.05
Fast-guess RT			
M	266.77 ms		290
SD	49.099 ms		100
Total χ^2 fit error	1,020.104		
Stimulus probability ratio	60:40	75:25	90:10
Threshold z			
RSI: 500 ms	0.049545	0.055035	0.062455
1 s	0.052885	0.052220	0.059620
2 s	0.062745	0.062890	0.064055
Optimal threshold			
RSI: 500 ms	0.043370	0.040829	0.033304
1 s	0.056341	0.054623	0.049731
2 s	0.072819	0.071804	0.069020
Starting point x_0			
RSI: 500 ms	0.000915	0.012183	0.021868
1 s	0.004557	0.010700	0.011437
2 s	0.003909	0.005578	0.011804
Optimal threshold			
RSI: 500 ms	0.011686	0.031660	0.063330
1 s	0.011686	0.031660	0.063330
2 s	0.011686	0.031660	0.063330
Mixture weight			
RSI: 500 ms	0.079295	0.481880	0.841250
1 s	0.031668	0.224710	0.700610
2 s	0.016402	0.135260	0.454670
Drift increment			
RSI: 500 ms	0.015360	0.003770	0.000000
1 s	0.003040	0.000100	0.083964
2 s	0.000000	0.035220	0.057137
χ^2 fit error			
RSI: 500 ms	92.91	110.90	297.86
1 s	54.31	30.06	112.14
2 s	97.53	74.12	150.27

Note. Data from Experiments 1 and 2 were fit simultaneously, leading to parameter values identical to those in Table 1 for all parameters other than threshold, starting point, mixture weight, and drift increment. RSI = response–stimulus interval.
[a] Compare to average signal detection response time (RT) of 301 ms.

[7] In contrast to our difficulty in fitting the DDM to data produced by unequally likely stimuli, Ratcliff and McKoon (2008) were able to fit data reliably with RSI values comparable to our fastest condition and with a stimulus ratio of 75:25. However, their experiment involved several differences in design: response deadline bands, with "too fast" and "too slow" messages for feedback along with correct/error feedback; course credit for undergraduates as payment instead of payment for correct responses; and explicit instruction about the stimulus proportions within each block. Nevertheless, SAT adjustment was observed and was qualitatively consistent with the predicted starting point and threshold adjustments of Bogacz et al. (2006).

Although a mixture of integrative and nonintegrative responding is not predicted by an optimally parameterized, pure DDM, this result should be expected if there is variability in the model's parameters from trial to trial; this is precisely what is assumed in the extended DDM. To fit the data from Experiment 2, we therefore fit a model that was a mixture of a nonintegrative or fast-guess distribution (consisting only of guesses that the more likely stimulus was present) together with an RT distribution generated by the DDM. Because the fast responses that were made in the signal detection blocks at the end of each session appeared almost normally distributed, we modeled the nonintegrative mixture component as coming from a normal distribution.[8]

We also fit a model that allowed an increment to be added to the drift term; in this way, response biasing could be achieved by increasing drift toward the more likely response threshold, no matter which stimulus was presented. This is equivalent to changing the reference point in the one-dimensional stimulus space that determines a drift value of 0 (see Ratcliff, 1985, for discussion of how the 0-point of drift relates to the criterion parameter of signal detection theory). This type of model has been successfully fit to monkey behavioral and neurophysiological data in tasks that vary signal discriminability from trial to trial (e.g., Yang et al., 2005). Adapting the average drift across conditions may also be the optimal strategy in tasks with constant discriminability if the variability parameters of the extended DDM are large enough and if participants cannot act to reduce this variability below a given level; the current lack of analytical results for the extended DDM makes this result (or its opposite) difficult to prove. Empirically, however, including a drift increment term that can vary across conditions allowed us to test whether human participants can be modeled as adapting drift across conditions when signal discriminability is constant from trial to trial (a circumstance in which optimal performance, in contrast, requires a pure DDM and no drift adaptation).

Results

Quantile probability plots. The top row of panels in Figure 11 displays the quantile probability plots for trials in which the favored stimulus is presented; the bottom row displays the quantile probability plot for the unfavored-stimulus trials.

In the superimposed scatterplot of RTs, correct response RTs are plotted in green and error RTs are plotted in red. This makes visible the shift of error and correct RT probabilities in response to unfavored stimuli as Π increases (see bottom row of panels in Figure 11). This approach also highlights the occurrence and relative frequency of anticipatory responding across conditions.

For a stimulus ratio of 60:40 ($\Pi = 0.6$), quantile probability plots (shown in the leftmost column of Figure 11) continue to show SAT adaptation of the type shown in Experiment 1: For both types of stimuli, the plots retain roughly the U shape seen in Figure 6, consistent with Prediction 2d. In response to unfavored stimuli, accuracy decreased (as indicated by the shift of the quantile columns toward the middle of the graph). In response to favored stimuli, correct responses in a given condition tended to be faster than error responses. Conversely, in response to unfavored stimuli, error responses were typically faster than correct responses.

For a stimulus ratio of 75:25, performance resulted in quantile probability plots with radically different shapes (middle column of Figure 11). Accuracy in response to favored stimuli increased markedly relative to the 60:40 condition, moving correct quantile columns to the right edge of the plot and error columns to the left. Unfavored stimuli, in contrast, produced quantile columns that are shifted further toward the center, and correct responses became less likely than errors when RSI was 500 ms. Furthermore, errors were much faster than correct responses to the unfavored stimulus, and this asymmetry in RT was more exaggerated for shorter RSIs. Both of these phenomena are consistent with an optimally tuned DDM, in which threshold magnitudes decrease as RSI decreases, and the starting point moves closer to the response threshold for the favored stimulus as Π increases. Similar starting point shifts and relative constancy of drift (but not anticipatory responding) were observed by Ratcliff and McKoon (2008) in their investigation of stimulus probability effects in a two-alternative motion discrimination task that was similar to Experiment 2 but used a fixed RSI and response deadline bands.

The quantile probability plots in the 90:10 conditions show more exaggerated versions of the patterns in the 75:25 conditions. The rightmost column of panels in Figure 11 shows favored stimulus quantile columns pushed even farther to the extreme right and left of the plot than in the 75:25 case. Overall correct RT was also faster, and error RT was slower, than in the 75:25 case. In response to unfavored stimuli in the 90:10 conditions, correct responses were less likely than error responses for all values of RSI. These results are also consistent with an optimally tuned DDM, in which the starting point is shifted near or beyond the response threshold for the favored stimulus.

Model fits. Table 2 lists the parameters estimated by fitting a mixture model consisting of a nonintegrative component (a normal distribution) and an integrative component (an extended DDM first-passage time distribution). Figure 12 shows predicted RT densities based on these fits superimposed on the empirical histograms; correct and error RTs to favored and unfavored stimuli are plotted separately. The first feature to notice in these plots is that the proportion of correct responses to unfavored stimuli decreases as RSI decreases and as Π increases, which is consistent with the increase in response bias predicted by the optimally parameterized DDM (Prediction 2c).

Also noteworthy in these plots are the clear signs of bimodality in the favored-correct panels of all conditions (except the 60:40 stimulus-ratio/2-s RSI condition), with the earlier peak decreasing in height relative to the later peak as RSI increases and as stimulus ratios approach 50:50. In all 90:10 stimulus-ratio conditions and in the 500-ms RSI/75:25 stimulus-ratio condition, the narrower, faster, nonintegrative component of the bimodal mixture had a larger RT density peak than the slower integrative component in

[8] The DDM with a single absorbing boundary, which has previously been used to model simple reaction times (Pacut, 1977), might be a suitable model for nonintegrative responses. The Wald distribution describes this model's first-passage times (Luce, 1986), but this approximates a normal distribution when drift is large, as we should expect for the highly salient signals in our tasks. There is also reason to suppose that Grice's (1968) deterministic accumulation model with random thresholds might be a good model for such simple RTs, and the RT distribution for this model is exactly the normal distribution.

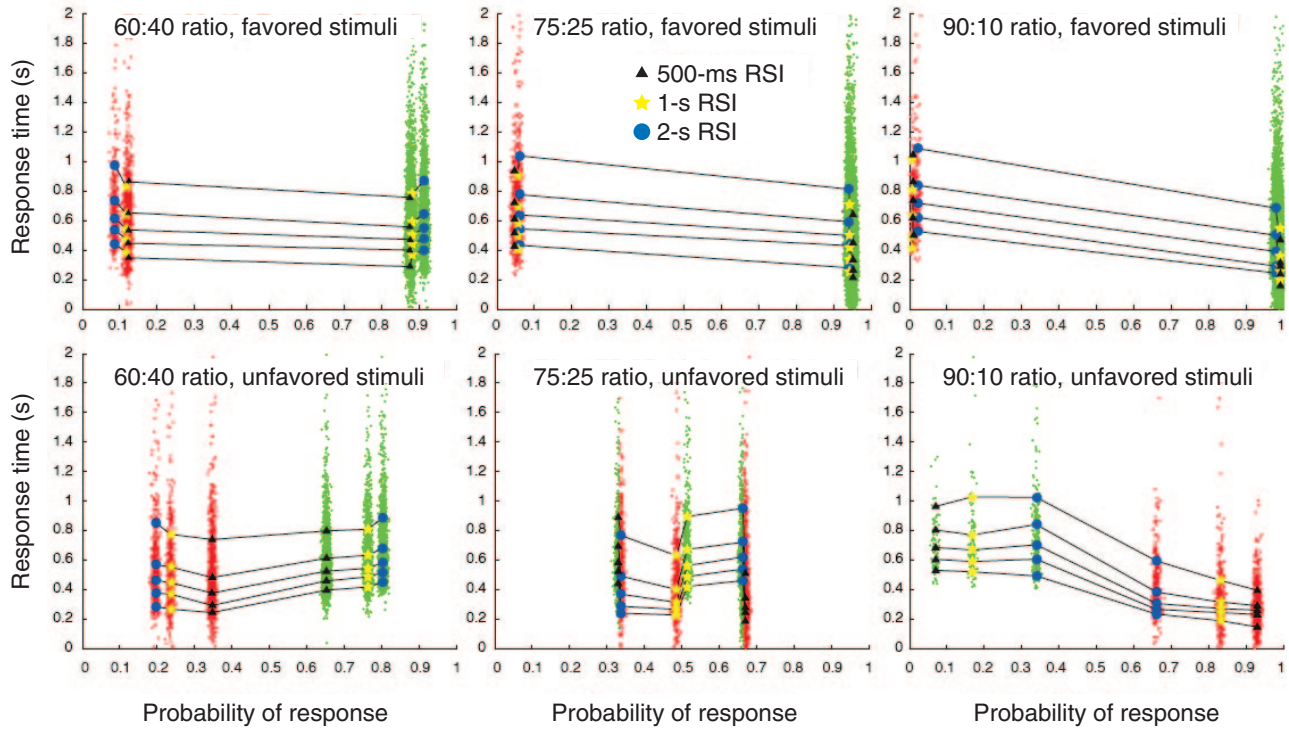

Figure 11. Quantile probability plots for all conditions of Experiment 2. Superimposed scatterplots of response time data are plotted in green for correct responses and red for errors. The left column shows quantile probability plots for a 60:40 stimulus ratio, the middle column shows quantile probability plots for a 75:25 stimulus ratio, and the right column shows quantile probability plots for a 90:10 stimulus ratio. The top row of panels shows quantile probability plots for responses to the more likely stimulus. The bottom row plots responses to the less likely stimulus; note the exchange of correct and error probabilities as stimulus-ratio asymmetry increases. RSI = response–stimulus interval.

the favored, correct RT panel (upper right corner of the 2 × 2 plot panels in Figure 12). Nonintegrative peaks that were smaller than the integrative peaks occurred in the 75:25 stimulus-ratio conditions with RSI equal to 1 s and 2 s, and in all 60:40 stimulus-ratio conditions.

Table 2 lists the fitted mixture weights on the nonintegrative component (with integrative weights equal to 1 minus the nonintegrative weights). These weights are greater than 0.5 in the 90:10 stimulus-ratio conditions with RSI equal to 500 ms and 1 s, and they decrease in the same condition order as the nonintegrative RT density peaks. Qualitatively, this is the pattern of shifting relative weights on the two mixture components that should be expected for a model that approximates an optimally parameterized DDM but which has variability in its parameters across trials (and may therefore switch between integrative and nonintegrative responding as the starting point crosses back and forth across the favored response threshold from trial to trial).

Extended DDM parameter estimates for the integrative mixture component are also listed in Table 2.[9] In contrast to Experiment 1, thresholds were frequently smaller than the optimal value for the pure DDM (see Figure 13), except, again, in the 500-ms RSI conditions and in the 90:10 stimulus-ratio conditions with the shortest RSIs. However, these estimates may become less reliable as the mixture weight on the nonintegrative component becomes large, because of the relatively small number of integrative responses in these conditions and possibly because the fast RTs may not be properly apportioned to the two mixture components. In the 90:10 conditions, for example, thresholds do not appear to be modulated at all across conditions (in contrast, the effect of RSI on the mixture weight is enormous for these 90:10 conditions).

Starting points were similarly smaller than pure DDM optimal in all cases. Figure 14 shows fitted starting points, normalized by the distance from the lower threshold to the upper threshold, plotted against similarly normalized optimal starting points. In this figure, values greater than 0.5 imply a bias toward upper threshold crossings. As predicted, starting points increased as Π increased, but not as much as predicted. Note, however, that many of the data points come from mixture component fits in conditions in which responses were primarily nonintegrative and were predicted to be so (all data points were to the right of the vertical line, indicating

[9] Data from Experiments 1 and 2 were fit simultaneously, leading to identical values for A, T_0, s_A, s_x, s_t and p_0 in both experiments. Drift-increment and mixture weight parameters were not significantly different from 0 in all conditions of Experiment 1, and fits to data from Experiment 1 (data not listed) that constrained these parameters to 0 led to very similar values of A, T_0, s_A, s_x, s_t and p_0.

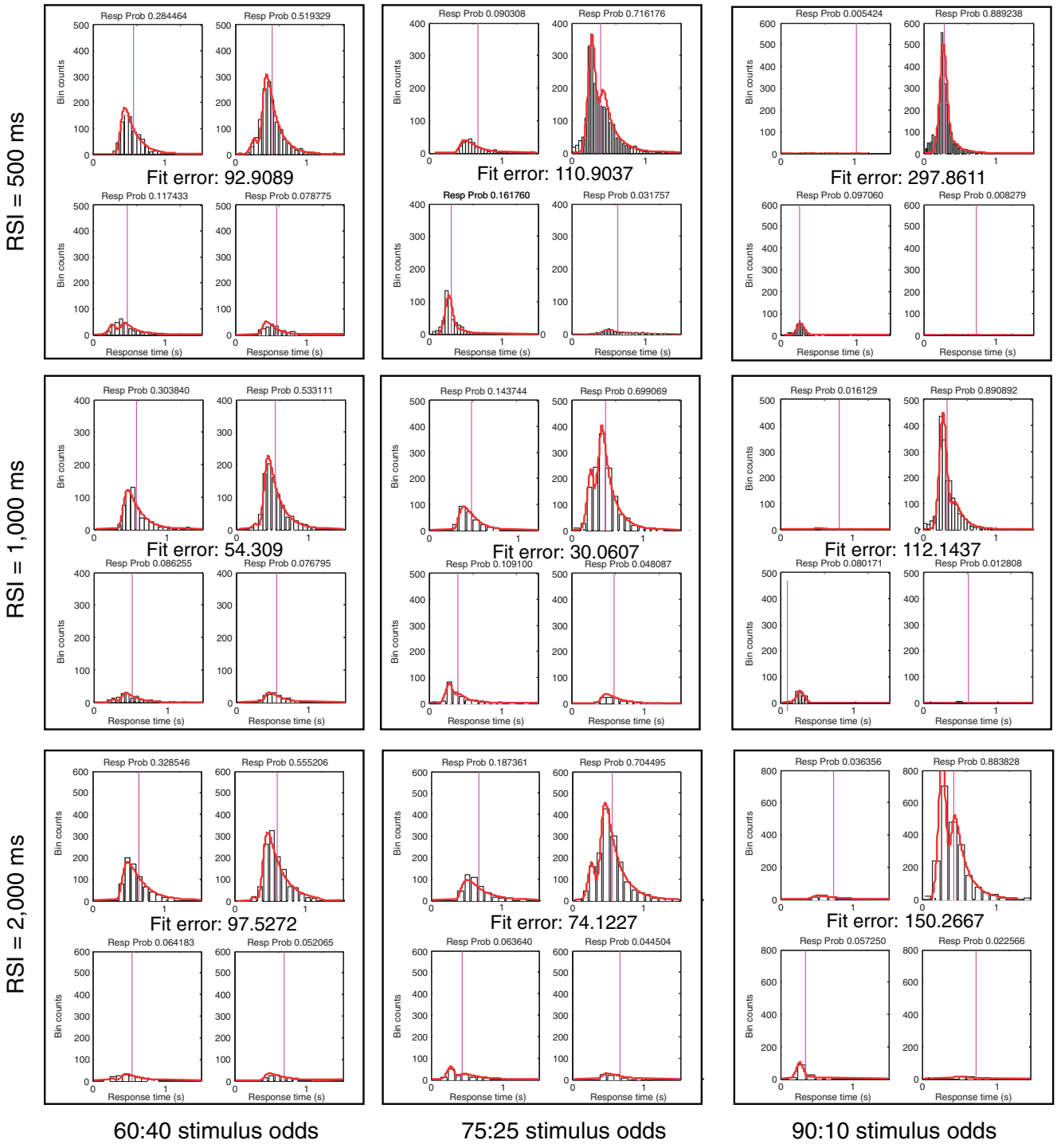

Figure 12. Fits to response time (RT) distributions in Experiment 2. Each response–stimulus interval (RSI)/stimulus–probability condition is represented by a panel consisting of a 2 × 2 set of four plots: RTs for correct responses to favored stimuli (upper right of panel); correct responses to unfavored stimuli (upper left); errors for favored stimuli (lower right); and errors for unfavored stimuli (lower left). Three columns of these 2 × 2 plot panels correspond to three stimulus probability conditions—60:40, 75:25, and 90:10 stimulus odds—and three rows correspond to three RSI conditions: 500 ms, 1 s, and 2 s. Resp. Prob = response probability.

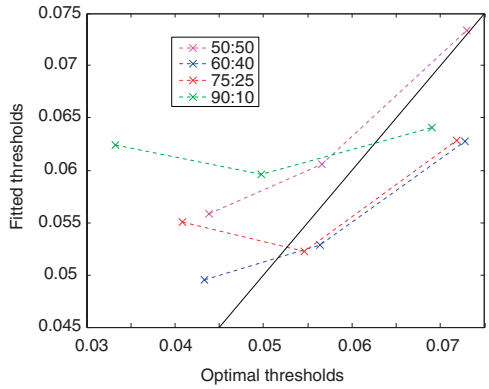

Figure 13. Comparison of fitted thresholds to optimal thresholds; key identifies different stimulus–probability conditions, and the black identity line indicates what would be a perfect match. The 50:50 ratio data are from Experiment 1.

starting points that exceed the upper threshold). A possible explanation for the difference in both starting point and threshold patterns between Experiments 1 and 2 derives from the fact that participants know that Experiment 2 involves manipulations of stimulus probability and that Experiment 1 does not. Thus, they may consciously try to develop a response bias in Experiment 2 that involves both lowering thresholds and shifting starting points.

Finally, the drift increment value did not show the systematic pattern across conditions that would be expected for a parameter that was strategically adapted to produce a response bias (i.e., growing as Π increased and perhaps as RSI decreased). It was significantly different from 0 in four conditions, but two of these involved 90:10 stimulus ratios and, thus, a relatively small number of integrative responses. Therefore, there appears to be little evidence of strategic drift adaptation across conditions.

Comparing decision RT densities to the signal detection RT density. Another test of the hypothesis that RT distributions are a mixture of integrative and nonintegrative responding is to compare the RT distribution for a given condition of discrimination trials with the RT distribution for signal detection trials (signal detection blocks occurred at the end of each session of Experiments 1 and 2). We predicted that these distributions would be comparable in conditions eliciting nonintegrative responding.

The three panels of Figure 15 compare empirical decision RT densities with the signal detection RT density (these are Gaussian-kernel-smoothed densities of empirical data rather than predicted densities based on parameter fits; using smoothed densities rather than histograms makes it easier to superimpose data plots from multiple conditions). Within each panel, decision RT densities are plotted for a single RSI and the full range of stimulus ratios. Data from the 90:10 ratio conditions show how nonintegrative responding created relatively peaked densities in the 0.5-s and 1-s RSI conditions, with peaks located near the peak of the signal detection RT density. In contrast, the densities for the 60:40 ratio conditions, and for the 50:50 conditions of Experiment 1, are located about 200 ms to the right. The 75:25 ratio conditions are particularly interesting: They have a wide spread, providing a transitional form between the 60:40 densities and the 90:10 densities, consistent with a mixture between integrative and nonintegrative responding.

These plots show a very clear pattern: As RSI decreases and Π increases, decision RT densities develop a second mode with the same location as the signal detection density. This mode increases in amplitude while the mode located farther to the right (closer to

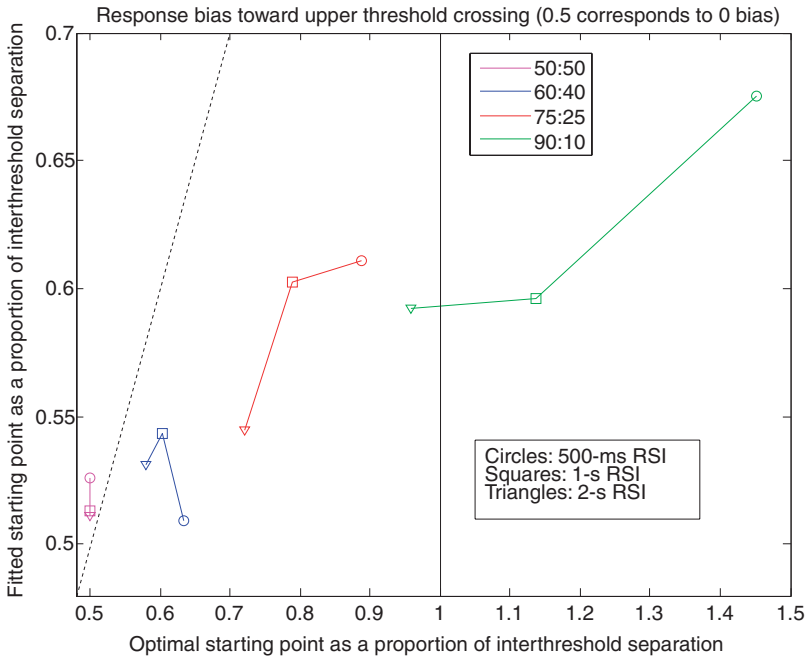

Figure 14. Comparison of fitted starting points to optimal starting points; key identifies different stimulus–probability conditions. The 50:50 data are from Experiment 1.

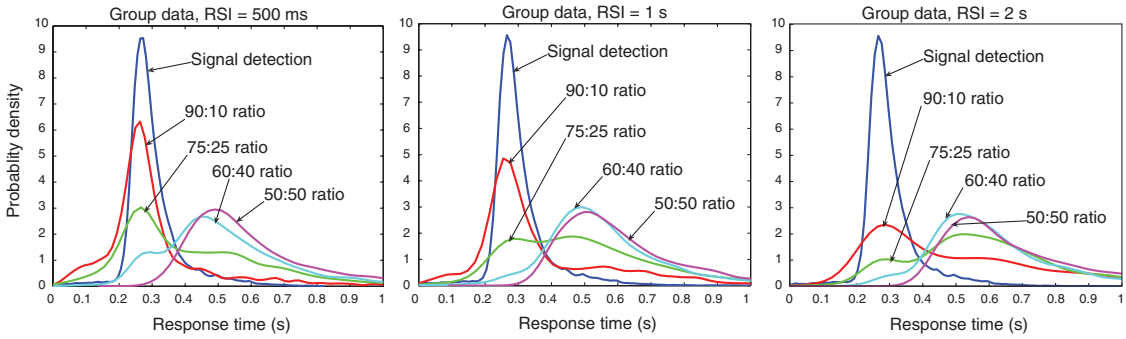

Figure 15. Comparison of signal detection response time (RT) density with RT densities in all two-alternative forced-choice (2AFC) conditions of Experiment 2. Mean response–stimulus intervals (RSIs) are 500 ms, 1 s, and 2 s in the left, middle, and right panels, respectively. As the stimulus–ratio asymmetry increases, the RT density for two-alternative decisions approaches that for the signal detection condition. This change in the RT density becomes more pronounced as the mean RSI decreases. In addition, a bimodal density appears for conditions with unequally likely stimuli, suggesting a mixture of integrative and nonintegrative responding. The nonintegrative modes increase in amplitude (and the integrative modes decrease) as the asymmetry in stimulus ratios increases and as the mean RSI decreases.

the 50:50-ratio/2-s RSI location) decreases in amplitude. Ultimately, the density becomes unimodal and very similar to the signal detection density (the main difference being the presence of anticipatory responses indicated by a tail to the left of the signal detection density).

Response proportions, RT, and accuracy. We now compare the observed response proportions, RTs, and error percentages to their predicted values, given fitted values of A and T_0. As Π increases, maximizing reward rate should cause a bias toward the favored response to develop as a result of starting point shifts (determined by Equation 6); as RSI decreases, this bias should at some point cause nonintegrative responding; that is, exclusive choice of the favored response, with RTs that are comparable with signal detection RTs. The particular values of RSI and Π that are predicted to produce this nonintegrative responding (i.e., solutions for z of Equation 5 that equal the x_0 values predicted by Equation 6) depend on fitted values of drift A and residual latency T_0, which may differ from participant to participant.

If correlations among parameter estimates or wide confidence intervals around them make it difficult to assess whether thresholds and starting points are near their optimal values, then comparing such qualitative features of observed behavior with the same features of behavior predicted by an optimally tuned DDM can help answer this question. Recall also that the conditions for optimality discussed in Bogacz et al. (2006) depend on the pure DDM. Because the extended DDM was far easier to fit to the data, and because variability parameters in these fits tended to be far from 0, our predictions regarding optimal threshold and starting point values are only approximations to optimal tuning for the extended DDM (although our extended DDM simulations suggest that these approximations may be reasonably accurate). Thus, an examination of qualitative features of behavior may be particularly helpful.

Figure 16 illustrates the comparison of observed average RT, accuracy, and response proportions in all conditions of Experiments 1 and 2 with predictions based on fitted A and T_0 values. The top row of plots shows a close match between the response proportions predicted by the pure DDM (heavy dashed line) and the proportions observed (solid line). For comparison, a thin dashed line depicts the predictions of a simple probability-matching hypothesis, which specifies that response proportions should equal stimulus proportions; this alternative hypothesis is not well supported by the data. Good matches also occurred between the predicted and observed RT averages (middle row of plots). Quantitatively, the match between the predicted and observed error percentages (bottom row of plots) is not as close in the 60:40 and 75:25 stimulus-ratio columns, but the overall shape of the error percentage curves is reflected in the observations, and there was a decrease in the average magnitude of these observed percentages as RSI increased, as predicted. The proportion asymmetry Π defining the critical probability surface (the point at which a transition to nonintegrative responding is predicted to occur) is plotted as a thick vertical line in all conditions where it falls within the corresponding plot's x-axis limits. As predicted, the average participant approached nonintegrative responding in the 90:10 ratio conditions but appeared to achieve this type of behavior fully only when both the ratios were 90:10 and RSI was either 0.5 s or 1 s.

Discussion

Evidence from Experiment 2 provides support for a nearly optimally tuned DDM as a model of decision making in this task. Empirical response proportions closely matched the proportions predicted by an optimally tuned DDM in the case of pooled data from all participants (Figure 16). RT and accuracy data also qualitatively matched the shape of the predicted RT and accuracy curves plotted in each RSI condition in Figure 16 (similar results for an individual participant are given in Appendix E). Parameter fits showed clear threshold and starting point shifts in the expected directions across conditions (Figures 13–14), although these parameters often deviated from their optimal values (especially in those conditions with a large proportion of nonintegrative re-

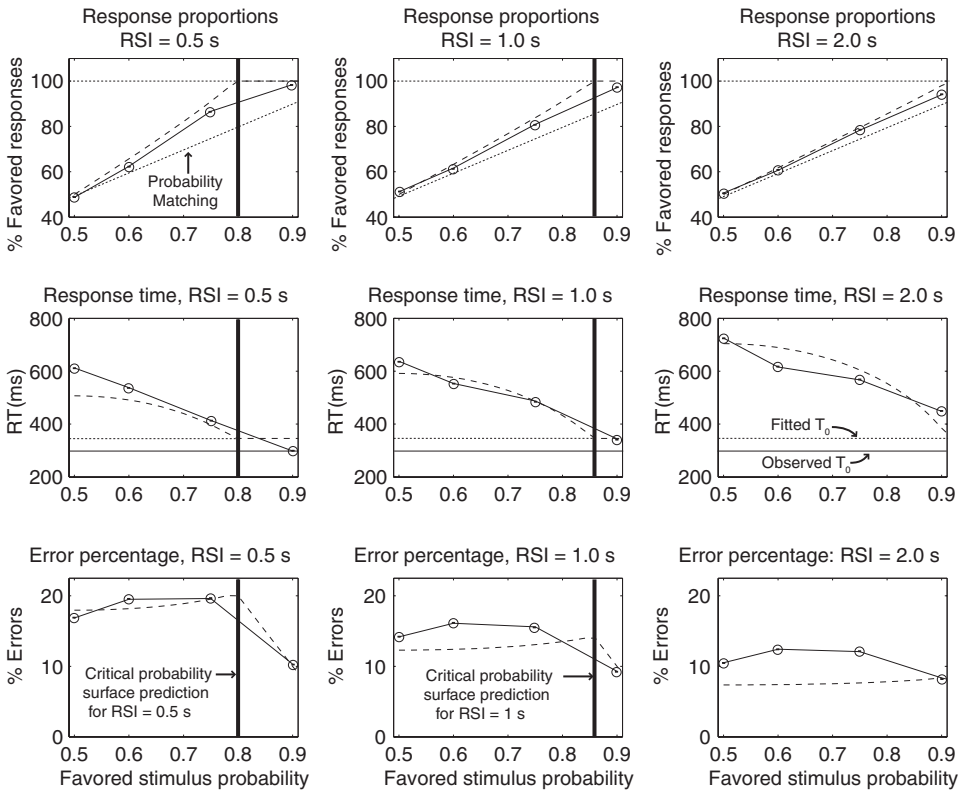

Figure 16. Comparison of predicted and observed response proportions (top row of plots), response times (RTs) (middle row) and error percentages (bottom row) in all conditions of Experiment 2, based on fits of drift and residual latency. The horizontal axis in each plot denotes the stimulus proportions (0.6 indicates a 60:40 ratio, 0.75 indicates a 75:25 ratio, and 0.9 indicates a 90:10 ratio). The left column of plots corresponds to a mean response–stimulus interval (RSI) of 500 ms, the middle column corresponds to a mean RSI of 1 s, and the right column corresponds to a mean RSI of 2 s. Standard error bars are plotted, but are barely visible, in all plots.

sponses, which may make parameter estimation more imprecise for the remaining proportion of integrative responses).

Examining the RT densities as task conditions changed in Figures 12 and 15 provided a clear picture of the way in which RT distributions were transformed as the stimulus–ratio asymmetry increased and the mean RSI decreased: Unimodal integrative RT densities took on a transitional bimodal shape, followed by a unimodal, nonintegrative density shape that was very similar to the density for signal detection responses.

Experiment 3

The theory of optimal decision making applies also to the case in which the two responses are not equally rewarded (i.e., a proportion r of some unit of reward is assigned to one response when correct, and $1 - r$ is assigned to the other). The assumption of optimality when $r = 0.5$ leads to specific predicted values for the starting point x_0 and the threshold z, and corresponding predictions regarding speed, accuracy, response bias, and a shift to nonintegrative responding:

Prediction 3a: Estimates of the starting point should be greater than 0 (i.e., closer to the threshold corresponding to the more rewarded response); starting point should be shifted from 0 into the range defined by Equations 9–10, which are approximations analogous to Equations 6 for unequal stimulus proportions.

Prediction 3b: As a consequence, the decision maker should choose the more rewarded response more frequently than the alternative, and the average RT should be shorter and the accuracy lower for that response. (Qualitatively similar predictions were borne out in a study by Voss et al., 2004.)

Prediction 3c: With other task factors held constant, numerical results in Bogacz et al. (2006) show that thresholds should decrease more dramatically than in Experiment 2 for equal values of Π and r as r increases.

Prediction 3d: As in the case of unequal stimulus proportions, sufficiently large reward asymmetries and sufficiently short RSIs should shift the starting point beyond the threshold for the more rewarded response, implying the existence of a critical reward–ratio surface. Numerical results (Bogacz et al., 2006) show that this surface is similar in shape to the critical probability surface in Figure 3 but that it predicts nonintegrative responding at smaller values of r than of Π, all other parameters being equal. (In contrast to the relative reward

ratio, the absolute magnitude of the reward scale is predicted to have no effect on behavior.)

To test these predictions, we attempted to leverage the results of Experiment 2 to develop a task involving a single RSI, a single motion coherence, and a single reward asymmetry that would define a point near the critical reward–ratio surface for most participants. Ideally, some participants would lie on one side of the surface, and the remainder would lie on the other because of individual differences in the acuity of motion perception (modeled as A/c).

Method

Participants. Fifteen participants, ranging in age from 18 to 27 ($M = 22$), were recruited through the paid experiments website of the Department of Psychology, Princeton University. None of these individuals participated in Experiment 1 or 2.

Stimulus and apparatus. Apparatus and stimuli were identical to those used in Experiments 1 and 2.

Procedure. Participants engaged in a single 50-min session in which leftward and rightward motion stimuli were presented with equal probabilities. Coherence was set to 10%. The session consisted of one 4-min practice block followed by twelve 3-min blocks with self-paced rest periods in between. RSI was constant across blocks and equal to 1 s. Participants received 3 cents for every correct response on one key (Z or M, counterbalanced across participants) and 1 cent for a correct response on the other key (i.e., the reward ratio was set to 3:1). They earned nothing for incorrect responses. Participants were informed of the score that they earned (3, 1 or 0 cents) after each trial. They were not explicitly informed that one response would be rewarded more than the other when correct. Participants were paid the total amount accrued during the experimental session or $10, whichever was higher (all the participants earned more than $10).

In this experiment, the mean RSI for each trial was the sum of a fixed 300-ms interval plus an exponentially distributed delay with a mean of 700 ms (truncated at 1.91 s) to discourage anticipatory responding. As in Experiments 1 and 2, a penalty delay was enforced whenever a response was made less than 100 ms after stimulus onset to discourage anticipations. Participants were also informed that the block durations were fixed so that faster responding would lead to more trials. They were once again encouraged to earn as much as possible.

Results

Response proportions, RT, and accuracy. Consistent with Prediction 3b, 12 of 15 participants chose the favored response more frequently than the unfavored response (a one-tailed binomial test yields $p = .018$). Consistent with Prediction 3d, 4 of the 15 participants chose the favored response almost exclusively, in proportions greater than 0.90. The remaining 11 participants had proportions in the range between 0.46 and 0.63. This pattern suggests that 4 of the participants performed the task mostly in nonintegrative mode, whereas the rest performed mainly in integrative mode.

Consistent with Prediction 3b, median RTs were significantly smaller for the favored than for the unfavored response, $t(14) = 2.79$, $p = .014$; see Maddox and Bohil (1998) for similar results with fixed viewing times. The difference remained significant even after removing the four nonintegrative participants and the first block of trials (discussed earlier) of the remaining participants, $t(10) = 2.43$, $p = .035$). For those participants who performed the task in integrative mode, the error percentages for the favored response were, on average, higher than those for the unfavored response, although the difference did not reach statistical significance, possibly because of the small sample size, $t(10) = 1.74$, $p = .11$.

Model fits. After we removed the 4 nonintegrative participants and the first block of trials of the remaining participants, the pooled data was amenable to fitting by the extended DDM (although fit errors were higher than for the participants in Experiment 1, who had much more practice). Parameter values for the fits are presented in Table 3. The fits to the pooled data from Experiment 3 were computed in MATLAB using the Diffusion Model Analysis Toolbox (DMAT) software (Vandekerckhove & Tuerlinckx, 2007a, 2007b). We explored the use of DMAT to corroborate the model-fitting performance of our own software and got similar results; however, with a smaller data set and less practiced participants, we got substantially larger fit errors for a single condition, which our constrained approach would make even worse. We therefore report the results obtained with DMAT.

The fitted value of the starting point x_0 was 0.0063. Consistent with Prediction 3a, the starting point is shifted toward the more rewarded response threshold. To test whether the shift was statistically significant, we computed a 95% confidence interval for the starting point by generating 1,000 bootstrap samples using the parametric bootstrap method implemented in DMAT. The confidence interval (0.0055, 0.0071) does not contain 0, indicating a significant shift of x_0 away from the point of zero response bias. The optimal range of x_0 values (0.0148, 0.0295) was obtained by substituting the fitted value of A into Equations 9–10 (Prediction 3a). The fitted value of x_0 is too small to be optimal, at about half the value of the lower limit of the interval. Thresholds were suboptimally large, violating Prediction 3c, because they would be large even for the comparable 1-s RSI condition of Experiment 1 (i.e., with $r = .5$), given that fitted drift values are similar in both experiments.

Decision versus signal detection. Similar to the results of Experiment 2, the fact that some participants exhibited noninte-

Table 3
Fitted Parameter Values for the Average Integrative Participant, With 3:1 Asymmetric Reward Proportions (3 Cents vs. 1 Cent)

Parameter	Value
Threshold (z)	0.0741
Starting point (x_0)	0.0063
Drift (A)	0.1861
Residual latency (T_0)	356.4 ms[a]
Drift SD (s_A)	0.1199
Starting point range (s_x)	0.0771
T_0 range (s_t)	0.2123
χ^2 fit error	244.83

[a] Compare to average signal detection response time of 301 ms in Experiments 1 and 2.

grative behavior during most of the experimental session whereas others integrated led to a bimodal shape for the empirical density of pooled RTs.[10] This density had an early mode matching the peak of the RT density for the signal detection task in Experiments 1 and 2. Figure 17 presents the empirical RT densities for the pooled data of all 15 participants, with separate densities for favored correct, favored error, unfavored correct, and unfavored error responses. These were superimposed on the signal detection RT density obtained in Experiments 1 and 2. The favored correct RT density shows two clearly discernible modes, with the earlier mode almost aligned to the peak of the detection task density. This indicates that, during nonintegrative performance, participants pressed the more rewarded key almost exclusively, consistent with Prediction 3d. While in integrative mode, however, they made favored and unfavored responses in similar proportions (both density curves are rescaled so that the area below them is proportional to the number of responses of each type). The plots resemble those in Figure 15, illustrating the predicted similarity between the unequal probability conditions of Experiment 2 and the unequal reward condition of Experiment 3.

Discussion

Experiment 3 demonstrates that, when reward inequality is introduced in two-alternative decision-making tasks, participants are able to adjust their decision behavior within a single session in a way that qualitatively matches the predictions of the theory of optimal decision-making. The magnitude of the observed adjustment, however, was smaller than predicted, perhaps because more practice was required before optimal control strategies could develop. As predicted by the theory, participants were more likely to make the more rewarded response than the alternative and were faster and made more mistakes when making favored rather than unfavored responses.

The theory also predicts that, for sufficiently large values of the reward ratio, participants will select the more rewarded response exclusively, regardless of the identity of the stimulus. In Experiment 3, we chose a reward ratio (3:1) that was likely to be large enough to trigger nonintegrative behavior for some of the participants, based on what we had observed with similar motion coherences and RSIs in Experiment 2 and based on the similarity of the effects predicted by Equation 6 for stimulus proportion manipulations and the effects of reward asymmetry predicted by Equations 9–10. The results of the experiment matched this qualitative prediction for most participants (excluding the 3 who displayed no response bias): Some participants exhibited nonintegrative behavior (they chose the favored response for almost the whole session), whereas the majority showed integrative behavior (they chose both responses with frequencies that were similar, but with a bias toward the favored response).

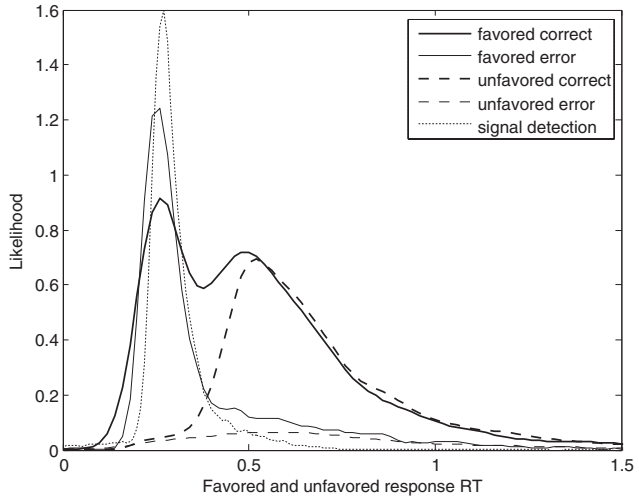

Figure 17. Distribution of response times (RTs) for the favored and unfavored responses in Experiment 3, plotted against the RT distribution for signal detection obtained in Experiment 2. The RT distribution for correct favored responses is bimodal, with the earlier mode almost aligned to the maximum in the signal detection curve. The distribution of incorrect favored responses is concentrated around that early mode. Conversely, the distributions of unfavored responses show almost no sign of an early mode, and their maximum is roughly aligned with the second mode of the distribution for correct favored responses, which indicates that early responses were almost exclusively favored ones.

General Discussion

We evaluated quantitative predictions of an optimal model of 2AFC decision making. These predictions (Bogacz et al., 2006) focused specifically on the behavioral effects produced by manipulations of mean RSI, stimulus probability, and relative reward magnitude—factors that enter into a wide range of decision-making tasks.

In a motion discrimination task with equally likely stimuli and equally rewarded responses, a reduction of the mean RSI was predicted to cause participants to place a greater emphasis on speed and less on accuracy. More specifically, this shift in SAT was predicted to occur as a result of specific threshold reductions, which could be identified by fits of the DDM to the observed RT distributions. Evidence from Experiment 1 supported these predictions, although the degree of threshold adaptation was less than predicted, and thresholds appeared suboptimally large in two of three conditions.

When one stimulus was more frequent than the other, a response bias was predicted to develop as a result of specific starting point shifts, producing more errors when the less likely stimulus appeared but faster RTs when the more likely stimulus appeared. When the stimulus probabilities were sufficiently asymmetric and RSIs were sufficiently short, an extreme response bias was predicted that would involve nonintegrative responding; that is, exclusive responding in favor of the more likely stimulus, with RT distributions comparable to those in a signal detection task involving the same, easily detectable stimuli. Evidence from Experiment 2 showed that such biases developed. Furthermore, RT, accuracy, and response proportions manifesting these optimal biases could be accurately predicted in Experiment 2, purely on the basis of fits to data collected in Experiment 1, as well as on simultaneous fits to data from Experiments 1 and 2. Finally, when correct responses to one stimulus were more rewarded than correct responses to the

[10] In Experiment 2, in contrast, bimodality appeared even within the RT distributions for individual participants.

other, similar biases were predicted to develop. Evidence from Experiment 3 suggested that these biases (including nonintegrative responding) developed as predicted, although model fits suggested suboptimally small starting-point shifts and threshold reductions for those participants who did not switch to nonintegrative responding.

These findings raise an important theoretical question involving adaptation to changing task conditions: Because task parameters such as mean RSI that determine optimal thresholds must be repeatedly sampled to optimize the DDM or any other model, how quickly can adaptation be accomplished? Also, how accurate (i.e., how close to optimal) can learned thresholds and starting points become? Human participants, of course, cannot plausibly adapt thresholds instantaneously: New RSI values must be experienced before adaptation can occur. There is empirical evidence that human participants, performing well-practiced tasks, are capable of adapting performance over relatively short intervals (e.g., in as few as 5–10 trials) after a change in task conditions (R. Bogacz, personal communication; R. Ratcliff, personal communication; Ratcliff et al., 1999; but see, e.g., Myung & Busemeyer, 1989, where evidence was found only for slow adaptation). The well-known phenomena of posterror slowing (Rabbitt, 1969) and of recovery of speed after multiple correct responses (Rabbitt & Vyas, 1970) are also consistent with rapid adjustments of decision thresholds. Simen, Cohen, and Holmes (2006) proposed a rapid threshold adaptation algorithm that can achieve nearly optimal thresholds within this 5–10-trial time frame. This algorithm adjusts thresholds $\pm z$ continuously, setting them equal at every moment to a decreasing, linear function of a running estimate of recent reward rate that is estimated by an exponentially weighted average of recent rewards: $z(t) = z_{max} - w \cdot$ reward rate. Future work will investigate whether this or some other process is at work in adapting thresholds and starting points (or perhaps other DDM parameters).

Recent empirical work (Bogacz et al., in press) may also help to determine whether parameter correlation is an appropriate explanation for the observed pattern of suboptimally large thresholds in many conditions or whether the use of objective functions other than reward rate (e.g., ones that include an emphasis on accuracy) can better explain such findings. This empirical work aims specifically to apply another prediction of the theory in Bogacz et al. (2006)—the prediction of optimal performance curves relating RT to accuracy—to data involving a wide variation in error percentages to distinguish between the possible objective functions governing behavior. Because error percentages never exceeded 20% in our experiments, we could not test this prediction.

An alternative explanation for suboptimally high thresholds is that estimates of reward rate may be subject to temporal uncertainty. The asymmetric functional relationship of reward rate to threshold noted earlier (Figure 2A), as well as recent theoretical work (Zacksenhouse, Holmes, & Bogacz, in review), suggests that efforts to maximize reward rate in the presence of timing uncertainty should lead to overestimation of the optimal threshold. This suggests that individuals with less accurate ability to estimate interval duration should overestimate optimal thresholds to a greater extent. We are currently investigating this prediction.

We now address the final theoretical issue raised by our findings. This involves the frequently observed phenomenon in which fitted values of residual latency appear to be unreasonably large under the simple additivity and pure insertion assumptions of Donders's subtraction method. Under these assumptions, the residual latency should be equal to the average signal detection RT in our task, as the signal detection task was identical to the decision-making tasks, except insofar as it required no discrimination between leftward and rightward motion.

In many different fits to our data with a variety of constraints among parameters and upper bounds on variability parameters, as well as with different subsets of the data itself, the fitted residual latencies for individual participants and for the group as a whole were usually 50 ms or more longer than the average signal detection RT observed in the last two blocks of all sessions of Experiments 1 and 2. Differences of more than 25 ms were produced even in fits of the pure DDM to the pooled data, so correlations of T_0 with overly inflated variability parameters (cf. Ratcliff & Tuerlinckx, 2002) cannot entirely explain the phenomenon (although such correlations seem to explain why the discrepancies between T_0 and the average signal detection RT were about 70 ms larger in completely unconstrained fits of the extended DDM to data from Experiments 1 and 2). There have been many criticisms of Donders's subtraction method and the related assumption of pure insertion in models involving stages of processing. Nevertheless, these stages-of-processing approaches continue to exert a strong conceptual influence on RT research (Sternberg, 2001); in particular, such an approach is embodied in the typical interpretation of the DDM's residual latency parameter.

Keeping these caveats regarding additivity in mind, what our data seem to suggest is that there may be an irreducible increment of roughly 50 ms that is incurred when participants integrate, relative to nonintegrative responding. This may reflect the overhead of an additional stage of processing that can be eliminated when integration is not needed. In this conception, integrating information automatically causes the residual latency component of RT to increase by a positive quantity, Δ, in addition to requiring additional time (DT) for the DD process to cross threshold ($RT = DT + T_0 + \Delta$). Evidence for this comes from the transitional shapes of RT densities in Experiments 2 and 3: Rather than simply shifting leftward and diminishing in width, the RT densities from integrative response conditions go through a bimodal stage before converging to the shape of the signal detection RT density (Figures 15 and 17).

Bimodality of this type is consistent with a mixture model of integrative and nonintegrative responses, and the gap between modes in almost all of the correct/favored distributions in Experiment 2 is furthermore consistent with an irreducible T_0 increment: After all, a mixture model need not produce bimodality, but a large enough T_0 increment would tend to keep the two mixture components sufficiently separated so that bimodality would result. Furthermore, in the 90:10 stimulus-ratio conditions in which nonintegrative responding is evident, average RT is not equal to the fitted T_0 value: Instead, average nonintegrative 2AFC RT is statistically indistinguishable from the average signal detection RT. The data suggest that participants make an all-or-none decision either to integrate or not to integrate and thereby to reduce RT substantially (by $\Delta + DT$; Figure E2 in Appendix E shows an individual participant's performance from trial to trial that illustrates what appears to be precisely a switching from integrative to primarily nonintegrative behavior in fast, asymmetric blocks of Experiment 2).

In terms of neural processing, a time overhead Δ might be incurred by requiring processing by an additional, intermediate signal discrimination layer between a sensory input layer and a motor output layer (see the network models in Bogacz et al., 2006; Shadlen, Britten, Newsome, & Movshon, 1996; Simen et al., 2006; and Usher & McClelland, 2001, which all model the decision process as occurring in a specific network layer). This layer might require some nonzero "startup time" before carrying out its signal discrimination function, perhaps as a result of conduction times between distant brain areas or as a result of the smearing of sudden stimulus onsets into relatively gradual rises in the decision layer's inputs due to the effects of sluggish processing in earlier network layers. Time might, therefore, be saved if stimulus information could skip over this intermediate layer, the function of which would be unnecessary if the participant was precommitted to a particular response before the stimulus appeared. More behavioral and physiological work would be needed to evaluate this hypothesis. Fortunately, however, very little machinery would be required to achieve such a precommitment in the previously mentioned network models: If their discrimination layers incurred a one-time startup cost but then remained committed to a single response (equivalently, if the DDM starting point remained beyond one of the thresholds), then the models would perform signal detection as quickly as models lacking a discrimination layer. These linear systems models would need to be augmented, however, to include propagation delays or nonlinear activation dynamics to account for the hypothetical startup delay during integrative responding.

Conclusion

The theory of optimal decision making makes quantitative predictions that can be tested by model fitting and qualitative predictions that can be directly observed. Both types of prediction were supported by the data in our experiments. This theory also appears to provide leverage even with models that only approximate the optimal decision process. The extended DDM, for example, involves variability in starting point, drift, and residual latency that deviate from the optimal SPRT. Nevertheless, we were able to predict RTs, accuracy, and response proportions based on extended DDM fits of A and T_0 by computing thresholds and starting points from expressions developed for the pure DDM. In principle, the same approach is applicable to models such as the Ornstein–Uhlenbeck process incorporated into decision field theory (Busemeyer & Townsend, 1993) and the leaky-competing accumulator model of Usher and McClelland (2001); these models, like the extended DDM, contain the pure DDM as a special case. At the very least, when the parameters of these models are not too far from those that implement the pure DDM, the same phenomena (threshold and starting-point shifting, and transitions to nonintegrative responding) should occur in similar task conditions. Thus, the normative theory of 2AFC decision making may be applicable even to models for which it was not expressly developed.

For this reason, future experiments designed to discriminate between competing models of decision making may benefit from the type of manipulations involved in our experiments; that is, the type of manipulations affecting reward rate that are typically undertaken in studies of instrumental conditioning (e.g., Herrnstein, 1997). Different models may make dramatically different predictions when coupled with the assumption that reward rate (or some other objective function, e.g., a linear combination of RT and accuracy or of reward rate and accuracy) is being maximized: For example, they may make different predictions about when a transition to nonintegrative responding should occur, and this transition should be clearly identifiable in the data, as shown in Experiments 2 and 3. Therefore, a detailed, quantitative analysis of these predictions may help to tease apart what often appear to be subtle differences between alternative models of decision making.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

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, J., 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.

Bogacz, R., Hu, P., Cohen, J., & Holmes, P. (in press). Do humans produce the speed-accuracy tradeoff that maximizes reward rate? *Quarterly Journal of Experimental Psychology.*

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12,* 4745–4765.

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., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature, 377,* 59–62.

Edwards, W. (1961). Costs and payoffs are instructions. *Psychological Review, 68,* 275–284.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 2,* 312–329.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap.* New York: Chapman and Hall.

Estes, W. K., & Maddox, W. T. (2005). Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. *Psychonomic Bulletin and Review, 12,* 403–408.

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

Fitts, P. (1966). Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology, 71,* 849–857.

Freedman, D., & Diaconis, P. (1981). On the maximum deviation between the histogram and the underlying density. *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete, 58,* 139–167.

Gardiner, C. W. (2004). *Handbook of stochastic methods* (3rd ed.). New York: Springer-Verlag.

Garrett, H. (1922). A study of the relation of accuracy to speed. *Archives of Psychology, 56,* 1–105.

Gill, P. E., Murray, W., Saunders, M. A., & Wright, M. H. (1998). *User's guide for NPSOL 5.0: A FORTRAN package for nonlinear programming* (Tech. Rep. SOL 86-1.). Stanford, CA: Stanford University, Systems Optimization Laboratory.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Science, 5,* 10–16.

Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron, 36,* 299–308.

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York: Wiley.

Grice, G. R. (1968). Stimulus intensity and response evocation. *Psychological Review, 75,* 359–373.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science, 274,* 427–430.

Herrnstein, R. J. (1997). *The matching law: Papers in psychology and economics.* In H. Rachlin & D. I. Laibson (Eds.). Cambridge, MA: Harvard University Press.

Jentzsch, I., & Dudschig, C. (2009). Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *The Quarterly Journal of Experimental Psychology, 62,* 209–218.

LaBerge, D. (1962). A recruitment theory of simple behavior. *Psychometrika, 27,* 375–396.

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder-Mead Simplex method in low dimensions. *SIAM Journal on Optimization, 9,* 112–147.

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

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

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

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

Maddox, W. T., & Bohil, C. J. (1998). Base rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24,* 1459–1482.

McKinnon, K. (1998). Convergence of the Nelder-Mead Simplex method to a nonstationary point. *SIAM Journal on Optimization, 9,* 148–158.

Myung, I. J., & Busemeyer, J. R. (1989). Criterion learning in a deferred decision making task. *American Journal of Psychology, 102,* 1–16.

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

Ollman, R. T. (1966). Fast guesses in choice-reaction time. *Psychonomic Science, 6,* 155–156.

Pachella, R, & Pew, R. (1968). Speed-accuracy tradeoff in reaction time: Effect of discrete criterion times. *Journal of Experimental Psychology, 76,* 19–24.

Pacut, A. (1977). Some properties of threshold models of reaction latency. *Biological Cybernetics, 28,* 63–72.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Vision Research, 5,* 376–404.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

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

Rabbitt, P., & Vyas, S. (1970). An elementary preliminary taxonomy of errors in choice reaction time tasks. *Acta Psychologica, 33,* 56–76.

Rapoport, A., & Burkheimer, G. J. (1971). Models for deferred decision making. *Journal of Mathematical Psychology, 8,* 508–538.

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

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

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review, 92,* 212–225.

Ratcliff, R. (2001). Diffusion and random walk processes. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (Vol. 6, pp. 3668–3673). Oxford: Elsevier.

Ratcliff, R., Cherian, A., & Segraves, M. A. (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., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922.

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,* 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., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19,* 278–289.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin and Review, 9,* 438–481.

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

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience, 22,* 9475–9489.

Rouder, J. N., & Speckman, P. L. (2004). An evaluation of the vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin and Review, 11,* 419–427.

Schall, J. D. (2001). Neural basis of deciding, choosing and acting. *Nature Reviews Neuroscience, 2,* 33–42.

Schouten, J., & Bekker, A. (1967). Reaction time and accuracy. *Acta Psychologica, 27,* 143–153.

Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience, 16,* 1486–1510.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology, 86,* 1916–1936.

Simen, P., Cohen, J. D., & Holmes, P. (2006). Rapid decision threshold modulation by reward rate in a neural network. *Neural Networks, 19,* 1013–1026.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience, 27,* 161–168.

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.

Sommer, W., Leuthold, H., & Soetens, E. (1999). Covert signs of expectancy in serial reaction time tasks revealed by event-related potentials. *Perception and Psychophysics, 61,* 342–352.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 461–479.

Sternberg, S. (2001). Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica, 106,* 147–246.

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

Tanner, W., & Swets, J. (1954). A decision-making theory of visual detection. *Psychological Review, 61,* 401–409.

Tuerlinckx, F. (2004). The efficient computation of the cumulative distribution and probability density functions in the diffusion model. *Behavioral Research Methods, Instruments and Computers, 36,* 702–716.

Tuerlinckx, F., Maris, E., Ratcliff, R., & De Boeck, P. (2001). A comparison of four methods for simulating the diffusion process. *Behavioral Research Methods, Instruments and Computers, 33,* 443–456.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592.

Vandekerckhove, J., & Tuerlinckx, F. (2007a). The diffusion model analysis toolbox [computer software and manual]. Retrieved July 21, 2009, from http://ppw.kuleuven.be/okp/dmatoolbox/

Vandekerckhove, J., & Tuerlinckx, F. (2007b). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin and Review, 14,* 1011–1026.

Van Zandt, T. (2000) How to fit a response time distribution *Psychonomic Bulletin and Review, 7,* 424–465.

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

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory and Cognition, 32,* 1206–1220.

Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics, 16,* 117–186.

Wald, A, & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Annals of Mathematical Statistics, 19,* 326–339.

Wickelgren, W. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41,* 67–85.

Yang, T., Hanks, T., Mazurek, M. E., McKinley, M., Palmer, J., & Shadlen, M. N. (2005). *Incorporating prior probability into decision-making in the face of uncertain reliability of evidence.* Presented at the annual meeting of the Society for Neuroscience, Washington, DC.

Yellott, J. I. (1971). Correction for guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology, 8,* 159–199.

Zacksenhouse, M., Holmes, P., & Bogacz, R. (2007). Robust versus optimal strategies for two-alternative forced choice tasks. Manuscript submitted for publication.

Appendix A

Conversion Between Terminology of Bogacz et al. (2006) and Ratcliff and Tuerlinckx (2002)

Table A1 provides a parameter conversion table to assist readers more familiar with the parameter symbols used by Ratcliff and Tuerlinckx (2002).

Table A1
Parameters of the Pure and Extended Drift-Diffusion Models

Parameter	Pure	Extended	Ratcliff et al. (2006) terminology
Drift	A	A	$\zeta\ (= A)$
Threshold	z	z	$A\ (= 2\ z_{Bogacz})$
Starting point	x_0	x_0	$z\ (= z_{Bogacz} + x_0)$
Residual latency	T_0	T_0	$T_{er}\ (= T_0)$
Noise	c	c	$s\ (= c)$
Starting point variability	—	s_x	$s_z\ (= s_x)$
Drift variability	—	s_A	$\eta\ (= s_A)$
T_0 variability	—	s_t	$s_t\ (= s_t)$

Note. In the left and middle columns are the parameter symbols used by Bogacz, Brown, Moehlis, Holmes, and Cohen (2006); in the right column is the terminology used by Ratcliff and Tuerlinckx (2002).

(*Appendixes continue*)

Appendix B

Accuracy and DT for Nonzero Starting Points

When starting point x_0 is not 0 (i.e., not equidistant from both thresholds), the expressions for $\overline{\mathrm{ER}}$ and $\overline{\mathrm{DT}}$ are as follows (Equations A43 and A44 in Bogacz et al., 2006), where Π denotes the probability of the more likely stimulus:

$$\overline{\mathrm{ER}} = \frac{1}{1 + e^{(2Az/c^2)}} - \frac{1 - \Pi e^{-2x_0(A/c)^2} - (1 - \Pi)e^{2x_0(A/c)^2}}{e^{2zA/c^2} - e^{-2zA/c^2}};$$

$$\overline{\mathrm{DT}} = \frac{z}{A}\tanh\frac{Az}{c^2} + (1 - 2\Pi)x0 + \frac{2z}{A}\left\{\frac{1 - \Pi e^{-2x_0(A/c)^2} - (1 - \Pi)e^{2x_0(A/c)^2}}{e^{2zA/c^2} - e^{-2zA/c^2}}\right\}.$$

Appendix C

Reward-Maximizing SAT for Models With Concave SATF

SATs are often conceptualized as points along a SATF. An SATF defines the proportion of correct responses as a function of mean RT (Luce, 1986). According to the theory of optimal DDM parameterization, points along an SATF are selected in response to changing stimulus probabilities, rewards, and RSI. Changes in the SNR, in contrast, are predicted to change the SATF itself; we did not investigate changes in SNR in this article.

For any model producing a smooth, concave SATF (i.e., in which accuracy increases smoothly and monotonically with mean RT but has a strictly negative second derivative with respect to RT), the definition of reward rate in Equation 4 implies the existence of a unique SAT that maximizes reward rate. This is the case for the DDM but also for other models of decision making.

Equation 4 defines reward rate as follows, where $\overline{\mathrm{ER}}$ is error proportion, $\overline{\mathrm{DT}}$ is average DT, T_0 is residual latency, and $\overline{\mathrm{RSI}}$ is the average RSI:

$$\overline{\mathrm{RR}} = \frac{1 - \overline{\mathrm{ER}}}{\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}}}. \quad (11)$$

Let Acc represent accuracy: $\mathrm{Acc} = 1 - \overline{\mathrm{ER}}$. Then an SATF is given by $\mathrm{Acc}(\overline{\mathrm{DT}})$, which we assume is strictly increasing and concave as $\overline{\mathrm{DT}}$ (and therefore RT) increases. From Equation 11, we therefore have:

$$\overline{\mathrm{RR}}(\overline{\mathrm{DT}}) = \frac{\mathrm{Acc}(\overline{\mathrm{DT}})}{\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}}}. \quad (12)$$

Because $\mathrm{Acc}(\overline{\mathrm{DT}})$ is clearly bounded above by 1 (representing perfect accuracy), whereas $\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}}$ grows without bound as $\overline{\mathrm{DT}}$ increases, $\overline{\mathrm{RR}}(\overline{\mathrm{DT}})$ approaches 0 as $\overline{\mathrm{DT}}$ approaches infinity.

We can also assume that $\mathrm{Acc}(0)$ is near 0.5 (representing chance performance). Therefore, $\overline{\mathrm{RR}}(\overline{\mathrm{DT}})$ either decreases monotonically toward 0 as $\overline{\mathrm{DT}}$ increases (meaning that it has a maximum at $\overline{\mathrm{DT}} = 0$), or $\overline{\mathrm{RR}}(\overline{\mathrm{DT}})$ has one or more local maxima for $\overline{\mathrm{DT}}$ (0,).

To analyze how many possible local maxima exist, we take the derivative of $\overline{\mathrm{RR}}$ with respect to $\overline{\mathrm{DT}}$:

$$\overline{\mathrm{RR}}'(\overline{\mathrm{DT}}) = \frac{\mathrm{Acc}'(\overline{\mathrm{DT}})}{\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}}} - \frac{\mathrm{Acc}(\overline{\mathrm{DT}})}{(\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}})^2}. \quad (13)$$

Setting $\overline{\mathrm{RR}}'(\overline{\mathrm{DT}}) = 0$, we get:

$$\frac{\mathrm{Acc}'(\overline{\mathrm{DT}})}{\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}}} = \frac{\mathrm{Acc}(\overline{\mathrm{DT}})}{(\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}})^2}$$

$$\mathrm{Acc}'(\overline{\mathrm{DT}}) = \frac{\mathrm{Acc}(\overline{\mathrm{DT}})}{(\overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}})}. \quad (14)$$

Equation 14 states that the local maxima or minima of $\overline{\mathrm{RR}}$ must occur at values of DT where the derivative of the SATF equals the reward rate (up to a constant scaling factor involving the size of rewards—for the present discussion, we assume that rewards have a unit magnitude).

The second derivative of $\overline{\mathrm{RR}}$ determines whether the zeros of Equation 13 are local minima or maxima. The second derivative is given by the following, where we set $\overline{\mathrm{RT}} = \overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}}$:

$$\overline{\mathrm{RR}}''(\overline{\mathrm{DT}}) = \frac{\mathrm{Acc}''}{\overline{\mathrm{RT}}} - \frac{\mathrm{Acc}'}{(\overline{\mathrm{RT}})^2} - \frac{\mathrm{Acc}'}{(\overline{\mathrm{RT}})^2} + \frac{2\,\mathrm{Acc}}{(\overline{\mathrm{RT}})^3}$$

$$= \frac{\mathrm{Acc}''}{\overline{\mathrm{RT}}} - \frac{2\,\mathrm{Acc}'}{(\overline{\mathrm{RT}})^2} + \frac{2\,\mathrm{Acc}}{(\overline{\mathrm{RT}})^3}$$

$$= \frac{\mathrm{Acc}''}{\overline{\mathrm{RT}}} \text{ (by substitution of Eq.14).} \quad (15)$$

Because $\overline{\mathrm{RT}} = \overline{\mathrm{DT}} + T_0 + \overline{\mathrm{RSI}} > 0$ and $\mathrm{Acc}''(\overline{\mathrm{DT}}) < 0$ by concavity of the SATF, $\overline{\mathrm{RR}}''(\overline{\mathrm{DT}})$ must be strictly negative, and therefore any value of $\overline{\mathrm{DT}}$ for which $\overline{\mathrm{RR}}' = 0$ is a local maximum.

By the assumed continuity of the SATF, any two neighboring local maxima must be separated by a local minimum, or else a piecewise constant segment of $\overline{\mathrm{RR}}(\overline{\mathrm{DT}})$ is maximal. However, the existence of such a segment would imply that $\overline{\mathrm{RR}}'' \equiv 0$ over that segment, contradicting the assumption that $\overline{\mathrm{RR}}'' < 0$.

For models that produce monotonically increasing SATFs that are not concave, multiple local maxima are possible.

Appendix D

Data Fitting

Data in Experiments 1 and 2 were fit using the chi-square fitting method of Ratcliff and Tuerlinckx (2002), implemented in MATLAB software custom-written by the authors. We extended this method to incorporate upper bounds on certain parameters during fitting, as well as to allow fitting of a mixture model consisting of an RT distribution generated by the DDM and a normal RT distribution with smaller mean and variance. Data in Experiment 3 were fit in MATLAB with the Diffusion Model Analysis Toolbox (DMAT) (Vandekerckhove & Tuerlinckx, 2007a, 2007b). Here we focus on the details of the fitting methods used in Experiments 1 and 2.

Fit Error Function

In the chi-square fitting method of Ratcliff and Tuerlinckx (2002), the 0.1, 0.3, 0.5, 0.7, and 0.9 quantiles are used to define six bins of RTs, with the fastest RT bin and the slowest RT bin each containing 10% of the total number of trials and the other bins containing 20%. A given set of DDM parameters was then used at each iteration of the fitting process (discussed later) to generate a cumulative distribution function (CDF) for each of the two types of responses.[11] The CDF (computed for the extended DDM with the freely available MATLAB function *CDFDif.m* described in Tuerlinckx [2004]) is used to generate a prediction of the number of trials expected within each bin. The chi-square error function is given by the following equation:

$$\chi^2 = \sum_{conditions} \sum_{i=1}^{6} \frac{(\text{trials observed}_{\text{bin}_i} - \text{trials expected}_{\text{bin}_i})^2}{\text{trials expected}_{\text{bin}_i}}. \quad (16)$$

This fit error function was evaluated by Ratcliff and Tuerlinckx (2002) in a study comparing the maximum likelihood method, the chi-square method we use here, and a weighted least-squares method applied to quantiles. Simulated data was constructed for a set of parameter values, and the methods were evaluated for computational speed, bias, and robustness to contaminants. Contaminants are responses not generated by the diffusion process (perhaps because of failures to attend to the task). Lacking any more informed model of what the RT distribution should be for real contaminants, Ratcliff and Tuerlinckx (2002) simulated them as RTs generated by the extended DDM, with an additional increment drawn from a uniform distribution; in fitting this simulated data, they made the simplifying assumption that the contaminants were drawn from a uniform distribution spanning most of the observed RT range. The chi-square method applied to the extended DDM—with an additional parameter intended to capture contaminant RT proportions—was the method they recommended: It was faster and more robust than the maximum likelihood method and less biased than the weighted least-squares method that they investigated.

We added four additional parameters to this extended DDM and modified its first-passage time CDF before computing the fit error: Three parameters were needed to make data from Experiment 2 fittable, and one was used to test the hypothesis that drift is strategically adapted from condition to condition (thereby contradicting the prediction of constant drift for an optimized DDM). The first three parameters were the mean, variance, and mixture weight of a normal distribution intended to model a nonintegrative RT distribution; the fourth parameter defined an increment that could be added to the single drift term that was fit across all conditions.

Optimization Algorithm

To minimize the fit error over the space of parameter values, Ratcliff and colleagues (e.g., Ratcliff & McKoon, 2008) typically use the Simplex algorithm (Nelder & Mead, 1965). Tuerlinckx has used a constrained optimization algorithm, NPSOL (Gill, Murray, Saunders, & Wright, 1998) instead of Simplex (Ratcliff & Tuerlinckx, 2002), which allows the user to constrain search over the parameter space to a particular region and to supply information about the function being minimized that speeds the search process. However, Ratcliff and Tuerlinckx (2002) reported that this method suffered from numerical instability problems and failures to converge to minima.

Despite its demonstrated practical utility in a wide range of problems, theoretical understanding of the convergence properties of Simplex is limited (Lagarias, Reeds, Wright, & Wright, 1998). It can fail to converge to a minimum even for convex functions in two dimensions (McKinnon, 1998). In purely practical terms, however, it was our experience that fitting with Simplex took much longer than with the constrained optimization method implemented in MATLAB's fmincon.m function. The latter is the method we used for the analysis in this article. Furthermore, as we noted in the article's introduction, it turned out to be quite useful to constrain the variability parameters of the extended DDM to provide some amount of control over parameter inflation. Constrained optimization approaches are designed for this type of restriction, whereas constraining an unconstrained algorithm by assigning high fit errors to undesirable parameter regions is an art.

[11] Ratcliff and Tuerlinckx (2002) referred to this function and its corresponding density as defective, indicating that the CDF (the integral of the density) does not approach 1 as RT approaches infinity. The sum of CDFs for the two responses, evaluated at infinity, does equal 1, however. The two defective distributions can then be fit to correct and error distributions separately (and these two distributions can be further subdivided into favored and unfavored stimulus RTs in Experiments 2 and 3).

(*Appendixes continue*)

For our problem, MATLAB's fmincon.m automatically selected its medium-scale settings. Under these settings, the fitting process alternates between two phases. In the first phase, it estimates the curvature of the error surface around the current search point in parameter space using sequential quadratic programming. In the second phase, it uses line search (related to Newton's method) to minimize the function along a line in parameter space selected on the basis of the curvature estimate. Then the process repeats.

This algorithm has proven convergence properties for smooth error functions. As a sum of (normalized) squares, the chi-square error function is smooth as long as the expected number of trials in the denominator of each term in Equation 16 does not approach 0 for any term in the sum.[12] For fits of the extended DDM without contaminants to RT distributions in individual conditions, fitting a condition typically took less than 30 s on a 2.53-GHz Intel Pentium IV with 512-KB cache, 533-MHz bus, and 512 MB of RAM, as opposed to a typical fit time of several minutes for a Simplex approach.

As noted by Ratcliff and Tuerlinckx (2002), convergence problems can indeed pose a difficulty for this approach; however, judicious use of initial conditions and parameter bounds made these problems manageable. For example, because it was clear that the first mode in the bimodal, empirical RT densities of Experiments 2 and 3 occurred well before 300 ms, an upper bound of 290 ms could be applied to the nonintegrative mean parameter during fitting. The result was fast convergence to a value of 267 ms. In contrast, when this bound constraint was not imposed, the fitting algorithm was prone to wandering into a region of parameter space that assigned high mean and variance to the normal component of our mixture model; once this happened, it was extremely difficult for the algorithm to recover, and searches usually terminated with extremely high fit errors and nonsensical parameter estimates.

However, numerical instability may still affect the algorithm as it is implemented in fmincon.m, because the error surface defined by the χ^2 function, although smooth in theory, appears to be quite jagged in practice: Tiny changes in parameter values can create extremely large jumps in the error, especially when data from multiple participants is pooled together (presumably because the larger number of total trials resulting from pooling produces larger expected bin counts for the fastest and longest RT bins, and deviations from this expectation drive up the error dramatically). In fitting the extended DDM, the default parameter settings for fmincon.m were extremely effective; as noted, good fits were achieved remarkably quickly.[13]

Application of the Freedman–Diaconis Histogram Bin Size Rule

For the histograms in Figures 7 and 12, we used a fixed-width bin size governed by the Freedman–Diaconis rule (Freedman & Diaconis, 1981) for minimizing the error between the histogram and the actual density. This bin size rule adapts the bin width to data from a given experimental condition according to the following equation:

$$\text{bin size} = \frac{2 \times \text{Interquartile range}}{\text{Number of observations}^{1/3}}. \quad (17)$$

Interquartile range is the difference between the first and third quartiles. Because correct and error RTs were fit separately, making the histograms for errors and correct responses comparable required choosing which distribution to plot with this bin size rule and then applying the same bin size to the other distribution as well. Because there were far fewer error RTs in general, applying the bin size derived from the correct RTs to the error RTs tended to oversmooth the error RT data. We therefore applied the Freedman–Diaconis rule to the error RTs and used the derived bin size for both correct and error RTs.

[12] We followed the practice of Tuerlinckx (2004) and made sure that the denominator was never below 0.00001, although to do so we took the maximum of this small number and the expected bin count rather than adding 0.00001 to all expected bin counts as done by Tuerlinckx.

[13] The DMAT MATLAB Toolbox (Vandekerckhove & Tuerlinckx, 2007a, 2007b) for fitting the DDM appears to operate faster still for typical data sets, even relying on Simplex as its optimization algorithm, because of efficient, low-level code optimization.

Appendix E

Participant 305

Here we examined performance by an individual participant to demonstrate that the behavioral phenomena observed in pooled performance data from all participants (e.g., bimodal response time densities in Experiment 2) were not simply artifacts of pooling.

In Experiment 1, Participant 305 shifted from a relative emphasis on speed to a relative emphasis on accuracy in the 2-s RSI condition compared with the 500-ms and 1-s RSI conditions. Figure E1 plots RT densities for this participant from Experiments 1 and 2 that are consistent with nonintegrative responding (i.e., fast responding with one response exclusively) for large stimulus-ratio asymmetries Π. In this figure, only data from the 500-ms RSI conditions are plotted. As Π grows, the RT densities clearly undergo a transition toward the signal detection density. Ultimately, in the 90:10-ratio/500-ms RSI condition, the RT density is essentially identical to the signal detection RT density.

Further evidence for nonintegrative responding comes from examining response totals and RTs on a trial-by-trial basis. Figure E2 plots data from the first session of Experiment 2. Panel A plots cumulative favored responses as a function of trial number. Dashed lines plot the maximum possible cumulative total of favored responses within each block of trials. Purely nonintegrative responding causes the observed cumulative response plot to lie on top of the dashed line (most clearly observed in Block 9, with RSI = 1 s and 90% rightward stimuli); deviations from nonintegration cause the cumulative plot to err toward the horizontal. Panel B plots RT as a function of trial number. Dashed lines indicate the observed,

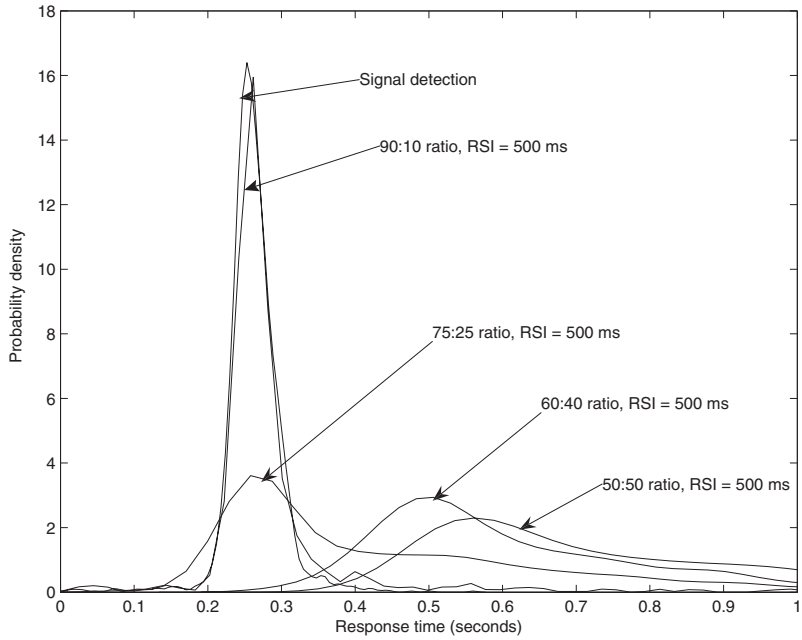

Figure E1. Comparison of signal detection response time (RT) density to RT densities in each two-alternative forced-choice condition of Experiment 2 for an individual participant, with an average response–stimulus interval (RSI) of 500 ms. As the stimulus proportion asymmetry increases, the RT density for two-alternative decisions approaches that for the signal detection condition.

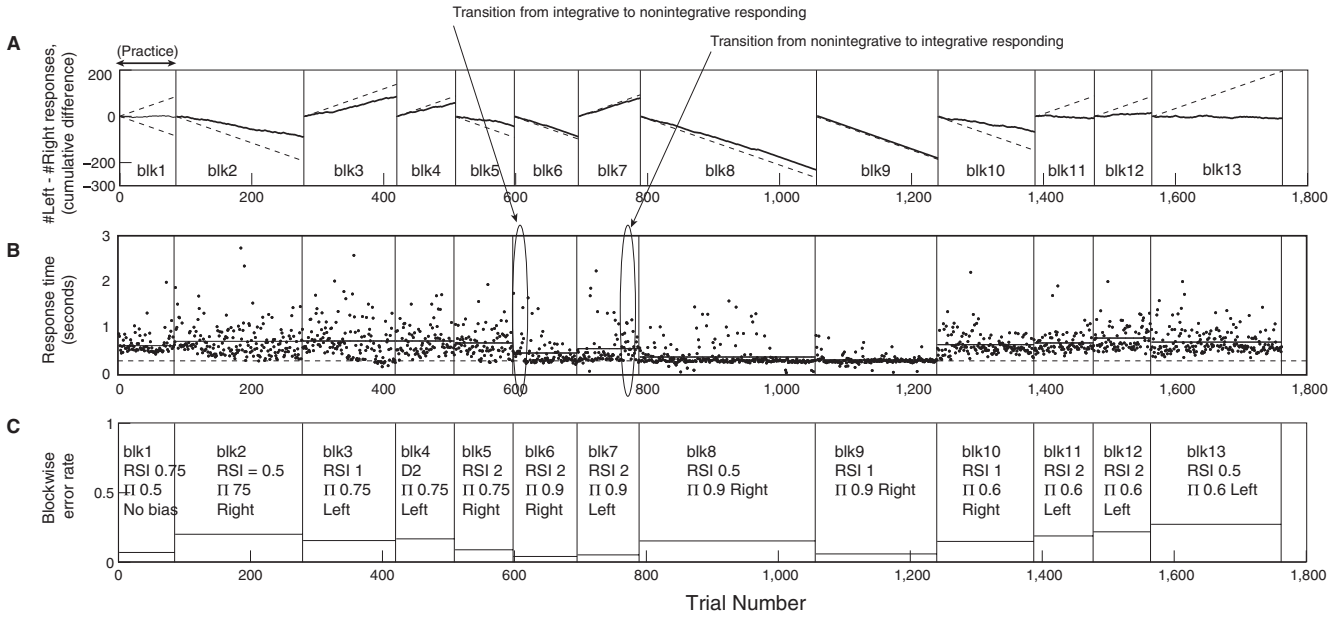

Figure E2. Trial-by-trial performance data from Participant 305 in the first session of Experiment 2 (after participation in the five sessions of Experiment 1). (A) Cumulative favored responses as a function of trial number. Dashed lines plot the maximum possible cumulative total of favored responses. (B) Response time (RT) as a function of trial number. Dashed lines indicate observed signal detection RT; superimposed solid, horizontal lines plot mean RT for the block. (C) The proportion of errors within each block as a function of trial number; text indicates RSI and Π conditions in each block.

(*Appendixes continue*)

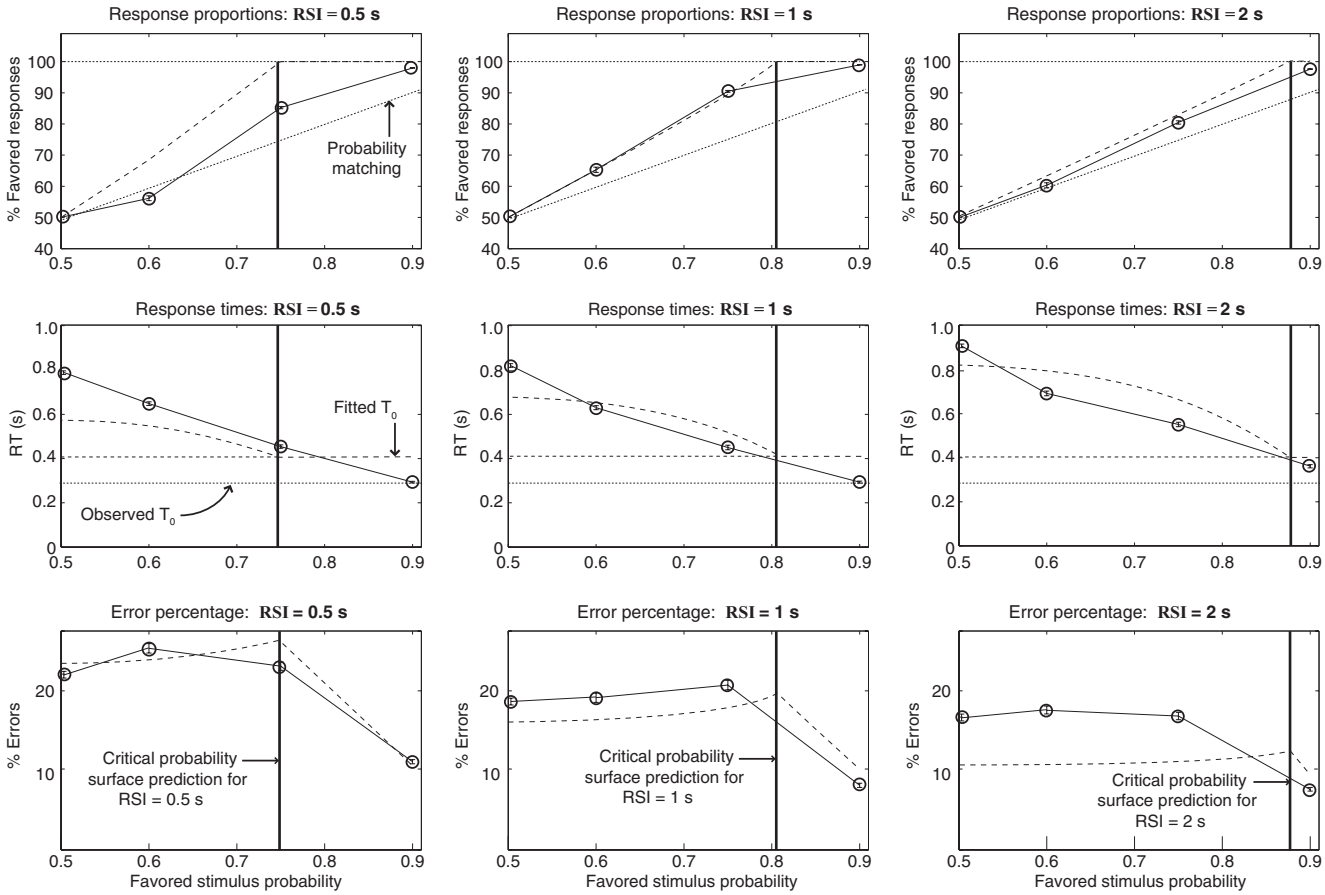

Figure E3. Comparison of predicted and observed response proportions, response times (RTs), and error percentages in all conditions of Experiment 2 based on unconstrained, extended drift-diffusion model fits of drift and residual latency to performance of Participant 305 in Experiment 1. The horizontal axis in each plot denotes the stimulus proportions (0.6 indicates a 60:40 ratio, 0.75 indicates a 75:25 ratio, and 0.9 indicates a 90:10 ratio). The left column of plots corresponds to a mean response–stimulus interval (RSI) of 500 ms; the middle column corresponds to a mean RSI of 1 s; the right column corresponds to a mean RSI of 2 s.

average signal detection RT; superimposed solid, horizontal lines indicate the average RT for the block. Panel C plots the proportion of errors within each block as a function of trial number; text indicates RSI and Π conditions in each block.

Block 7 (90:10 stimulus-ratio/500-ms RSI/rightward motion favored) and Block 8 (90:10 stimulus-ratio/1-s RSI/rightward motion favored) both show almost exclusive rightward responding, and responses are almost all near or below the average signal detection RT.

In later sessions, this participant showed evidence of nonintegrative responding in other conditions as well. This can be seen in Figure E3. In this comparison of response proportions, RTs and accuracy to the predictions of an optimally tuned DDM, Participant 305 also appears to have achieved nonintegrative responding in the 2-s RSI condition with a 90:10 stimulus ratio. However, given the fitted drift and residual latency terms, nonintegrative responding was also predicted (but not exclusively produced) in the 75:25-ratio/500-ms RSI condition.

Appendix F

Comparison of Unconstrained and Constrained Fits

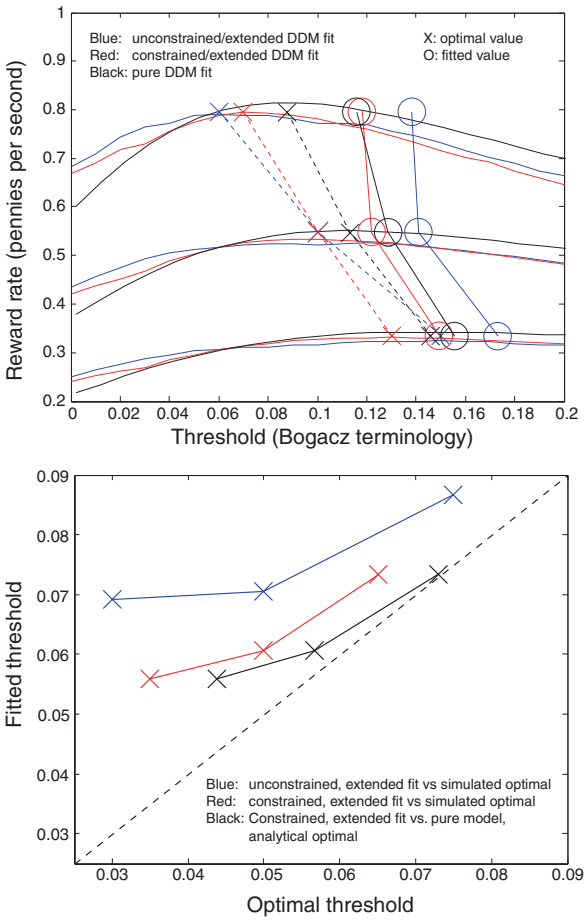

Figure F1. Comparison of pure drift-diffusion model (DDM) fits, constrained/extended DDM fits, and unconstrained/extended DDM fits in terms of harvesting efficiency. One set of reward rate curves corresponds to each of the response–stimulus interval values in Experiment 1. Unconstrained/extended fits are shown in blue; constrained/extended fits are in red; pure DDM are fits in black.

Constraining the additional parameters of the extended DDM during fitting as we have done forces the model to approximate the pure DDM. This approach to fitting our data resulted in starting-point and threshold values that were close to the optimal values, as defined by analytical functions of the fitted drift and residual latency values (Equations 6 and 7, respectively).

However, if unconstrained fitting of the extended DDM in fact achieves unbiased estimates of parameter values along with a reduced fit error relative to constrained fits—notwithstanding the parameter correlation problem that we have noted—then it is imperative to look for evidence of optimal SAT and response bias in these unconstrained fit values as well. Figure F1 shows the results of an unconstrained model fit in blue, a constrained model fit in red, and a pure DDM fit (with variability parameters set to 0) in black. In the top panel, the horizontal coordinate represents the reward maximizing threshold values (Xs) and fitted threshold values (Os) in each RSI condition; the vertical coordinate represents the average reward rate actually earned by participants in each condition. In the bottom panel, the fitted thresholds are plotted versus their optimal values.

The figure shows that, in all RSI conditions, unconstrained/extended fitting leads to parameter sets whose threshold values (blue Os) are much larger than the reward-maximizing values for the extended DDM (blue Xs), as indicated by simulations (noise and the flat maximum in the 2-s RSI condition make the position of the optimal value determined by simulations somewhat imprecise). These simulations show that more reward can be earned with the pure DDM than with the extended DDM, as the analytically derived black curves are greater than the numerically computed blue and red curves for the extended DDM, except at threshold values smaller than optimal.

The mismatch between fitted and optimal values for the constrained/extended DDM (red Os and Xs, respectively) is less pronounced. The mismatch between fitted and optimal values for the pure DDM (black Os and Xs) is the smallest of all. Thus, the choice of fitting procedure appears to determine whether the data favor the hypothesis of nearly optimal strategic control of decision making (at least when RSIs are greater than 500 ms), or, in contrast, an hypothesis of suboptimal emphasis on accuracy over reward rate in all conditions.

Received October 4, 2007
Revision received March 26, 2009
Accepted April 13, 2009 ■