frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 27 May 2011
doi: 10.3389/fnhum.2011.00048

Rational decision-making in inhibitory control

Pradeep Shenoy and Angela J. Yu*

Department of Cognitive Science, University of California, San Diego, La Jolla, CA, USA

Edited by:
Francisco Barcelo, University of Illes Balears, Spain

Reviewed by:
Christopher Summerfield, Oxford University, UK
Harriet Feldman, University College London, UK

****Correspondence:***
Angela J. Yu, Department of Cognitive Science, University of California, MC 0515, 9500 Gilman Drive, San Diego, La Jolla, 92093 CA, USA.
e-mail: ajyu@ucsd.edu

An important aspect of cognitive flexibility is inhibitory control, the ability to dynamically modify or cancel planned actions in response to changes in the sensory environment or task demands. We formulate a probabilistic, rational decision-making framework for inhibitory control in the stop signal paradigm. Our model posits that subjects maintain a Bayes-optimal, continually updated representation of sensory inputs, and repeatedly assess the relative value of stopping and going on a fine temporal scale, in order to make an optimal decision on when and whether to go on each trial. We further posit that they implement this continual evaluation with respect to a global objective function capturing the various reward and penalties associated with different behavioral outcomes, such as speed and accuracy, or the relative costs of stop errors and go errors. We demonstrate that our rational decision-making model naturally gives rise to basic behavioral characteristics consistently observed for this paradigm, as well as more subtle effects due to contextual factors such as reward contingencies or motivational factors. Furthermore, we show that the classical race model can be seen as a computationally simpler, perhaps neurally plausible, approximation to optimal decision-making. This conceptual link allows us to *predict* how the parameters of the race model, such as the stopping latency, should change with task parameters and individual experiences/ability.

Keywords: inhibitory control, optimal decision-making, speed-accuracy tradeoff, stop signal task

1 INTRODUCTION

Humans and animals often need to choose among actions with uncertain consequences, and to modify those choices according to ongoing sensory information and changing task demands. The requisite ability to dynamically modify or cancel planned actions is termed *inhibitory control*, considered a fundamental component of flexible cognitive control (Barkley, 1997; Nigg, 2000). In this paper, we examine optimal inhibitory control in the context of the widely studied stop signal paradigm (Logan and Cowan, 1984), where the subject's *go* response on a primary task, such as a two-alternative forced choice discrimination task, is interrupted by a *stop* signal on some trials. Subjects are instructed to withhold the go response on stop trials: a successful response cancelation is a stop success (SS), whereas a response is considered a stop error (SE, see **Figure 1**). Typically, SE rate increases as the presentation time of the stop signal is delayed with respect to the go stimulus onset, in a characteristic pattern known as the *inhibition function* (e.g., **Figure 5A**, adapted from Emeric et al., 2007). More subtly, SE reaction time (RT) tends to be faster than go RT (e.g., **Figure 5C**, adapted from Emeric et al., 2007).

The classical model for the stop signal task is the *race model* (Logan and Cowan, 1984), where behavioral outcome on each trial is conjectured to arise from the competition of two independent processes: *go* and *stop*. In this model, a stop trial results in error if the go process finishes before the stop process (**Figure 2A**). Thus, the average stopping latency, called stop signal reaction time (SSRT) by the race model, determines how successfully the observer can interrupt the go process. Since the stop process and its outcomes are unobservable, SSRT is estimated from the observed go RT distribution and the error rate on stop trials at a given stop signal delay (SSD). Specifically, if SSRT is constant, and the go and stop processes are independent, then the SE rate at a given SSD corresponds to the fraction of go RT distribution cut off by SSD + SSRT (the dark shaded region in **Figure 2A**). Conversely, given the SE rate for a particular SSD, SSRT can be estimated by subtracting SSD from the appropriate percentile of go RTs. Subject-specific SSRT is typically estimated by averaging the SSRT estimates from different SSDs, or simply taken as the SSRT best explaining the SSD giving rise to 50% error rate on stop strials. Consistent with the interpretation that SSRT measures some aspect of inhibitory ability, SSRT has been measured as longer in populations with presumed inhibitory deficits, such as attention-deficit hyperactivity disorder (Alderson et al., 2007), substance abuse (Nigg et al., 2006), and obsessive-compulsive disorder (Menzies et al., 2007). Although the race model yields an elegantly simple description of classical behavioral data, it does not address how different cognitive processes contribute to stopping behavior, thereby precluding the possibility of predicting how experimental manipulations of different cognitive factors, such as reward and context, should affect stopping behavior. Relatedly, it cannot readily differentiate the underlying cognitive deficits seen in the various clinical populations.

We present a rational decision-making framework for inhibitory control in the stop signal task. In our framework, sensory processing and action choice are optimized relative to a quantitative, global behavioral objective function that takes into account the costs associated with go errors, stop errors, and response delay. Classical behavioral results in the stop signal task are shown to be *natural consequences* of rational decision-making in the task. Moreover, the model can quantitatively predict changes in stopping behavior as a consequence of manipulations in task demands such as reward contingencies (Leotti and Wager, 2009). We show that a drift-diffusion model implementation of the race model

(e.g., Verbruggen and Logan, 2009b) can be seen as a simpler approximation to optimal decision-making, whereby parameters of the race model, such as the SSRT, must vary with task parameters in a systematic way to maintain the best approximation to optimal decision-making. Thus, race model-like behavior, including the well-studied SSRT, can be understood as emergent properties of rational decision-making. Altogether, our results suggest that cognitive control plays a critical role in stopping behavior, and the brain implements optimal or near-optimal decision-making, possibly via a race-model-like process, in an adaptive and context-dependent manner.

2 METHODS

Our computational model consists of two main components: (1) a *monitoring process*, which models sensory inference and learning about the identity, prevalence, and timing of the stimuli as hierarchical Bayesian inference, and (2) a *decision process*, formalized in terms of stochastic control theory, that translates the current expectations based on sensory evidence into a choice of action. In our model, the available actions at any given time include the two possible go responses and waiting one more time step. Repeated selection of the *wait* option results in a *stop* response on a given trial.

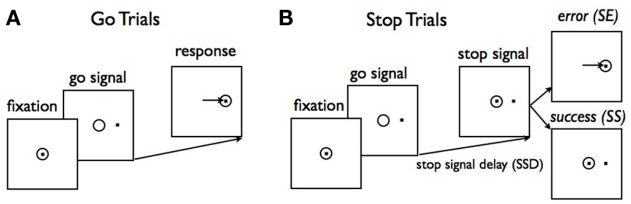

FIGURE 1 | Schematic illustration of a saccadic version of the stop signal task. (A) On a majority of trials (go trials), a central fixation dot is followed by one of two targets requiring a saccade to the indicated location. **(B)** On stop trials, the target presentation is followed after a short delay (SSD) by reappearance of the fixation point. A saccade on a stop trial is a stop error (SE), and a successfully canceled movement is a stop success (SS). Figure adapted from Hanes and Schall (1995).

Consistent with typical experimental design, the model assumes that subjects are given a deadline for responding, and that they must withhold the go response for the same amount of time to indicate a stop response.

2.1 THE MONITORING PROCESS: BAYESIAN STATISTICAL INFERENCE

The monitoring process in our model tracks sensory information about the go and stop stimuli during each trial, integrating it with prior belief about the distribution of go stimulus identity, and the prevalence and timing of the stop signal. **Figure 3** shows the generative model for sensory evidence in the task. The model assumes that subjects believe there are two hidden variables, corresponding to the identity of the go stimulus ($d = \{0,1\}$), and whether or not the current trial is a stop trial ($s = 1$ for stop, $s = 0$ for go). Priors over d and s reflect experimental parameters: $P(d = 1) = 0.5$, $P(s = 1) = r = 0.25$, where r is the *prior probability* that a trial is a stop trial. Conditioned on d, a stream of independent and identical (iid) inputs are generated on each trial, $x^1, \ldots, x^t$ where t indexes time within a trial from go signal onset, and the likelihood functions are $p(x^t|d = 0) = f_0(x^t)$ and $p(x^t|d = 1) = f_1(x^t)$. Without loss of generality, we assume f_0 and f_1 to be Bernoulli distributions with rate parameters q_d and $1 - q_d$, respectively. The dynamic variable z^t denotes the presence/absence of the stop signal: if the stop signal appears at time θ then $z^1 = \cdots = z^{\theta-1} = 0$ and $z^\theta = \cdots = 1$. On a go trial, $s = 0$, the stop signal of course never appears, $P(\theta = \infty) = 1$. On a stop trial, $s = 1$, we assume that the onset of the stop signal has a constant hazard rate, i.e., θ is generated from a geometric distribution: $p(\theta|s = 1) = (1 - \lambda)^{\theta - 1}\lambda$. Conditioned on z^t, there is a second, conditionally independent, stream of observations associated with the stop signal: $p(y^t|z^t = 0) = g_0(y^t)$, and $p(y^t|z^t = 1) = g_1(y^t)$. Again, we assume for simplicity that g_0 and g_1 are Bernoulli distributions with rate parameters q_s and $1 - q_s$, respectively.

The counterpart to the generative model is the *recognition model*, which specifies statistically optimal reverse-inference of the hidden variables based on the continual stream of sensory inputs. In the stop signal task, this means computing the posterior probability about go stimulus identity, $p_d^t \triangleq P(\mathrm{d} = 1 \,|\, \mathbf{x}^t)$, and the stop signal presence, $p_z^t \triangleq P\{\theta \le t \,|\, \mathbf{y}^t\}$, where $\mathbf{x}^t \triangleq \{x^1, \ldots, x^t\}$ denotes all the sensory inputs associated with the go stimulus, and

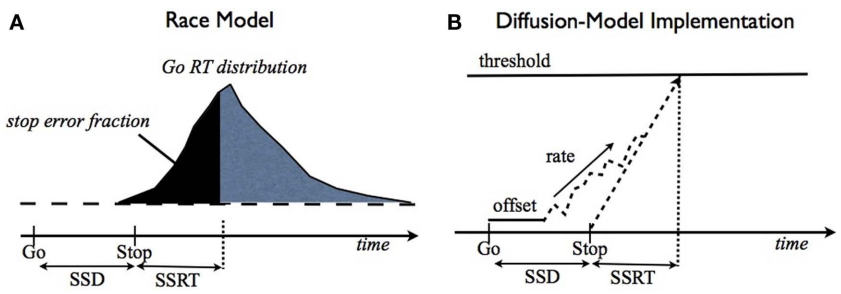

FIGURE 2 | Race model and drift-diffusion model for the stop signal task. (A) The race model (Logan and Cowan, 1984) proposes that the finishing times of two independent (go and stop) processes determine trial outcome: stop or go, depending on which finishes first. SSD + SSRT (stop signal reaction time) specifies the finishing time of the stop process, and determines what fraction of go trials will finish earlier and therefore result in a stop error. **(B)** An implementation of the race model using a drift-diffusion process, similar to (Verbruggen and Logan, 2009b). The go process consisting of a constant *drift rate* corrupted by additive Wiener (Gaussian, white) noise on each time step, a temporal *offset* (also known as the non-decision time), and a *threshold* for evoking the go response. The stop process is assumed to initiate at a time SSD after the go stimulus, and to take a time of SSRT (assumed to be fixed here) to reach the threshold. Whichever process finishes first determines trial outcome: *go* or *stop*.

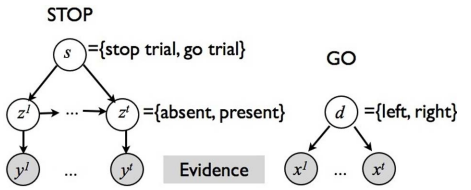

FIGURE 3 | Graphical model for sensory input generation in our Bayesian model. Two separate streams of observations, $\{x^1, \ldots, x^t, \}$ and $\{y^1, \ldots, y^t, \}$, are associated with the go and stop stimuli, respectively. x_t depend on the identity of the target, $d \in \{0,1\}$. y^t depends on whether the current trial is a stop trial, $s = \{0,1\}$, and whether the stop signal has already appeared by time $t, z^t \in \{0,1\}$.

$\mathbf{y}^t \triangleq \{y^1, \ldots, y^t\}$ denotes all the sensory inputs associated with the stop signal observed so far. Note that $P(d=0 \mid \mathbf{x}^t) = 1 - p_d^t$, and $P(\theta > t \mid \mathbf{y}^t) = 1 - p_z^t$.

The computation of p_d^t is a simple application of Bayes' rule:

$$p_d^t = \frac{p_d^{t-1} f_1(x^t)}{p_d^{t-1} f_1(x^t) + \left(1 - p_d^{t-1}\right) f_0(x^t)} = \frac{p_d^0 \Pi_{i=1}^t f_1(x^i)}{p_d^0 \Pi_{i=1}^t f_1(x^i) + (1 - p_d^0) \Pi_{i=1}^t f_0(x^i)}. \quad (1)$$

Inference about the stop signal is slightly more complicated due to the hidden dynamics in z^t (going from signal-absent to signal-present at a stochastic onset time). We first compute p_z^t, the posterior probability that the stop signal has already appeared.

$$p_z^t = \frac{g_1\left(y^t\right)\left(p_z^{t-1} + (1 - p_z^{t-1})h(t)\right)}{g_1\left(y^t\right)\left(p_z^{t-1} + (1 - p_z^{t-1})h(t)\right) + g_0(y^t)(1 - p_z^{t-1})(1 - h(t))} \quad (2)$$

where $p_z^0 = 0$ (stop signal never occurs at the same time as the go signal), and $h(t)$ is the conditional probability that the stop signal will appear in the next instant given it has not appeared already, $h(t) \triangleq P(\theta = t \mid \theta > t-1, \mathbf{y}^{t-1})$:

$$h(t) = \frac{rP\left(\theta = t \mid s = 1\right)}{rP\left(\theta > t-1 \mid s = 1\right) + \left(1 - r\right)} = \frac{r\left(1-\lambda\right)^{t-1}\lambda}{r\left(1-\lambda\right)^{t-2} + \left(1 - r\right)} \quad (3)$$

where, recall, r is the prior probability of a stop trial. $h(t)$ does not depend on the observations, since *given* that the stop signal has not yet appeared, whether it will appear in the next instant does not depend on previous observations.

In the stop signal task, a stop trial is typically considered a stop trial even if the subject makes the go response before the stop signal. Following this experimental convention, we need to compute the posterior probability that the current trial is a stop trial, denoted $p_s^t \triangleq P(s=1 \mid \mathbf{y}^t)$, which depends both on the current belief about the presence of the stop signal, and the expectation that it will appear in the future:

$$p_s^t = p_z^t \cdot 1 + \left(1 - p_z^t\right) \cdot P\left(s = 1 \mid \theta > t, \mathbf{y}^t\right) \quad (4)$$

where $P(s=1 \mid \theta > t, \mathbf{y}^t) = P(s=1 \mid \theta > t)$ again does not depend on past observations:

$$P\left(s = 1 \mid \theta > t\right) = \frac{P(\theta > t \mid s=1)P(s=1)}{P(\theta > t \mid s=1)P(s=1) + P(\theta > t \mid s=0)P(s=\mathbf{0})} = \frac{(1-\lambda)^{t-1} \cdot r}{(1-\lambda)^{t-1} \cdot r + 1 \cdot (1-r)} \quad (5)$$

The *belief state* at time t, the vector $\mathbf{b}^t = (p_d^t, p_s^t)$, represents all the information the ideal observer has about the stimulus properties on the current trial.

Figure 4A illustrates the behavior of this inference procedure, averaged across trials, for different types of trials. The evolution of the beliefs corresponding to the identity of the go stimulus (p_d), and whether the trial is a stop trial (p_s), are shown on trials without a stop signal (GO), as well as successful (SS) and error (SE) stop trials; in all examples, true $d = 1$. Over time, p_d increases in all three kinds of trials as sensory evidence about the go stimulus accumulates. On the other hand, p_s shows an initial rise due to prior expectation and then either decays to 0 on GO trials, or rises toward 1 on stop trials. Individual trajectories are stochastic due to noise in the sensory inputs. This stochasticity induces a go response on some stop trials and not others: stop error trials (non-canceled trials) are those on which the go stimulus belief state happens to be rising fast, and the stop signal is processed slower than average. Successful stop trials show the opposite trend.

2.2 THE DECISION PROCESS: OPTIMAL STOCHASTIC CONTROL

Based on the belief state, subjects have to make a decision at each moment about whether to go now or wait at least one more time point in case this is a stop trial; and if they wait, they need to make the same decision again using one more data point. To model this decision process, we again assume an ideal observer implementing a Bayes-optimal decision policy. To say what is *optimal*, we need to specify a loss function that captures the reward structure of the task, against which the decision policy can be optimized. We assume there is a time cost of c per unit time on each trial to capture the opportunity cost of not responding quickly. Consistent with experimental design, we also assume a deadline D for responding on go trials and for determining a subject has withheld a go response long enough on a stop trial. In addition, there is a penalty c_s for choosing to respond on a stop signal trial, and a unit cost for making an error on a go trial (by choosing the wrong discrimination response or exceeding the deadline for responding). Because only the relative costs matter in the optimization, we can normalize the coefficients associated with all the costs such that one of them is unit cost. Let τ denote the trial termination time, so that $\tau = D$ if no go response is made before the deadline, and $\tau < D$ if a response is made. On each trial, the policy π produces a response time $\tau \leq D$, as well as a binary response $\delta \in \{0,1\}$ if a go response is made ($\tau < D$). The loss function is:

$$l(\tau, \delta; d, s, \theta, D) = c\tau + c_s \mathbf{1}_{\{\tau < D, s=1\}} + \mathbf{1}_{\{\tau < D, \delta \neq d, s=0\}} + \mathbf{1}_{\{\tau = D, s=0\}} \quad (6)$$

where $\mathbf{1}\{\cdot\}$ is the indicator function, evaluating to 1 if the conditions in $\{\cdot\}$ are met, and 0 otherwise. The optimal decision policy minimizes the average or *expected* loss, $L_\pi \triangleq \left\langle l(\tau, \delta; d, s, D) \right\rangle$,

$$L_\pi = c\left\langle \tau \right\rangle + c_s r P(\tau < D \mid s=1) + (1-r)P(\tau < D, \delta \neq d \mid s=0) + (1-r)P(\tau = D \mid s=0).$$

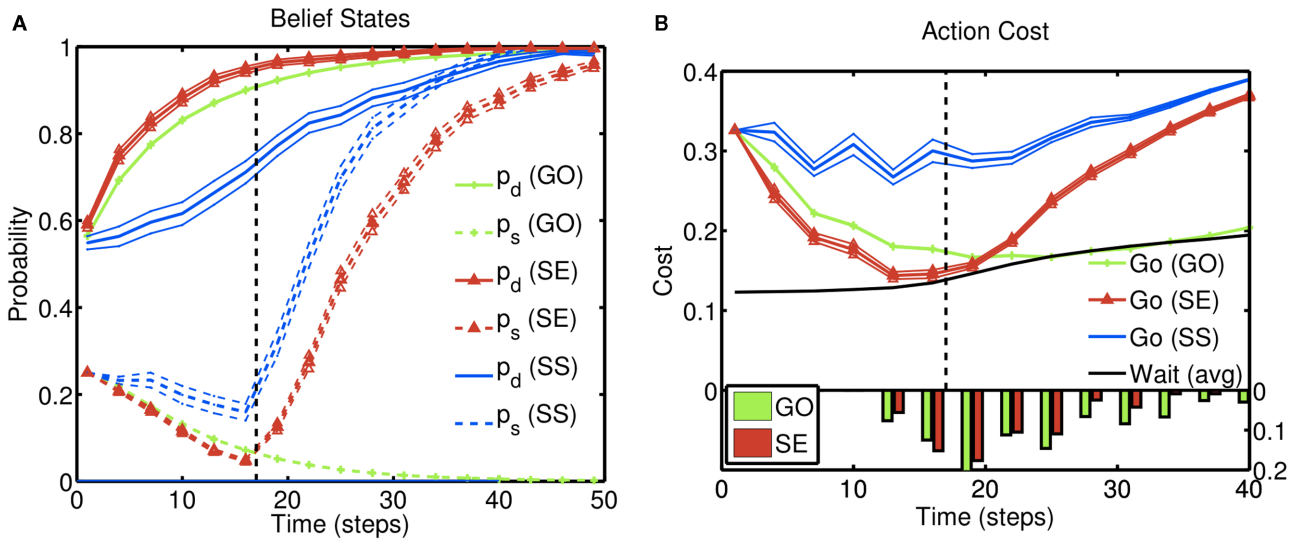

FIGURE 4 | Inference and action selection in the stop signal task. (A) Evolution of the belief state over time, on go trials (green), successful stop trials (SS; blue), and error stop trials (SE; red). Solid lines represent the posterior probabilities assigned to the true identity of the go stimulus (one of two possibilities) for the three types of trials – they all rise steadily toward the value 1, as sensory evidence accumulates. The dashed black vertical line represents the onset of the stop signal on stop trials. The probability of a stop signal being present (dashed lines) rises initially in a manner dependent upon prior expectations of frequency and timing of the stop signal, and subsequently rises farther toward the value 1 (stop trials), or drops to 0 (go trials), based on sensory evidence. **(B)** Average action costs corresponding to going (Q_g^t, see text) and waiting (Q_w^t, see text), using the same sets of trials as **(A)**. The black dashed vertical line denotes the onset of the stop signal. A response is initiated when the cost of going drops below the cost of waiting. The RT histograms for go and error stop trials (bottom) indicate the temporal distribution of when the go cost crosses the stop cost in each simulated trial. Each data point is an average of 10,000 simulated trials. Error bar = SEM. Simulation parameters: $q_d = 0.68$, $q_s = 0.72$, $\lambda = 0.2$, $r = 0.25$, $D = 50$, $c_s = 0.2$, $c_s = 0.004$. See section 2 for definition of parameters. Unless otherwise specified, these parameters were used in all subsequent simulations.

Minimizing L_π over the policy space directly can seem computationally daunting, since there is no obvious parameterization of the policy space; however, Bellman's dynamic programming principle (Bellman, 1952) provides an iterative relationship in terms of the *value function* (defined in terms of costs here), $V^t(\mathbf{b}^t)$

$$V^t(\mathbf{b}^t) = \min_a \left[\int p(\mathbf{b}^{t+1} \mid \mathbf{b}^t; a) V^{t+1}(\mathbf{b}^{t+1}) d\mathbf{b}^{t+1} \right],$$

where $\mathbf{b}^t = (p_d^t, p_z^t)$ is the belief state, and a ranges over all possible actions.

In our model, *the action (decision) space* consists of {*go*,*wait*}, with the corresponding expected costs (also known as Q-factors), $Q_g^t(\mathbf{b}^t)$ and $Q_w^t(\mathbf{b}^t)$, respectively. Note that our model produces a "stop" response on a trial by repeatedly deciding to *wait* rather than *go* until the deadline is reached.

$$Q_g^t(\mathbf{b}^t) = ct + c_s p_s^t + \left(1 - p_s^t\right) \min\left(p_d^t, 1 - p_d^t\right) \tag{7}$$

$$Q_w^t(\mathbf{b}^t) = 1_{\{D>t+1\}} \left\langle V^{t+1}(\mathbf{b}^{t+1}) \mid \mathbf{b}^t \right\rangle_{\mathbf{b}^{t+1}} + \mathbf{1}_{\{D=t+1\}} \left(c(t+1) + 1 - p_s^t\right) \tag{8}$$

$$V^t(\mathbf{b}^t) = \min\left(Q_g^t, Q_w^t\right) \tag{9}$$

The optimal decision policy chooses the action corresponding to the smaller Q-factor, and the value function is the smaller of the Q-factors Q_g^t and Q_w^t. Note that the *go* action results in either $\delta = 1$ or $\delta = 0$, depending on whether p_d^τ is greater or smaller than 0.5, respectively. The dependence of Q_w^t on V^{t+1} allows us to recursively compute the value function backward in time. Given V^{t+1} and $\mathbf{b}^t$, we can compute $\langle V^{t+1} \rangle$ by summing over the uncertainty about the next observations x^{t+1}, y^{t+1}, since the belief state $\mathbf{b}^{t+1}$ is a deterministic function of $\mathbf{b}^t$ and the observations (Eqs 1 and 4).

$$\left\langle V^{t+1}(\mathbf{b}^{t+1}) \mid \mathbf{b}^t \right\rangle_{\mathbf{b}^{t+1}} = \sum_{x^{t+1}, y^{t+1}} p\left(x^{t+1}, y^{t+1} \mid \mathbf{b}^t\right) V^{t+1}\left(\mathbf{b}^{t+1}; \mathbf{b}^t, x^{t+1}, y^{t+1}\right) \tag{10}$$

$$\begin{aligned} p\left(x^{t+1}, y^{t+1} \mid \mathbf{b}^t\right) &= p\left(x^{t+1} \mid p_d^t\right) p\left(y^{t+1} \mid p_s^t\right) \\ p\left(x^{t+1} \mid p_d^t\right) &= p_d^t f_1(x^{t+1}) + (1 - p_d^t) f_0(x^{t+1}) \\ p\left(y^{t+1} \mid p_s^t\right) &= \left(p_z^t + (1 - p_z^t) h(t+1)\right) g_1(y^{t+1}) \\ &\quad + (1 - p_z^t)(1 - h(t+1)) g_0(y^{t+1}) \end{aligned} \tag{11}$$

The initial condition of the value function can be computed exactly at the deadline since there is only one outcome (subject is no longer allowed to go or stop): $V^D(\mathbf{b}^D) = cD + (1 - p_s^D)$. We can then compute $\{V^t\}_{t=1}^D$ and the corresponding optimal decision policy backward in time from $t = D - 1$ to $t = 1$. In our simulations, we do so numerically by discretizing the probability space for p_s^t into 1000 bins; p_d^t is represented exactly using its sufficient statistics.

Figure 4B illustrates how these action costs evolve over the course of a trial, averaged across trials of different types as before: go (GO) trials, stop error (SE) trials, and successful stop (SS) trials. Since the probability associated with the (correct) go stimulus identity increases with accumulating sensory evidence, the cost of going drops, eventually crossing the cost of waiting and triggering a go response. On stop trials, the onset of the stop signal initiates an increase in the cost of going, when the cost of a stop error is factored in. In error stop trials, the go cost (Q_g^t) crosses the wait cost (Q_w^t) before the stop stimulus is fully processed. In successful stop trials, the go cost never dips below the wait cost. The RT histograms for go and error stop trials illustrate that, although the *average* go cost trajectories do not cross the average wait cost, *every individual trajectory* crosses over at some point on each trial.

2.3 APPROXIMATING OPTIMAL DECISION-MAKING BY A RACE MODEL

We make the relationship between optimal decision-making and race-like behavior concrete by considering a specific implementation of the race model. One reason for examining this connection is that race-like processes may serve as a neural implementation of behavior in the stopping task (Hanes et al., 1998; Pare and Hanes, 2003; Boucher et al., 2007). In particular, we examine a diffusion model implementation which has long been used to model reaction times (Stone, 1960; Laming, 1968; Ratcliff, 1978; Luce, 1986; Hanes and Schall, 1996; Gold and Shadlen, 2002; Mazurek et al., 2003; Bogacz et al., 2006). Variants of the drift diffusion model have also been applied specifically to the stop signal task (Hanes and Carpenter, 1999; Verbruggen and Logan, 2009b).

Our implementation is illustrated in **Figure 2B**, where the go process consists of a constant drift *rate* with a starting point or *offset,* and additive, cumulative Gaussian noise on each time step of a trial. The stop process is modeled as a fixed-latency process with a corresponding latency parameter (SSRT). Although we could easily consider a stochastic stop process with its attendant rate and threshold, we specifically wish to model SSRT as measured in practice, i.e., by using a constant-SSRT assumption (Logan and Cowan, 1984). Finally, go responses are initiated by the process crossing a *threshold*, unless it is at a time exceeding SSD + SSRT, which is the finishing time of the stop process – in the latter case, no response is produced. For each condition in the reward manipulation task (Section 3.2), we select values for these four parameters, rate, offset, threshold, and SSRT, in order to best approximate the cumulative RT and stop error distributions of the optimal model.

The basic drift-diffusion model has the following form:

$$d\kappa = \beta dt + \varepsilon d\mathrm{W} \tag{12}$$

where the rate parameter β determines the direction (positive or negative) and speed of the average movement of the dynamic parameter κ, and dW denotes Wiener noise, such that $\varepsilon\ d$W is normally distributed with mean 0 and variance $\varepsilon^2 dt$. In practice, we simulate this process by discretizing time and approximating it with a random-walk:

$$k(t+1) = k(t) + \beta + w \tag{13}$$

with drift parameter β, and normally distributed noise $w \sim \mathcal{N}(0, \sigma^2)$. We assume that a decision is made when $k(\tau)$ first crosses the threshold h or $-h$, whichever first. The simulated RT for a trial is $\tau + \tau_0$, where τ_0 is a temporal offset parameter denoting non-decision time (Ratcliff, 1978; Bogacz et al., 2006). For $\beta > 0$, the response is correct if h is first crossed, and incorrect if $-h$ is first crossed; vice versa if $\beta < 0$. We note that there is redundancy in the parameterization, such that β, σ, and h can all be scaled by the same constant and remain an identical process; we can therefore fix $\sigma = 1$ without loss of generality. Thus, the go process has three free parameters: β, h, and τ_0.

The stop process, as typically modeled in the literature, (Logan and Cowan, 1984), is assumed to have a fixed finishing time of SSD + SSRT. Since SSD is given by experimental design, SSRT is the only free parameter for the stop process. On stop trials, if $\tau + \tau_0 <$ SSD + SSRT, an error response occurs at $\tau + \tau_0$; otherwise, it is a correct stop trial and τ is assumed to take on the value ∞.

Our goal is to find a diffusion model approximation to optimal decision-making. To do this, we compare the joint distribution of RT and choice based on simulation of the optimal model, $p(\tau,\delta)$, and that from a parametrization of the race model $p(\tau,\delta|\boldsymbol{\theta})$. We look for settings of the diffusion model parameters $\boldsymbol{\theta} \triangleq (\beta, h, \tau_0, \mathrm{SSRT})$ that would minimize the KL divergence between the two distributions:

$$\begin{aligned}
\boldsymbol{\theta}^* &= \arg\min_{\boldsymbol{\theta}} \mathrm{KL}\left[p(\tau,\delta), p(\tau,\delta \mid \boldsymbol{\theta}) \right] \\
&= \arg\min_{\boldsymbol{\theta}} \int p(\tau,\delta) \log \frac{p(\tau,\delta)}{p(\tau,\delta \mid \boldsymbol{\theta})} d\tau d\delta \\
&\approx \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{n} \log \frac{p(\tau_i,\delta_i)}{p(\tau_i,\delta_i \mid \boldsymbol{\theta})} \ \text{ where } \ (\tau_i,\delta_i) \sim p(\tau,\delta) \\
&= \arg\min_{\boldsymbol{\theta}} - \sum_{i=1}^{n} \log p(\tau_i,\delta_i \mid \boldsymbol{\theta}) \\
&= \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{n} \log p(\tau_i,\delta_i \mid \boldsymbol{\theta})
\end{aligned} \tag{14}$$

where the approximation comes from the fact that the expectation of the log likelihood ratio $\langle \log p(\tau,\delta)/p(\tau,\delta|\boldsymbol{\theta}) \rangle$ under the distribution $p(\tau,\delta)$ can be approximated by a finite sum based on samples from $p(\tau,\delta)$ – the approximation becomes exact as $n \to \infty$. We note the interesting observation that minimizing the KL divergence is identical to minimizing the coding cost of the samples from the true (optimal) distribution by the approximate diffusion model distributions; it is also identical to maximum likelihood estimation of parameters of the approximate model as a function of the samples generated from the true (optimal) model.

To evaluate the sum, we generate n = 10,000 samples from the optimal model: $\{(\tau_i,\delta_i)\}_{i=1}^n$, where the probability of each trial being a stop trial is r, and the probability of the go stimulus being each of the two alternatives is 0.5, both given by the actual experimental design in question. We can compute $p(\tau,\delta|\boldsymbol{\theta})$ exactly, up to a discretization of values of $k(t)$, by convolving $p(k(t))$ with standard-normal noise and removing probability mass beyond both thresholds at the next timestep, to get $p(k(t+1))$. This gives $p(\tau,\delta|\boldsymbol{\theta})$ on go trials. On stop trials, we truncate the distribution at SSD + SSRT, which then gives us both the error stop trial RT distribution, as well as the error rate, for each SSD. We then sum

the log likelihood of each sample (τ_i,δ_i) under $p(\tau,\delta|\boldsymbol{\theta})$ over all 10,000 samples. We do so for each setting of the diffusion model parameters $\boldsymbol{\theta}$, and use Matlab's `fmincon` function to find the best-fitting parameters.

3 RESULTS

3.1 STOPPING BEHAVIOR AS A NATURAL CONSEQUENCE OF RATIONAL DECISION-MAKING

In the stop signal task, subjects typically make more stop errors when the stop signal delay (SSD) is longer, and response times on stop error trials are on average faster than go RTs (Logan and Cowan, 1984; Hanes and Schall, 1995; also see Verbruggen and Logan, 2009a for a recent review). **Figure 5** shows how this behavior arises naturally as a consequence of rational decision-making in the task. Data from human subjects performing a saccade version of the stop signal task (Emeric et al., 2007; **Figures 5A,C**), and from model simulations (**Figures 5B,D**), show the same characteristics: error rate increases as SSD increases, and RTs on stop error trials are on average faster than go trials.

Intuitively, the later the stop signal, the more likely that the cost of going has already dropped below the cost of waiting before the stop signal information can be factored in (see **Figure 4B**), leading to the increasing SE curve or *inhibition function* shown here. Faster RT on SE trials is an outcome of stochasticity in the processing of the go and stop signals; as shown in **Figures 4A,B**, stop error trials are those in which the go stimulus is processed faster than average (and the stop stimulus slower than average). This difference gives rise to the observed faster RT, illustrated by the histograms in **Figure 4B**.

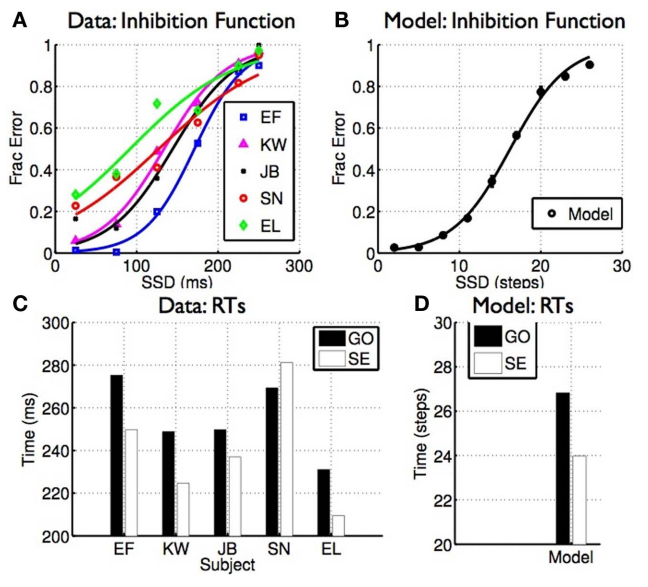

FIGURE 5 | Classical properties of stopping behavior arise naturally from optimal decision-making. (A) Inhibition function: errors on stop trials increase as a function of SSD. **(B)** Similar inhibition function seen for the model. **(C)** Discrimination RT is generally faster on stop error trials than go trials. **(D)** Similar results seen in the model. **(A,C)** Data adapted from Emeric et al. (2007) with permission from Elsevier.

The race model explains these results as well, utilizing a similar proximate explanation: later initiation of the stop process allows more go trials to "escape," giving rise to the form of the inhibition function; stochasticity in the go process allows the go process to sometimes escape the stop process, and those that do happen to escape have shorter finishing times (Logan and Cowan, 1984). A critical difference is that by focusing on the *finishing times* of the stop and go processes, but not their underlying computational import, the race model cannot predict *a priori* the effect of changes in experimental constraints on stopping behavior. We elaborate further on this contrast by considering the effect of reward manipulations on stopping behavior.

3.2 INFLUENCE OF REWARD STRUCTURE ON STOPPING

Leotti and Wager (2009) showed that subjects can be biased toward stopping or going by experimentally manipulating the relative penalties associated with go and stop errors. Their experiments associated a reward for fast go response times and penalty for stop errors, and manipulated these values in an iterative fashion to induce a particular degree of bias in each subject, as measured by the fraction of stop errors committed. **Figures 6A,B** shows that as subjects are biased toward stopping, they make fewer stop errors and have slower go responses. Since our model explicitly parametrizes the relative costs of go and stop errors (c_s in Eq. 6), we can easily simulate such a manipulation by setting c_s to a higher or lower value in Eq. 6. The new cost function induces different statistics in the trajectories of the action costs as in **Figure 4B**. In particular, making c_s larger makes expected go cost higher, as the same probability of a stop trial lead to a greater stop error cost, and this has the effect of slowing the initial downward trajectory of the go cost curve, and speed its repulsion away from the wait cost if later the stop signal is introduced – the over all effect, is to slow down go reaction times and lower stop error frequency. Simulation results from the optimal model (**Figures 6D,E** filled) confirm these intuitions and are similar to subjects' actual behavior (top row).

Also shown in **Figures 6C,F** is the measure of stopping latency (SSRT) assumed by the race model, for human behavior and for the optimal model. Since the race model's conjectured stop process is not observable, the SSRT must be *inferred* from the go RT distribution and the stop error distribution. In particular, if going and stopping are assumed independent, and the SSRT is approximated as constant, then the difference between the midpoints of the RT and stop error cumulative distribution functions is an estimate of the SSRT (Logan and Cowan, 1984). Note, however, that when this estimation process is applied to human data in the experiment, the SSRT *changes with reward manipulation* (**Figure 6C**), and therefore cannot be used in isolation as a subject-specific index of inhibition. Although SSRT is not an explicit component of our framework, nevertheless the same procedure outlined above can be used to estimate it for our model simulations, yielding the very same trend (**Figure 6F**, filled). The close match with human behavior suggests that SSRT is an *emergent property* of the interaction between going and stopping, and variations in SSRT are directly explained by optimal adjustments to the tradeoff between them.

3.3 RACE MODEL AND SSRT AS EMERGENT PROPERTIES OF OPTIMAL BEHAVIOR

We examine the relationship between the race model and optimal behavior by fitting a diffusion model implementation of the race model to output from the optimal model (**Figure 2B**, see section 2 for details). We examined how parameters of the best-fitting diffusion model vary as the reward structure of the task is manipulated (i.e., c_s takes on different values). The best-fitting parameters are shown in **Figure 7**, and indicate that the SSRT parameter indeed has to be adjusted in a manner consistent with our optimal model's predictions, as well as the experimental data (Leotti and Wager, 2009). The fit also shows that the rate and threshold do not vary substantially. However, the offset parameter (non-decision time) increases with increasing stop error cost – this is consistent with later response times, without apparent informational gain, as c_s increases. In general, the best-fitting race model for each c_s behaves very similarly to the optimal model (**Figures 6D–F**, unfilled).

Figure 8 shows the race model fits resulting from this optimization procedure, with (**Figures 8A,B**) showing the reaction time distribution of GO and stop error trials, as well as the cumulative SE distributions from the optimal model. **Figures 8C,D** show that the resulting race model fit is able to approximate the RT distributions and the stop error distribution functions qualitatively well, as a result of the optimization procedure selecting the appropriate race model parameters for each condition.

In summary, optimal decision-making may be implemented by a suitably parameterized race-diffusion model, suggesting one possible neural mechanism for behavior in the task. Furthermore, with an explicit procedure for fitting the race model to the optimal model, we can *predict a priori* how the race model parameters, such as SSRT, should change under different experimental manipulations, since the optimal model encodes the experimental parameters in a principled manner and gives precise predictions of associated behavioral changes.

4 DISCUSSION

We presented a rational decision-making framework for inhibitory control in the stop signal task. Our framework optimizes sensory processing and action choice relative to a quantitative, global behavioral objective function that explicitly takes into account the various costs associated with go errors, stop errors, and response

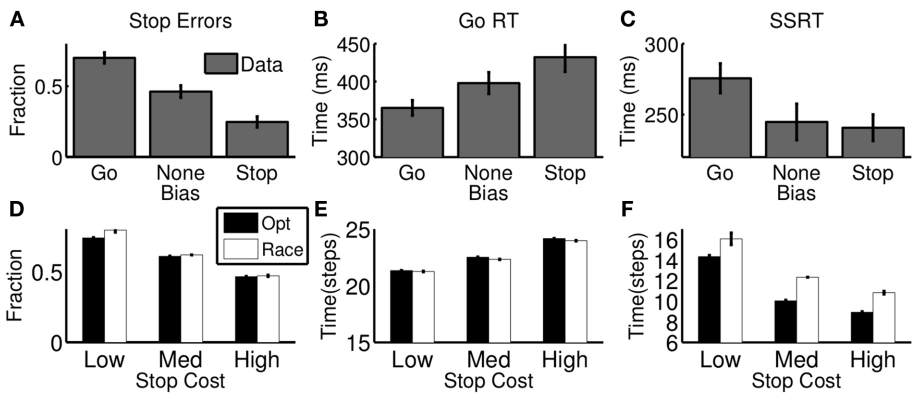

FIGURE 6 | Effect of reward manipulation on stopping behavior. (A–C) Data from human subjects performing a variant of the stop signal task where the ratio of rewards for quick go responses and successful stopping was varied, inducing a bias toward going or stopping (adapted from Leotti and Wager, 2009 with permission from APA). As stop errors are punished more severely, subjects have lower stop error rate **(A)**, slower go RT **(B)**, and faster SSRT **(C)**; low stop error penalty induces the opposite pattern. **(D–F)** The optimal decision model (black) and its best-fitting race model approximation (white) show similar trends as a function of stop error penalty (relative to go errors). "High," "Med," "Low" refer to high (c_s = 0.5, medium (c_s = 0.25), and low (c_s = 0.15) stop error penalty, respectively. For model simulations **(D-F)**, each bar denotes average of 10 simulated "sessions," each session consisting of 10,000 trials. Error bar = SEM.

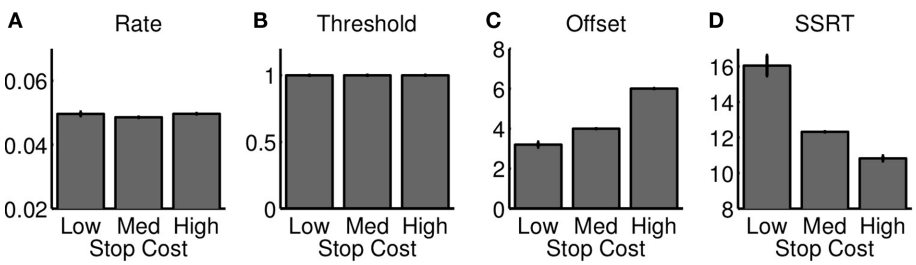

FIGURE 7 | Race model approximation to optimal decision-making as stop error penalty is varied. The figure shows the best race parameters, implemented as a diffusion model shown in **Figure 2**, that approximate the behavior of the optimal model, as the cost of stop errors is changed. The temporal offset to the start of the diffusion process increases **(C)**, and SSRT decreases **(D)**; rate and threshold parameters are unaffected **(A, B)**. Changes in SSRT are similar to those in optimal model and experimental data (**Figure 5**). Each bar denotes average of 10 simulated "sessions," each session consisting of 10,000 trials. Error bar = SEM.

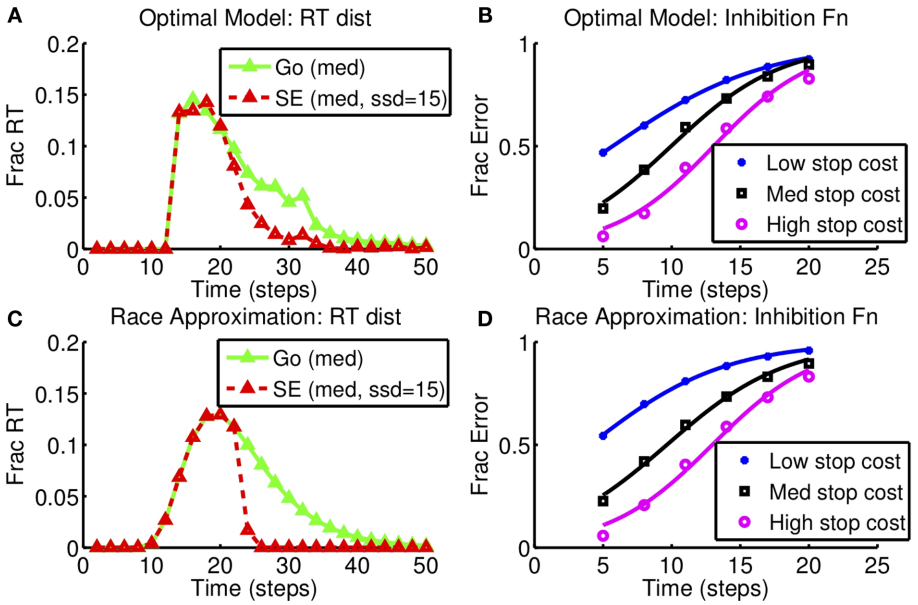

FIGURE 8 | Race model approximation to optimal decision-making as stop error penalty is varied: inhibition function and RT distribution. (A,B) Optimal model. **(C,D)** Diffusion race model. Left: RT distributions for GO and SE trials. SSD = 15, $c_s = 0.25$, as in **Figure 6**. Right: Inhibition function for different stop error costs (low: $c_s = 0.15$, med: $c_s = 0.25$, high: $c_s = 0.5$). Results based on 10,000 simulated trials from the optimal model, and also from the corresponding best-fitting diffusion race model.

delay. We show that classical behavioral results in the stop signal task are natural consequences of rational decision-making. Moreover, the model can quantitatively predict the influence of subtle manipulations of task parameters, such as reward contingencies (Leotti and Wager, 2009), on stopping behavior. Our results suggest, therefore, that cognitive processing in the task is a continual, intertwined choice between go and wait (stop), under the influence of multiple cognitive factors in a computationally optimal manner.

We also examined the relationship between the race model and the rational decision-making model. The two models are motivated by fundamentally different levels of analysis, corresponding to algorithmic and computational models in Marr's (1982) levels of analysis. Despite its elegant simplicity and ability to explain a number of classical behavioral results, the descriptive nature of the race model precludes an a priori prediction of how behavior should change in order to accommodate various cognitive goals and task constraints. On the other hand, the optimal model requires complex computations unlikely to be directly implemented by the brain. Even if subjects' behavior is similar to model predictions, the brain may well implement a simpler approximation to the optimal algorithm. Recent studies suggest that the activity of neurons in the frontal eye fields (FEF; Hanes et al., 1998) and superior colliculus (Pare and Hanes, 2003) of monkeys could be implementing a version of the race model (Hanes et al., 1998; Boucher et al., 2007; Wong-Lin et al., 2009). Specifically, movement and fixation neurons in the FEF show responses that diverge on go and correct stop trials, indicating that they may encode computations leading to the execution or cancelation of movement. If the race model is an appropriate description of these neural activities, however, we showed that the race model (and its diffusion model elaboration) will need its parameters, such as SSRT, carefully adjusted in different task conditions in order to best approximate the optimal model, and in order to account for experimental data (Emeric et al., 2007; Leotti and Wager, 2009). Our framework can therefore guide the search for, and provide a computational understanding of, the neural mechanisms underlying stopping behavior. For example, we conjecture that FEF neurons represent and track the relative values of various available actions such as going, waiting, and cancelation.

In our model, the RT distribution is the outcome of an adaptively optimal policy acting on accumulated noisy sensory evidence, in light of the global objective function. Notably, the optimal policy is deterministic given a particular sequence of sensory inputs, so that stochasticity in response latency is entirely driven by stochasticity in sensory inputs, which determine RT variance and all other higher-order moments in the RT distribution. A related but distinct framework (Daunizeau et al., 2010) considers the restricted space of non-adaptive policies where a fixed stopping time is chosen at the outset of the trial, based on minimizing the expected cost for the chosen stopping time. It is non-adaptive in the sense that it chooses a mean stop time without considering the actual sequence of sensory inputs observed, and assumes variability around that mean to arise independently from a non-sensory origin. However, substantial experimental data suggest that simple perceptual decisions involve accumulation of evidence up to a bound, related to a specific confidence level in the probability space (and therefore dependent on the actual sequence of noisy inputs observed), rather than up to a chosen stopping time (see e.g., Gold and Shadlen, 2007 for a review). Moreover, from a theoretical perspective, optimal policies for the type of stopping problems, including the stop signal task considered here as well as the simpler two-alternative forced choice tasks (e.g., Gold and Shadlen, 2007), are known to live within the

adaptive policy space, and not in the very restrictive sub-class of non-adaptive policies (Wald and Wolfowitz, 1948; Chow et al., 1971). In particular, adaptive policies can better accommodate moment-by-moment changes in perceived sensory information (Kiani et al., 2008). We note here that the original race model is agnostic with respect to to the source of stochasticity in reaction times, taking it as the consequence of some inherent stochasticity in the unspecified go and stop processes. However, the race model can be implemented using a drift-diffusion model to make explicit the role of sensory noise in decision-making, as we demonstrate in our simulations.

In our model, a stop decision is implemented as a sequence of wait actions. Neurophysiological evidence from monkeys (Hanes et al., 1998) and humans (Aron et al., 2007a,b) suggest that successfully stopped actions may involve increased activity in certain neural populations such as the fixation neurons of the FEF, or cortical regions such as the inferior frontal gyrus and subthalamic nucleus implicated by human imaging studies. Studies in humans involving fMRI and tractography data suggest that the inferior frontal gyrus may implement a stop action via a hyperdirect pathway to the subthalamic nucleus (Aron et al., 2007a,b). One important and planned line of inquiry for our work is to consider a rational model with an explicit stop action, in order to better account for what is known about the neurophysiology of stopping.

The stop signal task is traditionally thought of as probing behavioral inhibition, whereas other tasks such as the Stroop and Eriksen tasks (Stroop, 1935; Eriksen and Eriksen, 1974) are thought to engage cognitive inhibition (see e.g., Nigg, 2000 for a taxonomy). In contrast to this view, the close correspondence between our rational decision-making model and human behavior at the task demonstrates the influence of multiple cognitive factors on stopping behavior. Our previous work also showed that behavior in the Eriksen task (Yu et al., 2009) can arise from Bayesian statistical inference in a bounded rational manner (Simon, 1956). An interesting challenge is to explore how performance measures from these various inhibitory control tasks relate to each other within individuals, both empirically and from a computational perspective (Friedman and Miyake, 2004).

One major aim of our work is to understand how stopping ability and SSRT arise from various cognitive factors. Our work shows that SSRT arises from number of contributing elements: reward/penalty-sensitivity, sensory processing rate, and top-down expectations such as that of stop signal frequency. Thus, SSRT should not be viewed as a unique, invariant measure of stopping ability for each subject, but rather as an emergent property of the dynamic, context-dependent comparison between going and stopping. This more nuanced view of stopping ability and SSRT may aid in the careful analysis of impaired stopping ability, e.g., longer measured SSRTs, in a number of psychiatric and neurological conditions, such as substance abuse (Nigg et al., 2006), attention-deficit hyperactivity disorder (Alderson et al., 2007), schizophrenia (Badcock et al., 2002), obsessive-compulsive disorder (Menzies et al., 2007), Parkinson's disease (Gauggel et al., 2004), Alzheimer's disease (Amieva et al., 2002), and so on. It is unlikely that these various conditions share an identical set of underlying neural and cognitive deficits. In our framework, almost all model parameters, such as the fraction of stop trials, the SSD distribution, stop error cost, and go response deadline, are set directly by the experimental design. The only exceptions are parameters representing the sensory noise corrupting go stimulus and stop signal processing. These sensory parameters may be one important source of inter-subject differences. However, it is also likely that in practice, individuals have different estimates for the other parameter values on any given trial in an experiment, given their prior biases, memory capacity, individual experiences, and learning rates. Since our model makes explicit the dependence of subject behavior on subtle differences in these subject-specific parameters, these parameters can be inferred from behavioral data directly via model-fitting. In the future, we plan to use model-fitting techniques, in conjunction with calibration experiments for independent estimation of behavioral biases, to study individual and group differences in inhibitory control.

ACKNOWLEDGMENTS

We thank Jaime Ide, Chiang-shan Li, Gordon Logan, Martin Paulus, Rajesh Rao, Jeffrey Schall, Veit Stuphorn, and Nick Yeung for stimulating discussions and helpful suggestions. Funding was partly from CAN CTA (ARL/DCS).

REFERENCES

Alcock, J., and Sherman, P. (1994). The utility of the proximate-ultimate dichotomy in ethology. *Ethology* 96, 58–62.

Alderson, R., Rapport, M., and Kofler, M. (2007). Attention-deficit/hyperactivity disorder and behavioral inhibition: a meta-analytic review of the stop-signal paradigm. *J. Abnorm. Child Psychol.* 35, 745–758.

Amieva, H., Lafont, S., Auriacombe, S., Le Carret, N., Dartigues, J. F., Orgogozo, J. M., and Fabrigoule, C. (2002). Inhibitory breakdown and dementia of the Alzheimer type: a general phenomenon? *J. Clin. Exp. Neuropsychol.* 24, 503–516.

Aron, A., Behrens, T., Smith, S., Frank, M., and Poldrack, R. (2007a). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *J. Neurosci.* 27, 3743.

Aron, A., Durston, S., Eagle, D., Logan, G., Stinear, C., and Stuphorn, V. (2007b). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *J. Neurosci.* 27, 11860.

Badcock, J. C., Michie, P. T., Johnson, L., and Combrinck, J. (2002). Acts of control in schizophrenia: dissociating the components of inhibition. *Psychol. Med.* 32, 287–297.

Barkley, R. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121, 65–94.

Bellman, R. (1952). On the theory of dynamic programming. *Proc. Natl. Acad. Sci. U.S.A.* 38, 716–719.

Bogacz, R., Brown, E., Moehlis, J., Hu, P., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychol. Rev.* 113, 700–765.

Boucher, L., Palmeri, T., Logan, G., and Schall, J. (2007). Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* 114, 376–397.

Chow, Y., Robbins, H., and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Boston, MA: Houghton Mifflin.

Daunizeau, J., den Ouden, H., Pessiglione, M., Kiebel, S., Friston, K., and Stephan, K. (2010). Observing the observer (ii): deciding when to decide. *PLoS ONE* 5, e15555. doi: 10.1371/journal.pone.0015555

Emeric, E., Brown, J., Boucher, L., Carpenter, R., Hanes, D., Harris, R., Logan, G., Mashru, R., Paré, M., Pouget, P., Stuphorn, V., Taylor, T., and Schall, J. (2007). Influence of history on saccade countermanding performance in humans and macaque monkeys. *Vision Res.* 47, 35–49.

Eriksen, B., and Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.

Friedman, N., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol.* 133, 101.

Gauggel, S., Rieger, M., and Feghoff, T. (2004). Inhibition of ongoing responses in patients with Pakingson's disease. *J. Neurol. Neurosurg. Psychiatry* 75, 539–544.

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

Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.

Hanes, D., and Carpenter, R. (1999). Countermanding saccades in humans. *Vision Res.* 39, 2777–2791.

Hanes, D., Patterson, W., and Schall, J. (1998). The role of frontal eye field in countermanding saccades: visual, movement and fixation activity. *J. Neurophysiol.* 79, 817–834.

Hanes, D., and Schall, J. (1995). Countermanding saccades in macaque. *Vis. Neurosci.* 12, 929.

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

Kiani, R., Hanks, T., and Shadlen, M. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017.

Laming, D. R. J. (1968). *Information Theory of Choice Reaction Times*. London: Academic Press.

Leotti, L., and Wager, T. (2009). Motivational influences on response inhibition measures. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 430–447.

Logan, G., and Cowan, W. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychol. Rev.* 91, 295–327.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York: Oxford University Press.

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco, CA: WH Freeman.

Mazurek, M. E., Roitman, J. D., Ditterich, J., and Shadlen, M. (2003). A role for neural integrators in perceptual decision making. *Cerebr. Cortex* 13, 1257–1269.

Menzies, L., Achard, S., Chamberlain, S., Fineberg, N., Chen, C., del Campo, N., Sahakian, B., Robbins, T., and Bullmore, E. (2007). Neurocognitive endophenotypes of obsessive-compulsive disorder. *Brain* 130, 3223.

Nigg, J. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 126, 220–246.

Nigg, J., Wong, M., Martel, M., Jester, J., Puttler, L., Glass, J., Adams, K., Fitzgerald, H., and Zucker, R. (2006). Poor response inhibition as a predictor of problem drinking and illicit drug use in adolescents at risk for alcoholism and other substance use disorders. *J. Am. Acad. Child Adolesc. Psychiatr.* 45, 468.

Pare, M., and Hanes, D. (2003). Controlled movement processing: superior colliculus activity associated with countermanded saccades. *J. Neurosci.* 23, 6480–6489.

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

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychol. Rev.* 63, 129–138.

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

Stroop, J. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Verbruggen, F., and Logan, G. (2009a). Models of response inhibition in the stop-signal and stop-change paradigms. *Neurosci. Biobehav. Rev.* 33, 647–661.

Verbruggen, F., and Logan, G. (2009b). Proactive adjustments of response strategies in the stop-signal paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 835–854.

Wald, A., and Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Stat.* 19, 326–339.

Wong-Lin, K., Eckhoff, P., Holmes, P., and Cohen, J. (2009). Optimal performance in a countermanding saccade task. *Brain Res.* 1318, 178–187.

Yu, A. J., Dayan, P., and Cohen, J. D. (2009). Dynamics of attentional selection under conflict: toward a rational Bayesian account. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 700–717.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 December 2010; paper pending published: 31 January 2011; accepted: 08 May 2011; published online: 27 May 2011. Citation: Shenoy P and Yu AJ (2011) Rational decision-making in inhibitory control. Front. Hum. Neurosci. ***5****:48. doi: 10.3389/fnhum.2011.00048*

Rational impatience in perceptual decision-making: a Bayesian account of discrepancy between two-alternative forced choice and Go/NoGo Behavior

Pradeep Shenoy & Angela J. Yu

April 27, 2012

Contact Author: Angela J. Yu
Department of Cognitive Science,
University of California, San Diego
9500 Gilman Dr #0515 La Jolla, CA 92093-0515
ajyu@ucsd.edu
+1 858-822-3317
Running title: Rational impatience in Go/NoGo tasks
keywords:

Abstract

Two-alternative forced choice (2AFC) and Go/NoGo (GNG) paradigms are sometimes used interchangeably to study sensory and cognitive processing. While GNG is thought to isolate the decisional component by eliminating response selection, experimental data suggest a systematic bias toward the Go stimulus in GNG (more and faster Go responses), compared to the same two stimuli used in 2AFC (e.g., Bacon-Mace, 2007). We postulate that this "impatience" to go is due to the implicit asymmetry in the GNG cost structure. Using a Bayesian inference and risk-minimization framework, we show how minimizing both response error and delay leads to the Go bias: the NoGo response must wait until the response deadline has passed to register, while a Go response immediately terminates the trial. We show that the optimal decision policy is formally equivalent to a drift-diffusion model (DDM), with a time-varying threshold that decreases backwards toward stimulus onset, as the temporal cost of choosing NoGo increases when further away from the response deadline. Counterintuitively, previous work directly fitting a fixed-threshold DDM to behavioral data (Gomez et al., 2007) found the threshold to be higher in GNG than 2AFC. We show that fitting a fixed-threshold DDM to the optimal model produces the same effects. Our results demonstrate that the observed discrepancies between GNG and 2AFC decision-making may specifically arise from rational strategic adjustments due to the cost structure differences, and may well employ a common sensory and cognitive processing infrastructure otherwise. Two experimentally verifiable predictions are made by the optimal decision-making model with respect to the GNG task: (1) false alarm rate increases as a function of response time (the constant-threshold DDM approximation predicts a constant alarm rate), (2) lengthening the response deadline exacerbates the Go bias.

Introduction

The two-alternative forced-choice (2AFC) task is a standard experimental paradigm used in psychology and neuroscience to investigate all aspects of sensory, motor, and cognitive processing (Luce, 1991). Typically, the paradigm involves a forced choice between two responses based on a presented stimulus, with the measured response time and accuracy of choices shedding light on the cognitive and neural processes underlying behavior. Another paradigm that appears to share many features of the 2AFC task is the Go/NoGo (GNG) task ((Donders, 1969), see Luce(Luce, 1991) for a review), where one stimulus category is associated with an overt *Go* response that has to be executed before a response deadline, and the other stimulus (NoGo) requires withholding response for a certain amount of time. In principle, the GNG task could be used to probe the same decision-making problems as the 2AFC task, with the possible advantage of eliminating a "response selection stage" that may follow the decision in the 2AFC task (Donders, 1969; Gordon & Caramazza, 1982). Indeed, the GNG task has been used to study various aspects of human and animal cognition, e.g., lexical judgements (Hino & Lupker, 2000; Perea, Rosa, & Gomez, 2002), perceptual decision-making (Thorpe, Fize, Marlot, & Others, 1996; Delorme, Richard, & Fabre-Thorpe, 2000; Mack & Palmeri, 2010), and the neural basis of choice behavior (in particular, distinguishing among neural activations associated with stimulus, memory, and response)(Sommer & Wurtz, 2001; Hasegawa, Peterson, & Goldberg, 2004; Aston-Jones, Rajkowski, & Kubiak, 1994). However, experimental evidence indicates that there is a significant choice bias toward the overt (Go) response in the GNG task (Delorme et al., 2000; Bacon-Macé, Kirchner, Fabre-Thorpe, & Thorpe, 2007; Gomez, Ratcliff, & Perea, 2007; Aston-Jones et al., 1994), generating shorter response times and more false alarms in the overt response, than when compared to the same stimulus pairings in a 2AFC task (Gomez et al., 2007; Bacon-Macé et al., 2007). It has been suggested that this choice bias may reflect different sensory and cognitive processes underlying the two tasks, and thus making the two non-exchangeable in the study of perception and decision-making (see e.g., (Jones, Cho, Nystrom, Cohen, & Braver, 2002; Gibbs & Van Orden, 1998)).

We hypothesize that this discrepancy may simply arise from differences in the implicit reward (cost) structure of the two tasks: the NoGo response incurs a higher imposed waiting cost than the Go response, since the NoGo response must wait until the response deadline has passed to register, while a Go response immediately terminates the trial; in the 2AFC task, the cost function, whether in times of error or delay, is obviously symmetric across the two alternatives. We propose that the implicit cost structure difference in GNG may be able to fully account for the Go bias in GNG compared to 2AFC tasks, without the need to appeal to other differences in sensory or cognitive processing. To investigate this hypothesis, we adopt a Bayes-optimal inference and decision-making modeling framework for both the 2AFC and GNG tasks, whereby sensory processing is modeled as iterative Bayesian integration of noisy evidence, and the decision of when/how to respond rests on a decision policy that minimizes a behavioral cost function that is a linear combination of expect decision delay and response errors. This formulation has already yielded considerable insight into the behavioral and neural processes underlying perceptual decision-making in 2AFC (Gold & Shadlen, 2002; Mazurek, Roitman, Ditterich, & Shadlen, 2003; Bogacz, Brown, Moehlis, Holmes, & Cohen,

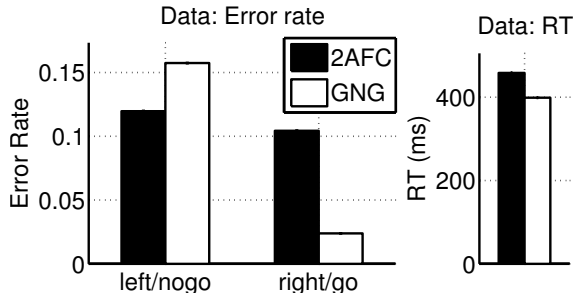

Figure 1: Systematic error biases in the GNG task. (A) The figure shows error rates associated with a perceptual decision-making task performed by subjects in both Go/NoGo and Yes/No (forced choice) settings. Although the error rates in the forced choice settings were similar for both classes, there was a significant bias towards the Go response in the GNG task, with more false alarms than omission errors. (B) Mean response time on the GNG task was lower than that on the 2AFC task. (Data adapted from Bacon-Mace et al., 2007).

2006). Here, we generalize this framework to also incorporate the GNG task, by assuming that a subject's sensory discrimination process and valuation of decision accuracy and speed are *identical* between 2AFC and GNG, with the only difference between them being the asymmetric temporal cost function in the GNG task, the Go response terminating a trial while a NoGO response is registered after the response deadline.

In the following sections, we describe our proposed Bayesian inference and decision-making model, and compare simulations of the optimal decision-making model with published experimental data of subjects performing perceptual decision-making in 2AFC and GNG tasks (Bacon-Macé et al., 2007). We also examine other evidence supporting rational decision-making in the GNG task (Nieuwenhuis, Yeung, Wildenberg, & Ridderinkhof, 2003). Next, we examine the formal relationship between the optimal model and a computationally simpler drift-diffusion model (DDM) with fixed decision threshold, which has been previously used to fit behavioral data from the GNG task (Gomez et al., 2007; Mack & Palmeri, 2010). Finally, we present experimentally verifiable predictions of the optimal decision-making model, including those that specifically differ from the fixed-threshold DDM approximation.

Bayesian inference and risk minimization in choice tasks

Figure 1 summarizes a critical difference in human choice behavior in the GNG/2AFC tasks. The figure shows data from subjects attempting to identify whether a briefly-presented noisy image contains an animal or not (Bacon-Macé et al., 2007). Subjects performed the task in two separate conditions: GNG (only respond to target images), and 2AFC (respond yes/no to each image). Even though error rates were comparable for both target and nontarget stimuli in the 2AFC task, subjects showed a significant bias towards the Go response in the GNG task, resulting in higher false alarms than omission errors (Figure 1A). Further, RT on the GNG task was significantly faster than RT on the 2AFC task (Figure 1B).

For the 2AFC task, a large body of literature supports the "accumulate-to-bound" model of

perceptual decision-making, (Ratcliff, 1978; Gold & Shadlen, 2002; Mazurek et al., 2003), where moment-to-moment sensory input ("evidence" in favor of either choice) is accumulated over time until it reaches a *bound*, at which point, a response is generated. Neural and behavioral data (Gold & Shadlen, 2007) support a specific probabilistic interpretation for this accumulated evidence, where neural activity in the lateral intraparietal area is thought to represent as the log odds in favor of one of the two choices. In such a framework, the bound for the accumulation process (Roitman & Shadlen, 2002) is equivalent to the decision-threshold as specified by the sequential probability ratio test (SPRT, (Wald, 1947)), which is known to optimize a trade-off between speed and accuracy (Wald & Wolfowitz, 1948; Bogacz et al., 2006). In particular, the optimal decision policy is a constant threshold on the posterior probability favoring either hypothesis in the case where there is no response deadline.

We use a generalization of this framework (Frazier & Yu, 2008) to account for response deadlines in both tasks; in the GNG task, the NoGo "response" is to withhold response for a specified interval, which forms an implicit deadline. We model evidence accumulation as iterative Bayesian inference over the identity of the stimulus, and decision-making as a choice procedure that minimizes overall cost given the inferred stimulus identity. We adapt this framework to jointly model behavior in both the 2AFC and GNG tasks, as detailed in the following sections.

Evidence integration as Bayesian inference

We model evidence accumulation as Bayes-optimal inference over the stimulus identity based on instant-by-instant sensory evidence. Specifically, we assume a *generative model* where the observations are a stream of sensory input $\mathbf{x}_t \triangleq (x_1, x_2, \ldots x_t)$ at time t, providing evidence of the hidden category label $d \in \{0, 1\}$. Conditioned on the label d, the instantaneous evidence x_t is generated i.i.d from one of two likelihood functions $f_0(\cdot)$ and $f_1(\cdot)$. We model the likelihod functions as Gaussian distributions with means $\pm\mu$, and a variance parameter σ^2 controlling the noisiness of the stimuli.

The *recognition model* specifies the mechanism by which stimulus identity is *inferred* from the noisy observations $\mathbf{x}_t$. In our model, we compute an optimal posterior belief over the category label conditioned on the evidence, $b_t \triangleq P\{d = 1|\mathbf{x}_t\}$, by iteratively applying Bayes' rule:

$$b_{t+1} = \frac{b_t f_1(x_{t+1})}{b_t f_1(x_{t+1}) + (1 - b_t) f_0(x_{t+1})}$$

where $b_0 \triangleq P\{d = 1\}$ is the prior probability over the category label. We hypothesize that the same evidence accumulation mechanism underlies decision-making in both tasks.

Action Selection as cost-minimization

We model behavior in the two tasks as a sequential decision-making process where, at each instant, the model needs to choose between responding (with the appropriate response), or

waiting one more time step. By waiting, more evidence can be acquired in order to better disambiguate the sensory stimulus and improve accuracy. In addition, in the GNG task, a NoGo stimulus requires a series of *wait* actions until a predefined response deadline D. We define optimal behavior in the task in terms of a *loss function* that is a linear combination of decision time and response errors. Thus, the optimal policy selects, at each time step, the action associated with the lowest *expected cost* given the instantaneous belief state.

Specifically, we assume a cost per unit time spent choosing a response (c), a penalty c_e for erroneous responses, and a penalty c_d for exceeding the deterministic response deadline D. Here we set $c_e = c_d = 1$ for simplicity. The *optimal Bayes-risk minimization policy* for the 2AFC task is the policy that minimizes the expected loss over all possible stimulus sequences:

$$L_\pi = c\langle\tau\rangle + P\{\delta \neq d\} + P\{\tau = D\}$$

Here $\langle\tau\rangle$ is the average of the response times τ for the model. If no response is registered by the time of the deadline, τ is set to D. This formulation is identical to the deadlined 2AFC response model we proposed in our earlier work (Frazier & Yu, 2008).

We define a loss function for the GNG task in an analogous manner. For the GNG task, choosing the overt (Go) response terminates the trial, but incurs a penalty if it was a NoGo trial. On the other hand, NoGo trials necessarily require the subject to continue to wait until the response deadline. The loss function for the GNG task is therefore:

$$L_\pi = c\langle\tau\rangle + P\{d = 0, \tau < D\} + P\{d = 1, \tau = D\}$$

The principal difference between the two tasks as formulated here is the loss function. In the 2AFC task, a trial is terminated by a response irrespective of the category of the stimulus. However, in the GNG version, subjects have to wait until the response deadline on trials that are associated with the "NoGo" response. This introduces a significant extra cost of time for NoGo responses, suggesting that it may in some cases be better to select the Go response despite the relative paucity of sensory evidence. We explore these aspects in detail in the following section.

Results

Opportunity cost and the Go/NoGo decision threshold

Figure 2A compares the optimal decision-making threshold for the two tasks, as a function of time. This is the threshold for the belief state b_t at which a response is generated. For the 2AFC task, the optimal policy is a pair of thresholds that are relatively constant over time, but collapse to a point as the deadline approaches, in order to avoid missing the response deadline (cf. (Frazier & Yu, 2008)). In contrast, the threshold for the GNG task (dotted line) is a single threshold that varies over time, and is *lower* at the beginning of the trial. This is a direct consequence of the opportunity cost involved with waiting until the deadline: if the deadline is far away, the cost of waiting may be more than the cost of an immediate

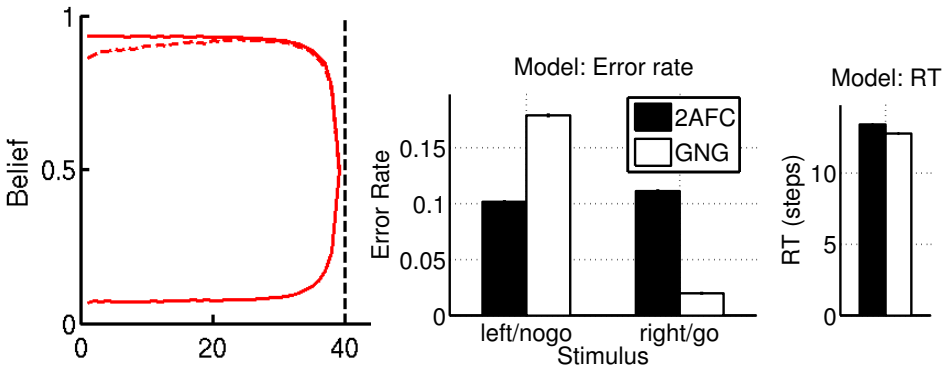

Figure 2: Rational behavior in 2AFC and GNG tasks. (A) The figure shows the decision threshold as a function of belief state across the 2AFC and GNG tasks. The optimal decision boundary for 2AFC is a pair of parallel thresholds that decrease and meet at the response deadline) (indicated by dashed vertical line). The Go/NoGo decision boundary is a single initially *increasing* threshold (dashed line), that decreases to 0.5 at the response deadline. (B) Model simulations show a bias towards the overt response in the GNG task, leading to reduced errors to the Go stimulus, and higher error rate on NoGo stimuil.

error that terminates the trial. This varying decision threshold is conceptually similar to the decreasing decision threshold that arises as a consequence of the *cost of exceeding the deadline* seen in the figure above (Frazier & Yu, 2008).

Decision-making in 2AFC and GNG tasks

Figure 2B;C shows the effect of the time-varying threshold on RT and accuracy in an example model simulation. Figure 2B shows that the GNG model is significantly biased towards the Go response, with a higher fraction of false alarms than misses. This asymmetry is absent in the 2AFC model performance. In addition, GNG response times are *faster* than 2AFC response times (Figure 2C). This bias is a direct result of the time-varying threshold in the GNG task; early on in the trial, the decision threshold is lower, and produces fast, error-prone responses.

This model prediction is confirmed by data from human perceptual decision-making. Figure 1 shows behavioral data in the two tasks (Bacon-Macé et al., 2007)– subjects determined from a brief presentation of a noisy visual stimulus whether or not the image contained an animal. The same task was performed in two response conditions: 2AFC, where each stimulus required a yes/no response, and GNG, where subjects only responded to image containing the target. Figure 1A shows that in the 2AFC condition, subjects are not significantly biased towards either response, with both false alarms and miss rates being similar to each other. On the other hand, in the Go/NoGo condition, subjects showed a significant bias towards the overt response, thus producing substantially more false alarms and fewer misses. In the GNG task, their RT was significantly shorter than in the 2AFC task (Figure 1B). Similar results are reported by Gomez et al. in the context of lexical decision-making (Gomez et al., 2007) – although our model is not directly applicable to the lexical decision-making tasks

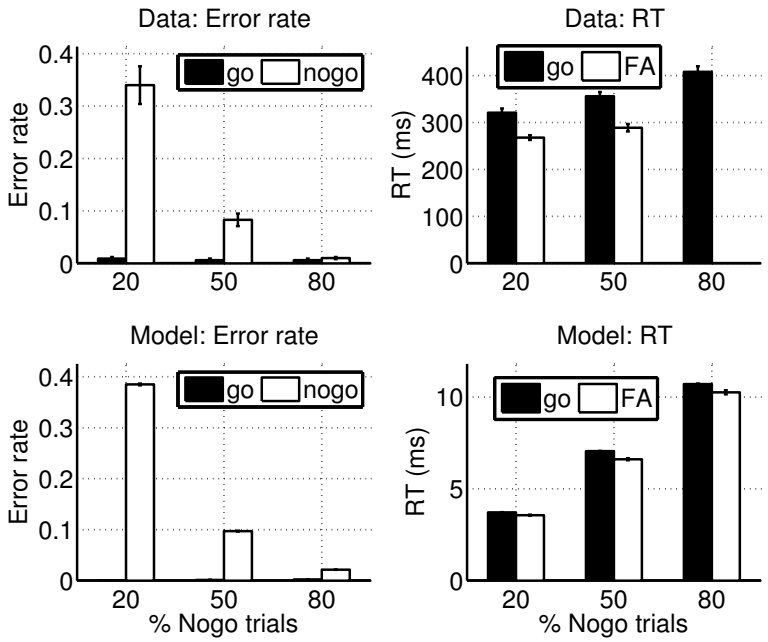

Figure 3: Influence of stimulus statistics on Go bias. Our model predicts that alse alarms are more frequent than misses (A), and are also faster than correct Go RTs (B). The Go bias, which is apparent at 50% Go trials, is signficantly increased when Go trials are more frequent (80%), and reduced when Go trials are reduced to 20% of the trials. (C-D) Human subjects showed a trend in their choice behavior at a letter discrimination task (Data from Nieuwenhuis et al., 2003).

they study, the asymmetric cost function in the GNG task nevertheless produces a trend towards early, error-prone responses.

Influence of stimulus probability on Go bias

To further investigate the behavior of the GNG model, and the degree of Go bias, we examined data from experiments where the overall frequeny of trial types was varied. For instance, Nieuwenhuis et al. (Nieuwenhuis et al., 2003) examined the effect of trial type frequency on behavioral and and electrophysiological measures in the GNG task. They used a block design to compare choice accuracy and RT when the fraction of NoGo trials was set to 20%, 50%, and 80%. As shown in Figure 3A;B, subjects' behavior was reliably modulated by trial type frequency– as NoGo trials became more prevalent, the rate of false alarms dropped substantially. The choice data show that, consistent with Figure 2 and a host of other experimental data, there is a significant bias towards the Go response when Go and NoGo trials are equiprobable, and this bias is increased (respectively diminished) as NoGo trials are fewer or more frequent (see also (Low & Miller, 1999)). The figure also shows that RT for both correct Go and erroneous NoGo responses increase with the frequency of NoGo trials.

Optimal model simulations shown in Figure 3C;D capture the choice & RT trends seen in the

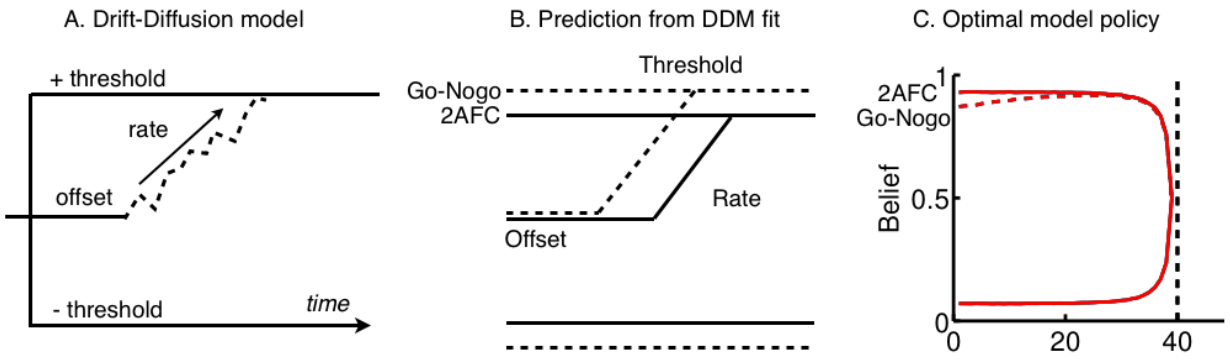

Figure 4: Drift-diffusion model (DDM) for 2AFC and GNG tasks (A) A simplified version of the DDM for 2-choice tasks, where a noisy accumulation process with a certain *rate* produces one of two responses when it reaches a positive or negative *threshold*. In addition to the rate and threshold parameters, a third parameter (the temporal *offset* to the start of the accumulation process) represents the nondecision processes associated with visual and motor delays. (B) DDM fits to 2AFC and GNG choice data(Gomez et al., 2007, Mack & Palmeri, 2010) suggest that the GNG task is associated with a higher threshold and shorter offset than the 2AFC task. (C) Optimal decision-making model predicts a lower, time-varying threshold for the GNG task.

data. In our formulation, although the decision boundary is unchanged by the experimental manipulation, the stimulus frequency induces a *prior belief* over the identity of the stimulus, and thus represents the starting point for the evidence accumulation process. When Go trials are rare, the starting point is far from the decision boundary, and. it takes longer for a response to be generated. Further, due to the extra evidence needed to overcome the prior, choices are less likely to be erroneous.

Drift-diffusion models and optimal behavior

By examining RT and choice data from lexical judgment, numerosity judgment, and memory-based decision making tasks, Gomez et al. (Gomez et al., 2007) found that a DDM with an *implicit negative boundary* associated with the NoGo stimulus provided a good fit to RT data. Further, joint parameter fits to 2AFC and GNG choice data indicated that the principal difference in the two tasks was in the nondecision time and decision threshold; the rate parameter (representing the evidence accumulation process) was similar in both tasks. In particular, they suggested that the nondecision time was shorter, and the decision threshold higher than in the 2AFC task (Figure 4B). These results were replicated by Mack & Palmeri by fitting DDMs to behavioral data from a visual categorization task performed in both 2AFC and GNG versions (Mack & Palmeri, 2010).

Although DDMs are formally equivalent to optimal decision-making in a restricted class of sequential choice problems (Wald & Wolfowitz, 1948), they do not explicitly represent and manipulate uncertainty and cost, as we do in our Bayesian risk-minimization framework. In particular, our framework allows us to *predict* that optimal behavior is well-characterized by

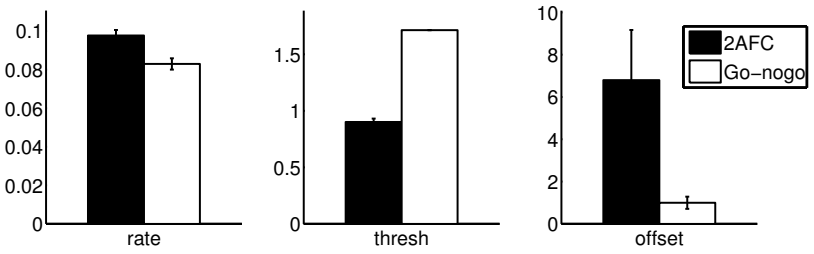

Figure 5: DDM approximation to optimal decision-making model. Simplified DDMs were fit to optimal model simulations of 2AFC and GNG behavior, and the best-fit parameters compared between tasks. The DDM approximation for optimal GNG behavior shows a higher decision threshold (B), and lower nondecision time (C), than the DDM approximation for the 2AFC task. In addition, the rate of evidence accumulation was also lower for the GNG fit (A).

a DDM with a time-varying threshold (Figure 4C), and that the restricted class of constant-threshold DDMs are insufficient to fully explain observed behavior. Nevertheless, we can ask whether our prediction is *consistent* with the empirical results obtained from DDM fits with constant decision thresholds.

To answer this question,

we computed the best constant-threshold DDM approximations to optimal decision making in the two tasks. We simulated the optimal model with a fixed set of parameters for both the 2AFC and GNG tasks, and then fit simplified random-walk models with 3 free parameters (Figure 4A) to the output of our optimal model's simulations. Figure5 shows that the best-fitting DDM approximation for optimal GNG behavior has a higher threshold and a lower offset parameter than the best-fitting DDM for optimal 2-choice behavior.

Thus, our results and those of Gomez et al. (Gomez et al., 2007) are conceptually consistent; a prinicipal difference in the two tasks is the decision threshold, whereas the evidence accumulation process is similar across tasks. However, our analysis explains precisely *how* and *why* the thresholds in the two tasks are different: the GNG task has a time-varying threshold that is lower than the 2-choice threshold, due to the difference in loss functions in the two tasks. When optimal behavior is approximated by a simpler class of models (i.e., models with fixed decision threshold), the best fit to optimal GNG behavior turns out to be a higher threshold and shorter nondecision time, as found by previous work (Gomez et al., 2007; Mack & Palmeri, 2010).

Model Predictions

The Bayes-risk-minimization framework we propose readily lends itself to novel predictions of choice behavior that can be verified experimentally. We present below some specific predictions arising from our model of the GNG task. In particular, these predictions will help distinguish between our proposed model and constant-threshold models such as the DDM for behavior in the GNG task.

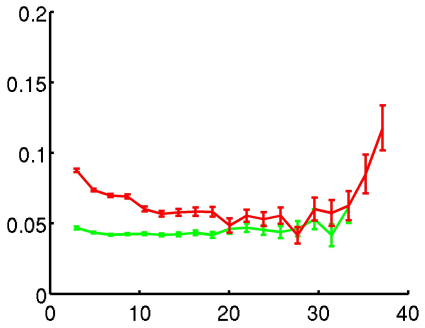

Figure 6: Relationship between RT and error rate. The figure shows average error rate of model simulations on trials binned according to the associated response times, separately for GNG and 2AFC. Model predicts higher error rates for early responses in the GNG task.

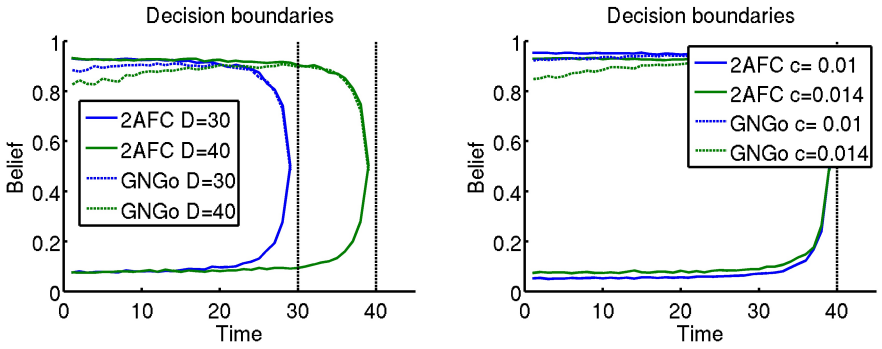

Figure 7: Influence of response deadline and cost of time on Go bias. The figures show the decision threshold as a function of belief state across the 2AFC and GNG tasks. (A) As the response deadline is increased, the decision threshold for GNG is lower at trial onset. (B) As the cost of time is increased, the relative cost of time and error is changed, and the difference between GNG and 2AFC thresholds is magnified.

A central prediction of our GNG model is the time-varying decision threshold, which is a consequence of the opportunity cost of waiting on NoGo trials. Since the decision threshold at any point in time is exactly equal to the accuracy of responses generated with that latency, the error rate in the GNG task is a function of response time, with a higher error rate for early responses. Figure 6 compares predicted average error rate for trials split into bins according to RT, for both the GNG and 2AFC tasks. For the 2AFC task, error rates are near-constant up until the response deadline where it increases steeply. In contrast, the GNG error rate shows a U-shaped relationship with RT, where both early and late respones are more error-prone. This nontrivial relationship between RT and choice accuracy is directly in contrast to the RT-independent error rate predicted by constant-threshold models (Gomez et al., 2007), and constitutes a testable prediction.

A second testable prediction of our model is that the decision threshold, and the associated Go bias, is modulated by any factor that affect the opportunity cost on NoGo trials. For example, Figure 7 shows predicted decision thresholds when extending the response deadline,

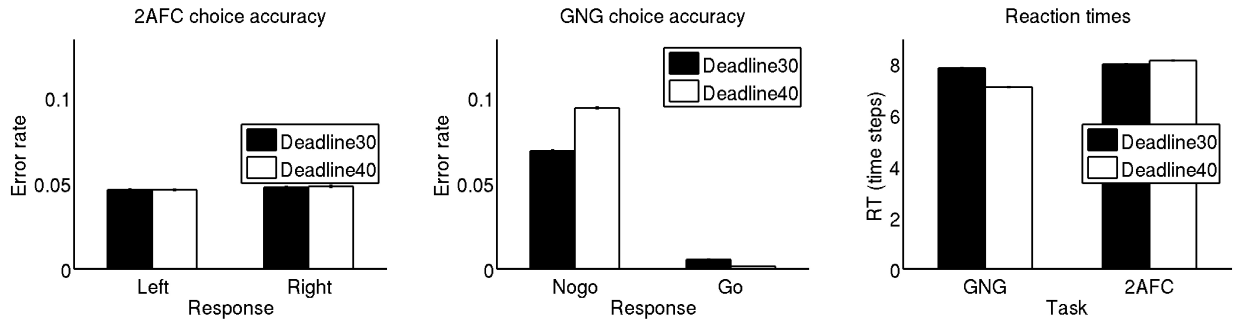

Figure 8: Influence of deadline on error rate and RT. (A) 2AFC error rates are not significantly affected by increasing the response deadline, for the parameter settings used in this simuluation. (B) Bias towards the Go response in the GNG task is significantly exacerbated by an increased response deadline. (C) Reaction times show a dissociation between the two tasks – RT on correct trials decreases for the GNG task, and increases for the 2AFC task.

and when increasing the relative cost of time. As expected, both of these manipulations produce a steep change in the decision boundary for the GNG task (Figure 7A-B). In contrast, the initial decision threshold for the 2AFC task does not change when the response deadline is extended, and is only minimally lowered when the cost of time is increased. These changes are directly reflected in the RT and error rates in the two tasks. For example, with an extended response deadline, 2AFC accuracy is not significantly affected, and RT shows a slight increase (Figure 8A; C). In contrast, GNG accuracy is asymmetrically affected, with a significant increase in the Go bias when the deadline is extended (Figure 8B). Again, in contrast to the 2AFC behavior, RT on correct trials *decreases* as a result of the extended deadline (Figure 8C). Similar results are predicted when the cost of time is increased (not shown).

Discussion

Forcing a choice between two alternatives is a fundamental technique used to study a wide variety of perceptual and cognitive phenomena. In many situations, the GNG variant of such tasks may appear to have advantages, e.g., simplicity of response selection for experiments in animals, or the removal of any potential response selection processes that may exist separately from the decision process. However, careful behavioral contrasts (Gomez et al., 2007; Bacon-Macé et al., 2007) and the theoretical analysis presented here show that, although the underlying decision-making processes are common across the two tasks, systematic biases are introduced by the use of the Go/NoGo paradigm.

When the optimal confidence threshold for this loss function is computed, our model makes the non-intuitive prediction that the confidence criterion for generating a response *increases over time* in the GNG task, favoring more error-prone, early responses. This time-varying threshold explains the observed bias towards overt responses, as well as the shorter response times in the GNG task. and find that the best-approximating GNG threshold is in fact higher than the threshold used for the 2AFC task, with a corresponding lower nondecision

time.

Taken together, our results suggest that systematic differences in RT and error rates between 2AFC and Go/NoGo tasks need not imply different underlying neural or cognitive processes. In fact, common Bayes-optimal sensory and decision processes in the two tasks produce asymmetries as a consequence of inherent reward structure asymmetries in the two tasks. Nevertheless, the 2AFC task is to be preferred when the objective is to study the cognitive process that distinguishes between two stimuli. If the GNG task is used instead, inferences made about the underlying cognitive phenomenon, or its neural correlates thereof, must be carefully scrutinized to account for these introduced biases.

Donders (Donders, 1969) originally suggested that various putative stages of mental processing could be measured by subtracting out pairs of reaction times on simple, choice and Go/Nogo tasks. For example, one might assume the existence of a *response selection stage* subsequent to a decision stage, and estimate its duration by subtracting GNG RT from 2AFC RT. However, this subtraction scheme depends on multiple problematic assumptions, including modular, isolated stages of processing, and the assumption of *pure insertion* (e.g., (Sternberg, 1969)), i.e., that processing stages are identical in duration across tasks. Using Donders' subtraction method would imply that the observed RT difference between the two tasks is the "response selection time", assuming motor execution times are similar (Miller & Low, 2001). We do not explicitly model a response selection stage or process, although such processes may indeed follow or accompany the decision-making process we model. Regardless, our analysis suggests that a principal source of the RT difference in the two tasks is in fact the *induced bias* towards early, error-prone responses in the GNG task.

Our model for the GNG task has a single decision threshold associated with the overt response, consistent with some early models proposed for the task (Smith, 2000; Sperling & DOSHER, 1986; Grainger & Jacobs, 1996). In contrast, the extended DDM framework proposed by Gomez et al. has an additional boundary associated with the NoGo response (corresponding to a *covert* NoGo response). Gomez et al. report that single-threshold variants of the DDM provided very poor fits to the data. Although computationally and behaviorally we do not *require* a covert-response or associated threshold, it is nevertheless possible that neural implementations of behavior in the task may involve an explicit "NoGo" response. Indeed, performance at the GNG task is also used as a measure of of *behavioral inhibition* (see e.g., (Eagle, Bari, & Robbins, 2008) for a summary). It is conjectured that behavior in the task recruits an additional executive control process that actively inhibits responses when inappropriate. We propose that differences in the responses to Go and NoGo stimuli may arise as a rational consequence of the asymmetric costs associated with the responses, and that inter-individual differences in the task may therefore be driven by differences in subjective valuations of time and accuracy. We have shown, in a closely related behavioral inhibition task called the stop signal task), that behavior in the task and its variants are be the outcome of rational decision-making based on cognitive factors such as perceived rewards and stimulus statistics.(Shenoy & Yu, 2011) Interestingly, the Go/NoGo and stop signal tasks show a very different neurophysiological and pharmacological response profile, although they purportedly probe related cognitive processes (Swick, Ashley, & Turken, 2011; Eagle et al., 2008); in other words, the brain activations and circuitry associated with successfully withholding a response share similarities and differences in the two tasks (Swick

et al., 2011). Further, the effect of pharmacological interventions are also different in the two tasks (Eagle et al., 2008), leading some researchers to suggest that there may be two different kinds of "behavioral inhibition" at work in the two tasks. A promising avenue for future work is a computational "theory of behavioral inhibition", that seeks to jointly explain the behavioral and neurophysiological similarities and differences observed using these two tasks.

References

Aston-Jones, G., Rajkowski, J., & Kubiak, P. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *The Journal of Neuroscience*, *14*, 4467–4480. Available from http://www.jneurosci.org/content/14/7/4467.short

Bacon-Macé, N., Kirchner, H., Fabre-Thorpe, M., & Thorpe, S. (2007). Effects of task requirements on rapid natural scene processing: From common sensory encoding to distinct decisional mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1013. Available from http://psycnet.apa.org/journals/xhp/33/5/1013/

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700. Available from http://psycnet.apa.org/psycinfo/2006-12689-002

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorization of natural scenes does not rely on colour cues: a study in monkeys and humans. *Vision Research*, *40*(16), 2187–2200. Available from http://www.sciencedirect.com/science/article/pii/s0042-6989(00)00083-3

Donders, F. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412. Available from http://ukpmc.ac.uk/abstract/MED/5811531

Eagle, D., Bari, A., & Robbins, T. (2008). The neuropsychopharmacology of action inhibition: cross-species translation of the stop-signal and go/no-go tasks. *Psychopharmacology*, *199*(3), 439–456. Available from http://www.springerlink.com/index/42410P80815K6X74.pdf

Frazier, P., & Yu, A. (2008). Sequential hypothesis testing under stochastic deadlines. *Advances in neural information processing systems*, *20*, 465–472. Available from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.71.8705&rep=rep1&t

Gibbs, P., & Van Orden, G. (1998). Pathway selection's utility for control of word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(4), 1162. Available from http://psycnet.apa.org/psycinfo/1998-04674-008

Gold, J., & Shadlen, M. (2002). Banburismus and the Brain:: Decoding the Relationship between Sensory Stimuli, Decisions, and Reward. *Neuron*, *36*(2), 299–308. Available from http://linkinghub.elsevier.com/retrieve/pii/S0896627302009716

Gold, J., & Shadlen, M. (2007). The neural basis of decision

making. *Annu. Rev. Neurosci.*, *30*, 535–574. Available from http://www.annualreviews.org/doi/abs/10.1146/annurev.neuro.29.051605.113038

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go-nogo task. *Journal of Experimental Psychology*, *136*(3), 389–413.

Gordon, B., & Caramazza, A. (1982). Lexical decision for open-and closed-class words: Failure to replicate differential frequency sensitivity. *Brain and Language*, *15*(1), 143–160. Available from http://linkinghub.elsevier.com/retrieve/pii/0093934X82900530

Grainger, J., & Jacobs, A. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological review*, *103*(3), 518. Available from http://psycnet.apa.org/psycinfo/1996-01780-005

Hasegawa, R. P., Peterson, B. W., & Goldberg, M. E. (2004, August). Prefrontal neurons coding suppression of specific saccades. *Neuron*, *43*(3), 415–25. Available from http://www.ncbi.nlm.nih.gov/pubmed/15294148

Hino, Y., & Lupker, S. (2000). The effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of experimental psychology. Human perception and performance*, *26*, 166–183.

Jones, A., Cho, R., Nystrom, L., Cohen, J., & Braver, T. (2002). A computational model of anterior cingulate function in speeded response tasks: Effects of frequency, sequence and conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(4), 300–317. Available from http://www.springerlink.com/index/Y350418548791606.pdf

Low, K., & Miller, J. (1999). The usefulness of partial information: Effects of go probability in the choice/nogo task. *Psychophysiology*, *36*(3), 288–297. Available from http://onlinelibrary.wiley.com/doi/10.1017/S0048577299980332/abstract

Luce, R. (1991). *Response times: Their role in inferring elementary mental organization* (No. 8). Oxford University Press, USA. Available from http://books.google.com/books?hl=en&lr=&id=WSmpNN5WCw0C&oi=fnd&pg=P

Mack, M. L., & Palmeri, T. J. (2010). Modeling categorization of scenes containing consistent versus inconsistent objects. *Journal of Vision*, *10*, 1–11.

Mazurek, M., Roitman, J., Ditterich, J., & Shadlen, M. (2003). A role for neural integrators in perceptual decision making. *Cerebral cortex*, *13*(11), 1257. Available from http://cercor.oxfordjournals.org/content/13/11/1257.short

Miller, J., & Low, K. (2001). Motor processes in simple, go/no-go and choice reaction time tasks: a psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(2), 266. Available from http://psycnet.apa.org/journals/xhp/27/2/266/

Nieuwenhuis, S., Yeung, N., Wildenberg, W. van den, & Ridderinkhof, K. R. (2003, March). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, affective & behavioral neuroscience*, *3*(1), 17–26. Available from http://www.ncbi.nlm.nih.gov/pubmed/12822595

Perea, M., Rosa, E., & Gomez, C. (2002). Is the go/nogo lexical decision task an alternative to the yes/no lexical decision task? *Memory and Cognition*, *30*(34-45).

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59. Available from http://psycnet.apa.org/journals/rev/85/2/59/

Roitman, J., & Shadlen, M. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *The Journal of neuroscience*, *22*(21), 9475. Available from http://neuro.cjb.net/content/22/21/9475.short

Shenoy, P., & Yu, A. J. (2011). Rational decision-making in inhibitory control. *Frontiers in human neuroscience*, *5*(48).

Smith, P. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*, *44*(3), 408–463. Available from http://www.sciencedirect.com/science/article/pii/S0022249699912609

Sommer, M., & Wurtz, R. (2001). Frontal eye field sends delay activity related to movement, memory, and vision to the superior colliculus. *Journal of Neurophysiology*, *85*(4), 1673–1685. Available from http://neuro.cjb.net/content/18/18/7519.short http://jn.physiology.org/content/85/4/1673.short

Sperling, G., & DOSHER, B. (1986). Strategy and optimization in human information processing. *Handbook of perception and human performance.*, *1*, 2–1. Available from http://www.csa.com/partners/viewrecord.php?requester=gs&collection=TRD&reci

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta psychologica*, *30*, 276–315. Available from http://www.sciencedirect.com/science/article/pii/0001691869900559

Swick, D., Ashley, V., & Turken, U. (2011). Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. *NeuroImage*. Available from http://linkinghub.elsevier.com/retrieve/pii/S1053811911002473

Thorpe, S., Fize, D., Marlot, C., & Others. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. Available from http://fias.uni-frankfurt.de/ triesch/courses/260object/papers/SpeedOfProcessing.pd

Wald, A. (1947). *Sequential analysis*. Dover publications.

Wald, A., & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, *19*(3), 326–339. Available from http://www.jstor.org/stable/2235638

Stimulus Expectancy, Norepinephrine, and Inhibitory Control

Pradeep Shenoy Angela J. Yu

April 23, 2012

Contact Author:
Angela J. Yu
Department of Cognitive Science,
University of California, San Diego
9500 Gilman Dr #0515
La Jolla, CA 92093-0515
ajyu@ucsd.edu
+1 858-822-3317

keywords: inhibitory control, norepinephrine, atomoxetine, stimulus expectancy, sequential effects, Bayesian inference, decision-making

Abstract

Inhibitory control, a key component of executive function, is thought to be dysregulated in conditions such as attention-deficit hyperactivity disorder and substance abuse. Recent studies show that inhibitory control is modulated by various contextual factors such as stimulus frequency and trial history, as well as atomoxetine (a norepinephrine reuptake inhibitor). We propose that these effects reflect strategic adjustment to fluctuations in stimulus expectancy. Specifically, we propose a rational decision-making model for the stop signal paradigm, and show that dynamic adjustments in subjects' internal expectancy of stop trials drive trial-by-trial variability in reaction time and accuracy in the task. Consistent with our previous proposal that norepinephrine helps to detect unexpected events, we hypothesize that it signals stop signal expectancy in the context of the stop signal task. The model predicts that norepinephrine depletion/elevation should increase/decrease stop errors and stopping latency, respectively. These predictions match with observed inhibitory impairments in ADHD patients and improvement upon atomoxetine administration. These results delineate the modulatory role of stimulus expectancy in inhibitory control, as well as the functional basis of norepinephrine involvement in inhibitory control and dysfunction.

Author Summary

Whether biting our tongue to stop an inappropriate comment, or resisting the temptation of an extra piece of dessert, the healthy brain constantly employs inhibitory control in everyday life to negotiate competing behavioral needs and objectives. However, inhibitory capacity, as measured in experimental paradigms such as the stop signal task, varies significantly across individuals, and even within an individual in different settings. In this work, we use a Bayesian inference and decision-making framework to identify the underlying factors that lead to across-subject and within-subject variability in inhibitory control. Our results suggest that, in general, subjects maintain and continuously update expectancy about upcoming stimuli (and the requisite responses), and that much of the across- and within-subject variability can be traced to differences in this expectancy updating process. Furthermore, we hypothesize that the neuromodulator norepinephrine plays a critical role in representing stimulus expectancy in the brain. This explains the impaired inhibitory capacity of individuals with Attention Deficit Hyperactivity Disorder, which is associated with norepinephrine depletion. Conversely, it also explains why atomoxetine, an ADHD drug that elevates effective norepinephrine level, ameliorates such inhibitory deficits, not only in ADHD patients, but also healthy humans and laboratory rats.

Introduction

Inhibitory control, the ability to withhold or modify actions in response to dynamically changing task demands, is a key element of executive control. Whether biting our tongue to stop an inappropriate comment, or resisting the temptation of an extra piece of dessert, the healthy brain constantly employs inhibitory control in everyday life to negotiate competing behavioral objectives. Deficits in inhibitory control appear to underlie a number of psychiatric conditions, including attention-deficit hyperactivity disorder (ADHD) [1, 2], obsessive-compulsive disorder [3], and substance abuse [4]. Understanding the cognitive and neurophysiological processes underlying inhibitory control is important for both basic and clinical neuroscience.

While traditionally inhibitory control was often thought of comprising rather automatic processes, a confluence of recent studies indicates that it is under the control of various contextual and cognitive factors. For example, in the classic stop signal task [5], in which a rare *stop* signal cues the withholding of a prepotent *go* response, an individual's stopping latency systematically varies according to stop signal frequency [6, 7, 8], trial history [6], and the administration of atomoxetine [9, 10], a norepinephrine-reuptake inhibitor [11, 12].

Previously, we proposed that subjects decide whether to go or stop in the stop-signal task by acting as rational decision-makers – continually weighing the relative values of *stop* and *go* on a millisecond timescale based on an accumulating stream of noisy sensory evidence [13]. We also showed that reward/motivational effects in the stop-signal task [14] arise naturally when subjects correctly internalize the experimentally imposed cost of stop and go errors [13].

In this work, we utilize the same rational decision-making framework to examine how other cognitive factors, e.g. stimulus expectancy and experienced stimulus history, affect stopping behavior. This normative approach makes it possible to infer subjects' prior knowledge and behavioral goals from observed behavioral patterns. We collect data from human subjects performing a version of the stop-signal task, and infer how they adjust their stimulus expectancy and decision policy on a trial-to-trial basis based on recent trial history. Building on our previous proposal that norepinephrine signals unexpected uncertainty [15, 16], we hypothesize that the impaired stopping ability in ADHD patients may arise from a deficient anticipation of the unexpected stop signal, and atomoxetine improves stopping behavior via increased subjective expectancy of the stop signal.

Results

The stop signal task typically involves discriminating a go stimulus that requires one of two responses on each trial, followed by a *stop signal*, on a small fraction of trials, that instructs subjects to withhold the go response (Figure 1A). On trials where a stop signal is presented (stop trials), a go response is considered to be a *stop error*. Stop errors become increasingly frequent as the onset asynchrony between the go and stop stimuli, the *stop signal delay* (SSD), is increased [5, 17, 18]. Subjects are naturally discouraged from going too early because they are more likely to make stop errors; they are also experimentally discouraged

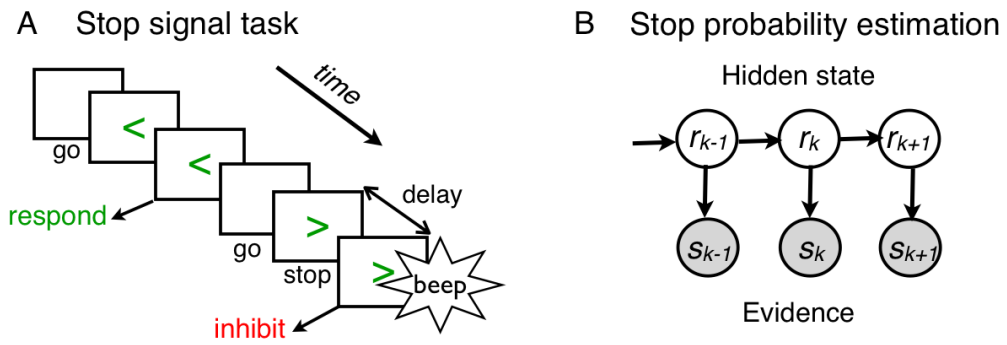

Figure 1: Task and Bayesian model schematics. (A) In a typical stop signal task, subjects discriminate between two *go* stimuli (left and right arrows) on most trials, but have to withhold their response on a small number of stop trials at the cue of a subsequently presented *stop signal*. (B) Subjects' trial-by-trial learning of stop trial frequency is modeled using the illustrated generative process. The probability r_k of a stop signal appearing on trial k is assumed to be changeable in an unsignaled manner, and consequently evolves across trials. Subjects receive veridical feedback on each trial indicating trial type (stop trial: $s_k = 1$, go trial: $s_k = 0$), which can in turn be used perform Bayesian inference of the underlying stop trial frequency r_k.

to go too slow either via explicit feedback or some form of a response deadline on go trials. Our rational decision-making model [13] provides a normative framework for negotiating the various competing behavioral objectives in the task, thus explaining not just *how* subjects respond to different contextual variables such as stop signal frequency [6, 7, 8], trial history [6], and reward/motivational factors [14], but *why* they do so, as well as what the observed behavioral patterns imply about the underlying cognitive and neural processes.

More concretely, this rational decision-making model [13] hypothesizes that subjects maintain a continually evolving, Bayes-optimal representation of stimulus category based on a noisy sensory input stream, and make an instant-by-instant choice of whether to go or wait one more time point, so as to minimize an average or expect cost function that takes into account response delay and stop/go errors (see [13] and Methods for details). We have shown that this rational decision-making model naturally gives rise to basic behavioral patterns consistently observed in the stop-signal task, as well as strategic adjustments due to reward/motivational manipulations [13].

In this work, we investigate the role of stimulus expectancy and the noradrenergic neuromodulatory system in inhibitory control. We expand the rational decision-making model to include trial-by-trial processing, as well as presenting novel behavioral data, to explain the contextual influences of stop signal frequency [6, 7, 8] and trial history [6] on inhibitory control. We also connect these contextual factors to the role of norepinephrine in inhibitory control, and, in particular, to inhibitory deficits in ADHD.

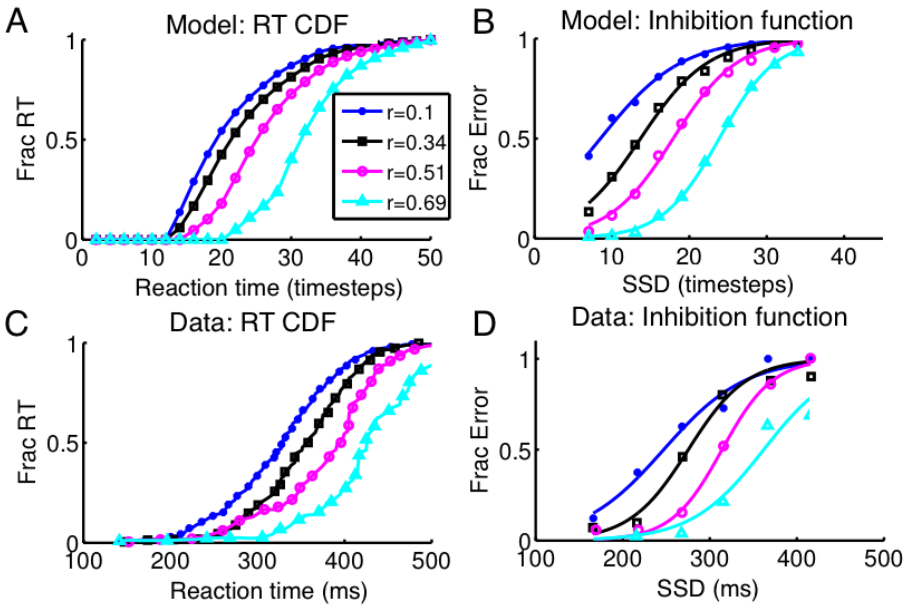

Figure 2: Influence of stop signal frequency on stopping behavior. Top panels: model predicts that as stop signals become more frequent (increasing values of r, the stop trial fraction), the higher prior expectation of a stop signal increases go RT while decreasing stop errors. Bottom panels: similar patterns observed in a monkey performing a saccade version of the task, adapted from Emeric et al (2007).

Influence of stop signal frequency on stopping behavior

Manipulation of the base rate of stop trials has consistently shown that increasing stop trial frequency results in an increase in go RT and decrease in stop error rate, whether the manipulation is implicit through experienced trials [6, 14] or explicit through instruction or a cue [8]. Within the context of our optimal decision-making model of inhibitory control, this can be understood as a rational consequence of a higher expectancy (prior probability) of encountering a stop signal: the ideal observer more readily detects a stop signal on a stop trial, and requires more sensory data to override its prior bias toward a stop signal on a go trial. Consequently, when we simulated the rational decision-making model, as stop trial frequency is increased, the rational model's go RT increases (Figure 2A), and stop error rate decreases (shifts to the right as a function of SSD, Figure 2B). Figure 2C;D shows that a macaque monkey behaves as predicted by the model in a saccade version of the stop signal task (data from Emeric et al. [6]). Human subjects also show similar behavioral adjustments as a function of stop trial frequency [14, 8].

Influence of trial history on stopping behavior

While the previous section established that prior expectation about stop signal appearance can modulate stopping behavior, here we examine how such prior expectations can arise based on trial-by-trial experience of stimulus type in the stop signal task. Go RT consistently

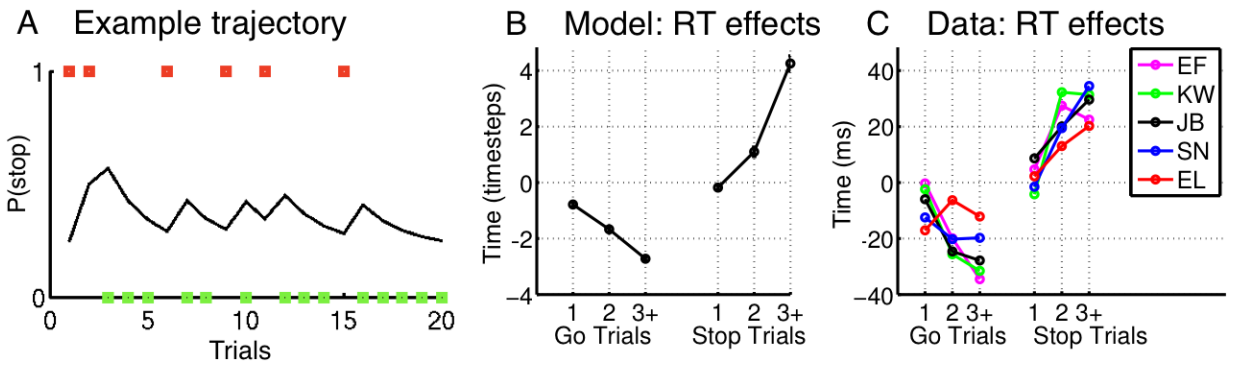

Figure 3: Influence of trial history on go RT. (A) The Bayesian sequential prediction model produces an *iterative* prior probability of the next trial being a stop trial (denoted as $P(stop)$, black line), which rises after a run of stop trials (red squares) and drops after a run of go trials (green squares). (B) Bayes-optimal inference and decision model predicts go RT to increase after a run of go trials, and decrease after a run of stop trials. (C) Similar sequential effects are observed in human subjects. RT changes are shown relative to subject-specific mean go RT; negative value indicates RT faster than a subject's average. Data ($n = 5$) adapted from Emeric et al (2007). The set of stop trial fraction values used in model and experiment were identical.

speeds up after a run of go trials and slows down after a run of stop trials (Figure 3C), even when these runs arise by chance [6, 7]. We interpret these sequential effects as reflecting a dynamic adjustment of internal expectations based on recent trial history: a run of stop trials elevates perceived stop trial frequency, while a run of go trials lowers it.

To examine sequential adjustments quantitatively, we adapt our previous Dynamic Belief Model (DBM; see Figure 1B; based on Yu & Dayan [19]) to model sequential prediction of stop trial likelihood as Bayesian inversion of a hidden Markov model that assumes trial type (stop or go) to be stochastically generated according to an underlying Bernoulli rate parameter. The rate parameter undergoes Markovian non-stationarity, remaining the same from trial to trial with probability α, and changing to a new value drawn from a constant prior distribution with probability $1 - \alpha$ (see Methods for details). We then insert the expected value of stop signal frequency, which we denote as $P(stop)$, into the rational decision-making model [13], in place of the actual frequency r as in the last section, to generate predictions for behavioral measures such as RT and stop accuracy. Such a dynamic adjustment policy of stimulus expectation can enable the subject to learn stimulus statistics and improve performance when stimulus frequencies can change over time (as in the base rate manipulation experiments of last section). However, it can also leads subjects to "superstitiously" latch onto chance patterns when the underlying statistics are truly stationary (as in Figure 3C). Previously, we presented theoretical arguments for why subjects persist in these superstitious tendencies in stationary context [19].

Figure 3A shows an example of this estimation process applied to a sample sequence of simulated data: the expected likelihood of seeing a stop trial, $P(stop)$, increases after each observation of a stop trial, and decreases after a go trial. Using a Monte Carlo simulation of the DBM, in combination with the rational decision-making model [13], we predict that go RT

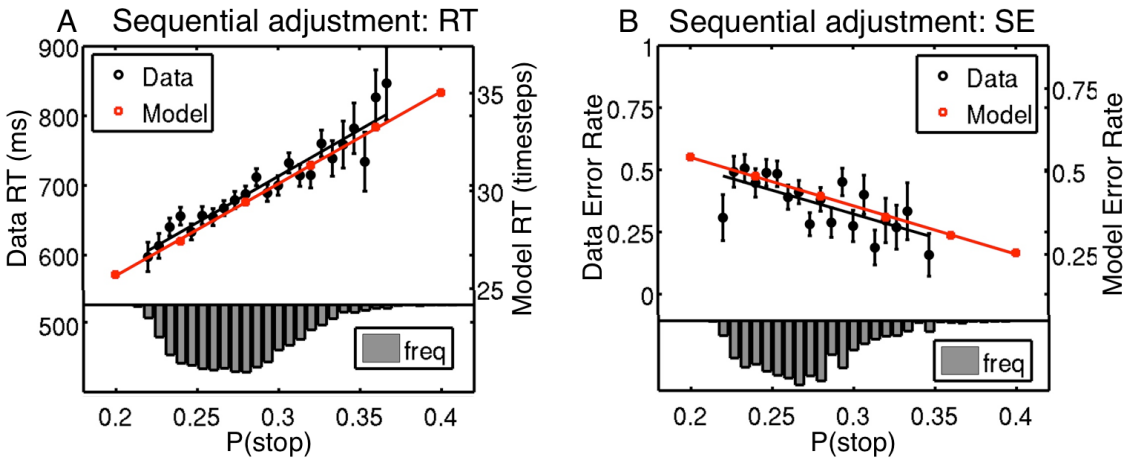

Figure 4: Sequential adjustment of RT and stop error (SE) rates. (A) Go trials were binned by model prediction of subjects' perceived trial-by-trial estimates of stop frequency (notated here as $P(stop)$ for convenience). Data are collapsed across 8 subjects, with the histogram (inverted, lower panel) indicating the empirical distribution of inferred $P(stop)$ (150 bins for the 0-1 range of probability) across trials. The Bayesian model predicts a positive linear relationship between RT and $P(stop)$ (red circles are exact simulation results, red line is best linear fit; y-axis scale on the right). Subjects also show a strong linear increase in go RT with increasing $P(stop)$ (black circles, errorbars = SEM); black line shows best linear fit ($r^2 = 0.9$, $p < 10^{-10}$; y-axis scale on the left). (B) The model predicts a linear decrease in SE rate with increasing $P(stop)$ (values binned as in (A); red circles are exact simulation results, red line is best linear fit; y-axis scale on the right). Subjects' stop error rate was also negatively correlated with $P(stop)$ (black circles, errorbars = SEM); black line shows best linear fit ($r^2 = 0.48$, $p < 0.001$); y-axis scale on the left.

should decrease after a run of go trials, and increase after a run of stop trials (Figure 3B, see Methods). A similar pattern of sequential effects is observed in human subjects performing the stop-signal task (Figure 3C; data from Emeric et al. [6]).

The comparisons in Figure 3 establish a qualitative connection between the sequential stop-frequency estimation method we propose, and go RT effects seen in behavioral data. To examine this more quantitatively, we collected data from 8 subjects performing the stop signal task, and used the DBM to infer their trial-by-trial expectation of the stop signal (see Methods). We then examined their go RT and stop error rate as a function of estimated trial-by-trial expectation of stop trial likelihood. Figure 4 summarizes the findings, comparing experimental data to the Bayesian model's predictions. The model predicts a positive, linear relationship between go RT and expected probability of encountering a stop trial, and a negative, linear relationship between stop error rate and $P(stop)$ (Figure 4A;B). The same strong linear relationships are observed in subjects' go RT (Figure 4A) and stop error rate (Figure 4B) data, plotted against the model's estimation of $P(stop)$ on each trial based on subject's experienced trial history. These data demonstrate that a significant proportion of trial-to-trial variability ("noise") in response time can be explained by human subjects' making dynamic, finely tuned, and rationally motivated adjustments in their perceptual and

decision strategy at a fine timescale.

Race model as neural approximation to Bayesian decision-making?

While the Bayesian model successfully accounts for trial-to-trial behavioral adjustments and within-trial decision-making, it is unlikely that the brain exactly implements Bayes-optimal computations in all scenarios. Instead, the brain is likely to have developed approximately optimal, computationally simpler strategies, learned over either evolutionary or developmental timescales, for ecologically relevant contexts. One possible neural mechanism is suggested by the popular race model [5], which assumes that the behavioral outcome (go or stop) on a trial depends on a race between the finishing times of a go process and a stop process. The go process is assumed to be stochastic, with a finishing time invariant between go and stop trials, and directly reflected in go RT empirical distribution. The stop process is assumed to have a mean processing latency from stop signal onset, termed the stop-signal reaction time (SSRT), whose value is *inferred* from observed go RT and stop error distributions. SSRT is a popular measure of stopping ability, with shorter SSRT implying greater facility at withholding the prepotent go response despite a late presentation of stop signal. SSRT has been measured to be longer in groups with inhibitory deficits: attention-deficit hyperactivity disorder (ADHD) [20], obsessive-compulsive disorder [3], and substance abuse [2]. Furthermore, there appear to be two pathways, "movement" and "fixation" neurons, in the superior colliculus and frontal eye field, that activate similar to the putative go and stop processes respectively, when monkeys are engaged in an oculomotor stop-signal task [21, 22, 23].

Given the neural evidence for a race-like process, we examine here the computational relationship between the Bayesian model and the race model. While algorithmically there is no intrinsic SSRT, nor independent go and stop processes, in the rational decision-making model, we can nevertheless parameterize a race model to best approximate rational decision-making in different experimental settings. This allows us to examine how much computational cost is incurred by adopting a race model approximation, as well as having the added benefit of being able to *predict* how parameters of the race model, and therefore related neural activities (e.g. movement and fixation neurons in oculomotor areas), *ought* to change as a function of task demands and behavioral goals. Without this formal mapping, the original race model alone could not have predicted *a priori* the parametric adjustments subjects might make in respond to different contextual factors, such as the clear influence of reward/motivation [14], stop signal expectancy [6, 8], trial history [6, 7], and atomoxetine administration [9].

While the original race model was agnostic with respect to how stochasticity arises in go RT [5], more recent work has hypothesized a drift-diffusion process [8]. We adopt a similar drift-diffusion implementation (Figure 5A), where go RT is the outcome of a noisy accumulate-to-threshold (diffusion) process, described by 3 parameters: the *rate* of diffusion, a *threshold* for the process at which a go response is initiated, and a temporal *offset* representing non-decision time (e.g. sensory and motor delays). SSRT, the single parameter controlling the finishing time of the stop process, constitutes the fourth parameter.

To examine how prior expectation of stop signal, $P(stop)$, affects parameters of the race

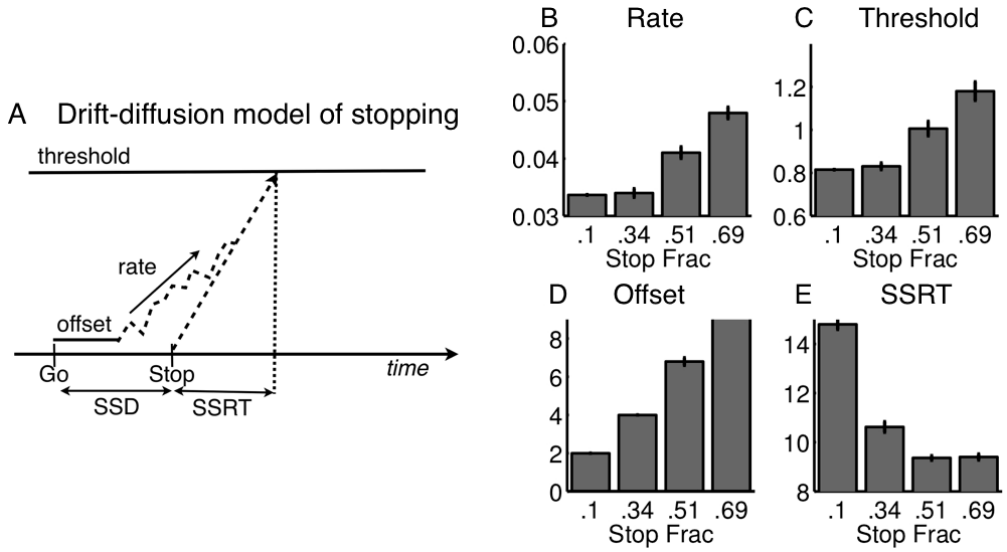

Figure 5: Stop signal frequency and the race model. (A) Drift-diffusion extension of the race model (Logan & Cowan, 1984) with four parameters: the rate of accumulation for the go process, a threshold at which responses are generated, the non-decision time (offset) for the go process, and SSRT. (B-E) Parameter settings for the DDM that best approximate Bayesian decision-making when the overall stop trial fraction is set at 0.1, 0.34, 0.51, and 0.69. The DDM fits predict that SSRT should decrease as stop trials are more frequent.

model, we generate synthetic RT and error data from the Bayesian model, and find the parameters for the race model that generates most similar empirical statistics (see Methods). Note that this is a *predictive* procedure, and not a data-fitting procedure, because the race model can leverage the fact that the Bayesian model explicitly represents $P(stop)$ and needs no tuning. Figure 5B-E show how the four parameters of the drift-diffusion race model *ought* to be adjusted in the face of changing expected stop trial frequency. In particular, higher stop signal expectancy should result in a graded decrease in SSRT, a prediction supported by some experimental data [14]; in addition, the go process is predicted to have longer non-decision latency (offset), higher threshold for responding, though also faster rate of drift toward the bound, as stop trial expectation is increased. Figure 6 shows that the rational decision-making model is well-approximated by the drift-diffusion approximation when its parameters are adjusted as in Figure 5. These results provide *testable predictions* about changes in activity profiles of neurons involved in stopping behavior [21, 22] as a function of stop trial expectancy.

Norepinephrine and subjective expectation of stop signals

The noradrenergic neuromodulatory system appears to play a critical in the stop signal task. Norepinephrine (NE) depletion, such as hypothesized in ADHD [24], has been associated with impaired stopping ability compared to healthy controls in the stop-signal task [20]. Conversely, the administration of atomoxetine, a NE-reuptake inhibitor, has been shown to improve stopping latency (sometimes with concomitant increase in go RT) not only in ADHD

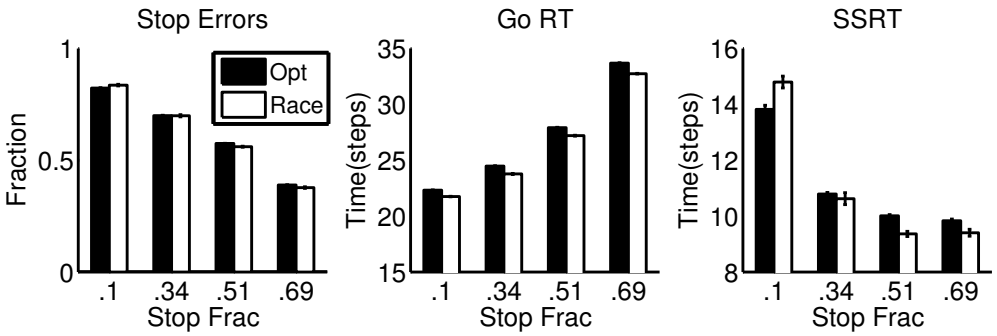

Figure 6: Rational decision-making and race model approximation. The figure compares the Bayesian model and drift-diffusion approximation as a function of the overall fraction of stop trials, on the measures of stop error rate (A), go RT (B), and SSRT (C).

patients [25], but healthy humans [9, 10] and rodents [26, 27]. To account for the modulatory influence of NE on stopping behavior, we hypothesize that NE modulates anticipated likelihood of a stop trial, a form of *unexpected uncertainty*. We previously proposed NE signals unexpected uncertainty, or the uncertainty induced by unexpected events violating default expectations [15, 16], as a unified computational explanation of its wide-ranging role in the sequential learning of stimulus and action contingencies, as well as anticipatory preparation and attentional allocation [28, 29, 30, 31, 32]. In the context of the stop-signal task, where the default expectation is the prepotent go stimulus-response, unexpected uncertainty is elicited by the infrequent stop signal. The associated release of norepinephrine could prepare downstream sensory, motor, and cognitive areas to select and execute an alternate plan of action – *stopping*.

We simulate the effects of elevated and depressed levels of norepinephrine in the model by inflating and deflating, respectively, the probabilistic expectation of a stop signal (i.e. $P(stop)$). By seeding the rational decision-making model with different prior expectations of stop trials, we make quantitative predictions of behavioral changes when norepinephrine is elevated (e.g. via atomoxetine administration) or depressed (e.g. as possibly in ADHD [24]). The model (see Methods) predicts improved stopping latency (SSRT) and increased go RT when atomoxetine is administered (Figure 7A), with the effects being dose-dependent (Figure 7C). Figure 7B;D show that rats administered atomoxetine show a similar pattern of behavioral alterations (data adapted from Bari et al. [27] and Robinson et al. [26]). ADHD patients [25] and healthy humans [9] respond similarly to atomoxetine (though also see Nandam et al. [33]).

Discussion

We have shown that Bayes-like stimulus prediction across trials, combined with rational decision-making within each trial, can account for a number of contextual influences in the stop signal task. In our computational model, an increase in perceived stop signal frequency affects inhibitory behavior in two ways: the stop stimulus is detected more rapidly (due to

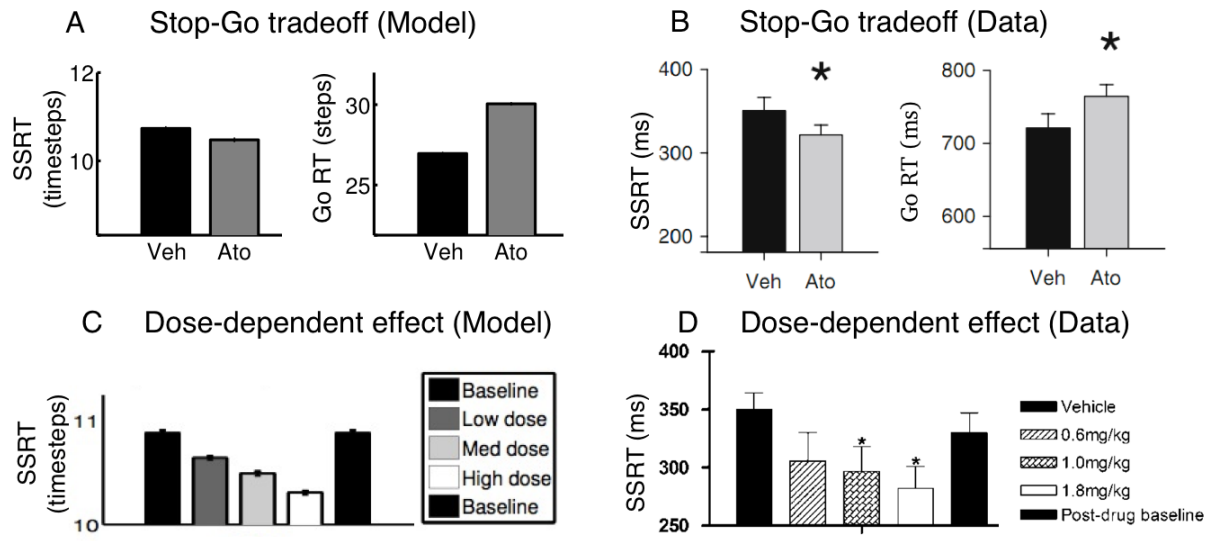

Figure 7: Effect of atomoxetine on stopping behavior. (A) The Bayesian model predicts that atomoxetine (Ato) administration), computationally realized as multiplicatively elevating stop signal anticipation, $P(stop)$, results in decreased stopping latency (SSRT; left) and increased go RT (right). (C) The model predicts SSRT to decrease dose-dependently with atomoxetine administration. (B) Atomoxetine decrease SSRT and increase go RT in rats (data adapted from Bari et al, 2009). (D) The effects of atomoxetine are dose dependent (data adapted from Robinson et al, 2008).

heightened prior expectation of stop signals), and (2) go RT is lengthened to avoid excessive stop errors. This is a rational tradeoff – the drop in stop errors on average win out over the increase in go RT, because of the anticipated relative preponderance of stop trials. Our experimental data demonstrates that subjects are highly sensitive to this tradeoff, and adjust their response strategy from trial to trial so as to continually maximize expected gain. This interpretation unifies a range of experimental results as a direct consequence of dynamically adjusted anticipatory control–behavioral adjustments in response to stop signal frequency [6], trial-by-trial behavioral "fluctuation" as a consequence of recent stimulus history [6, 7], the functional link between inhibitory deficit and noradrenergic depletion such as in ADHD patients, and improvement of inhibitory control by pharmacological agents such as atomoxetine [9, 25, 26].

It is important to note that while there appears to be a *rational basis* for contextual sensitivity in the stop-signal task, our work does not imply that subjects undergo conscious or cognitive reasoning to arrive at these "strategic" responses. Indeed, the rapid timescale of the behavioral adjustments, as well as the cross-species systematicity in anticipatory control, suggest that the necessary computations are implemented without cognitive awareness. It may well be that the brain developed these useful responses over the evolutionary or developmental timescale. Relatedly, despite the good correspondence between experimental data and predictions from the Bayesian inference and decision-making model, this work does not imply that the brain explicitly represents/computes probabilities or action costs. The correspondence merely suggests that the brain implements neural mechanisms that perform

comparable to the optimal procedure in specific experimental conditions, whether it is employing the service of a system-wide signal such as norepinephrine or two opponent neural pathways that respectively engage the go or stop responses.

Historically, the *race model* has often been used used to analyze behavior in the stop signal task, and the associated estimation of *stopping latency* or SSRT is often used as an index of inhibitory capacity. For instance, individuals affected by ADHD, obsessive-compulsive disorder, and schizophrenia have longer measured SSRTs (see Lipszyc & Schachar [20] for a meta-analysis). As we showed here and in our previous work [13], SSRT can be seen as an *emergent property* of rational decision-making, i.e., SSRT changes can be predicted by adjusting behavior rationally in various experimental contexts. This critical insight allows us to make the connection between pharmacological intervention and perceived stimulus probabilities in the task, and to predict the induced SSRT changes. In the case of ADHD patients, we suggest that potential noradrenergic depletion [24] results in an under-estimation of the perceived likelihood of the unexpected stop signal, and therefore the observed improvement in SSRT from atomoxetine administration arises from paying greater attention to rare task-relevant stimuli.

We now turn to the physiological mechanisms by which atomoxetine influences inhibitory control. It is known that atomoxetine increases extracellular levels of NE in the prefrontal cortex and other cortical areas [11, 12]. However, the effect of atomoxetine on *task-related* NE release is less well understood, although some evidence points to increased task-specific activation [10, 34, 35]. For example, Chamberlain et al. show that atomoxetine increases BOLD activity in the right inferior frontal gyrus on stop trials in the stop signal task [10]. Modafinil, a psychostimulant that also affects the norepinephrine transporter [36], is thought to increase task-related phasic activity of locus coeruleus neurons (LC, the principal source of NE in the cortex), while suppressing overall tonic activity [34]. How do the *tonic* and *phasic* modes of LC neural firing [37] cooperate in influencing inhibitory control? Previously, we proposed that phasic LC-NE activity represents unexpected uncertainty within a well-established task context in response to stimuli [16]–for example, the probability of a stop signal as modeled in this paper, or an oddball or rare target event [38, 29]. In contrast, tonic NE signals a broader uncertainty about the relevance of the task context or stimulus-action-reward mappings [15], encouraging further exploration of the environment [37]. This computational perspective, as well as experimental evidence, suggests that phasic and tonic NE are in cooperative balance (see Aston Jones et al. [37] for a review), enabling learning about task-relevant stimulus frequency at short and long timescales.

While this work has primarily focused on the role of norepinephrine in inhibitory control, in reality it works in concert with the cortical regions that it targets. Neurophysiological experiments have demonstrated that neurons in the supplementary eye field (SEF) *monitor* stimulus history in a saccadic stop signal task, with their firing rates changing trial-by-trial in response to trial history [39, 40]. SEF neuron activity strongly correlates with a moving-window estimate of stop trial frequency, and therefore tracks stimulus probabilities, either directly via or in cooperation with the LC. Another critical area implicated in monitoring behavior is the anterior cingulate cortex (ACC), which has dense projections to LC and is one of two principal cortical regions that do so [37]. The role of ACC in executive function has variously been attributed to *conflict* [41], and *error monitoring* [42]; however, in the

context of the stop signal task, conflict and errors are confounded with each other, and with the probability of encountering a stop signal. In recent work (Ide, PS, AJY, & Li, unpublished), we show that activity in dorsal ACC is explained by *expectation violation*– the difference between the expectation (probability) of the stop signal, and the observed event, whether go or stop trial. Given the strong reciprocal connectivity between ACC and LC [37], it is likely that dACC cooperates with the LC-NE system to encode information about unexpected uncertainty and the need for behavioral adaptation.

Besides norepinephrine, pharmacological studies have also implicated other neuromodulators, such as dopamine and serotonin, in inhibitory control. For example, Bari et al. [27] catalog a series of experiments in rats performing the stop signal task and demonstrate differential effects of selective reuptake inhibitors for dopamine, norepinephrine and serotonin on behavior. Further, these pharmacological manipulations have differential effects [43] on the stop signal task and the the Go/NoGo task [44], a closely related task also thought to engage the faculty of *behavioral inhibition*. An important goal for our future work is to analyze the computational role of other neuromodulatory systems in inhibitory control [33, 45].

Methods

Behavioral experiment

To examine the effects of sequential stimulus frequency estimation on inhibitory behavior, we collected data from 8 subjects performing a version of the stop signal task. Subjects were recruited from the undergraduate population in Cognitive Science at the University of California, San Diego, and received credit for participation. Subjects were trained on a random-dot motion discrimination task that has been used extensively to study the theoretical and neurobiological basis of perceptual decision-making (see e.g., Gold & Shadlen [46] for a recent review). On each trial, subjects were presented with a circular patch of dots, where a fraction of the dots moved coherently towards the left, or the right. The remaining dots moved in random directions. Trials were equally likely to contain right-moving or left-moving stimuli, and subjects were instructed to press the appropriate arrow key on a keyboard to indicate the direction of coherent motion. The coherence was chosen in a pilot study to produce low discrimination error rates (5-10% in our subjects). There was a deadline of 1.2s for the go response, indicated by a vertical bar that shrunk in height until it disappeared at the deadline. Not responding before the deadline on a go trial constituted a go omission error, while choosing the wrong alternative constituted a go choice error. Stop signals were presented as auditory beeps on a randomly selected 25% of the trials. If no response is recorded for 1.2s after go stimulus presentation, it is considered a correct stop trial; otherwise, it is considered a stop error trial. On stop trials, the stop signal was presented at a delay drawn uniformly at random from the following set: (0.1s, 0.2s, 0.3s, 0.4s, 0.5s, 0.6s). We established a clear reward structure using a points scheme: the 1.2s of trial time counted as 100 points, with subjects earning points equivalent to the time remaining before deadline when they respond. Each correct trial earned 50 points, while all error trials (go choice errors, go omission errors, and stop errors) resulted in -50 points. After the response

period, subjects received auditory feedback as well as text feedback with the points they earned on the trial.

While typical stop-signal experiments use square-vs-circle discrimination for the go stimulus, we chose the random dot task for ease of controlling and measuring noise/discriminability over time and across difficulty levels. Figure B shows that the average inhibition function, or stop error rate versus SSD, exhibits the classic sigmoidal trend of increasing SE rate with respect to SSD. The average stop error rate across subjects for the SSDs used was 40%, and the average SSRT was 229ms ± 29ms (SEM). Subjects maintained a high level of accuracy $(90 - 95\%)$ for the random-dot go stimulus discrimination component.

Learning and prediction of stop-signal frequency

A critical factor affecting behavior in the stop signal task is the prior expectation of stimulus type at the outside of the trial (denoted as P(stop) in the main text). Evidence of sequential adjustments in RT suggest that this quantity is learned and continually updated from experienced trial sequences. We previously developed a *dynamic belief model* (DBM) for estimating stimulus frequency, to account for sequential effects in human response times at simple decision-making tasks [47]. Here, we adapt DBM to the stop-signal task.

Figure 1B illustrates the generative model, based on which we propose subjects estimate the hidden variable r. The model assumes that the trial type on trial k, denoted as s_k (=1 for stop trial, =0 for go trial; for simplicity, we assume there is no perceptual confusion about trial type upon completion of the trial) is Bernoulli distributed with rate parameter r_k, and is independent of the other trials given r_k. The stop signal frequency is subject to unsignaled changes: with probability α it stays the same as last trial; with probability $1-\alpha$, it is re-sampled from a prior distribution $p_0(r)$. In simulations, we choose $p_0(\cdot)$ to be a Beta distribution with a bias towards small r ($Beta(2.5, 7.5)$).

To infer the current posterior distribution over r_k given experienced trials, $\mathbf{s}_k \triangleq (s_1, \ldots, s_k)$, we apply Bayes' Rule:

$$p(r_k|\mathbf{s}_k) \propto P(s_k|r_k)p(r_k|\mathbf{s}_{k-1}) \ . \tag{1}$$

The predictive distribution $p(r_k|\mathbf{s}_{k-1})$ on the right hand side is a mixture of the previous trial's posterior distribution and the prior distribution, with α and $1-\alpha$ acting as the respective mixing coefficients:

$$p(r_k|\mathbf{s}_{k-1}) = \alpha p(r_{k-1}|\mathbf{s}_{k-1}) + (1-\alpha)p_0(r_k) \ . \tag{2}$$

This reveals the iterative procedure by which the posterior distribution can be updated after each trial.

To compute the predicted likelihood of trial k being a stop trial based on observations up to the last trial $\mathbf{s}_{k-1}$, denoted as P(stop) in the main text, we note that

$$P(s_k = 1|\mathbf{s}_{k-1}) = \int P(s_k = 1|r_k)p(r_k|\mathbf{s}_{k-1})dr_k = \int r_k p(r_k|\mathbf{s}_{k-1})dr_k = \langle r_k|\mathbf{s}_{k-1}\rangle \ . \tag{3}$$

In other words, the predicted likelihood of a stop trial is just the expectation (mean) of the predictive distribution $p(r_k|\mathbf{s}_{k-1})$.

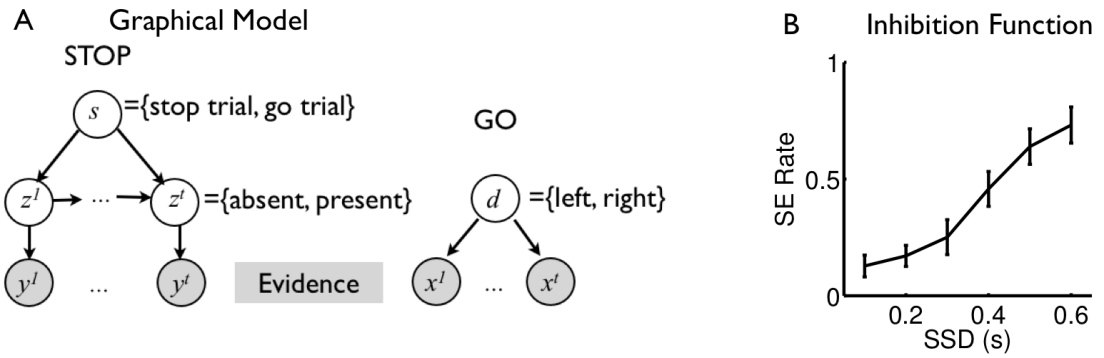

Figure 8: Random dot version of the stop signal task and basic behavioral data. (A) In the generative model for within-trial sensory inference, we assume two separate streams of observations, x^t, y^t, are associated with the go and stop stimuli respectively. x^t depends on the identity of the target, $d \in 0, 1$. y^t depends on whether the current trial is a stop trial, $s \in 0, 1$, and whether the stop signal has already appeared by time t, $z^t \in 0, 1$. (B) Inhibition function for human data we collected: stop error rate increased as a function of stop signal delay, consistent with classic results in the task.

In our experiments, we used a Beta prior with a preference towards low values of r_k, $Beta(2.5, 7.5)$, for the prior distribution$p_0(\cdot)$. The mixture coefficient α was set to 0.75 based on our previous results [19]. We also implemented a computationally simpler alternative to DBM using a best-fitting exponential linear filter on previous occurrences of the stimulus, instead of doing the nonlinear Bayesian computations. Previously, we showed that the computations required by DBM are well approximated by an exponential linear filter with appropriately tuned parameters [19]. We were therefore not surprised to find that the linear filter model produced behavioral predictions statistically indistinguishable from DBM (data not shown).

Rational decision-making in the stop signal task

We assume that subjects sequentially update their beliefs about stop signal frequency r_k according to the Bayesian procedure described in the last section, and incorporate this prior into a rational decision process we previously proposed to underlie within-trial computations [13]. Details of this rational decision-making model can be found in Shenoy & Yu (2011); we summarize the key elements here and highlight how the trial-by-trial DBM sequential prediction model interacts with the within-trial rational decision-making model.

Our Bayesian ideal observer assumes that subjects maintain a continually evolving, Bayes-optimal belief state about stimulus properties, and that they make an optimal decision between *go* and *wait* at each moment in time as a function of their belief state, in order to minimize the average (expected) cost of response delay (time+energy+opportunity cost), stop errors, and go errors. The belief state (p_d^t, p_s^t) consists of the posterior probability of the go stimulus requiring a response 1 (versus 0) and the stop signal being present (versus absent), respectively – it is updated at each instant t (where t indexes a small temporal increment on the order of $10 - 20ms$) by inverting the generative model in Fig. via Bayes'

Rule. The likelihood functions are assumed to be symmetric Bernoulli distributions, with rate parameters λ_g and λ_s specifying the detectability of go and stop stimuli, respectively. The prior over d reflects equal prevalence of the two alternatives in the go discrimination task: $P(d=1)=.5$. *The prior probability of the current trial being a stop trial is the predictive probability obtained from the DBM model for sequential estimation stop signal frequency:* $P(s=1)=P(s_k=1|\mathbf{s}_{k-1})$. The stop signal appears at a stochastic time θ (SSD) under an assumption of constant hazard rate.

The average cost function has the form:

$$L_\pi = c\langle\tau\rangle + c_s r P(\tau<D|s=1) + (1-r)P(\tau<D,\delta\neq d|s=0) + (1-r)P(\tau=D|s=0) \quad (4)$$

where τ is go response time, δ is the chosen go response, c parameterizes time cost, and c_s parameterizes stop error penalty (go error penalty is set to be 1, as only the relative magnitudes of these costs matter). The policy π maps each value of the belief state (p_d, p_s) at time t into the action space $(go, wait)$, thus controlling both the distribution of response times and errors, and in turn the overall expected cost. We compute the optimal policy using dynamic programming (up to discretization of the belief state space) [48].

SSRT estimation for the rational decision-making model

The race model assumes that behavioral outcome (go or stop) on a trial depends on a race between the finishing times of a go process and a stop process. It proposes some minimal assumptions under which the (not directly observable) SSRT can be estimated from behavioral data: go and stop processes are independent (specifically, the go process does not differ between go and stop trials), and that the stop process can be well characterized using a single parameter, its mean processing time or SSRT. The race model assumes that the finishing time of the stop process is a sum of SSD and SSRT, and that stop errors arise when the go process finishes before the stop process. Thus, the stop error rate at a particular SSD is the area under the go RT distribution up to SSD+SSRT. Although SSRT is not an intrinsic component of our optimal model, we can nevertheless estimate it as an empirical measure based on go RT and stop error distributions generated from Monte Carlo simulations of the model. We use a range of SSD's for each simulation condition, and find the SSD at which the model makes 50% stop errors, then subtract that SSD from the median of the go RT distribution to obtain the SSRT.

Race model approximation of the rational decision-making model

We approximate the rational decision-making model by extending the race model [5] as a random walk model, the discrete analog of the well-studied drift-diffusion model long used to model reaction times [49, 50, 51, 52, 46, 53]. In our implementation (Figure 5), the go process consists of a constant drift *rate* with a starting point or *offset,* and additive, cumulative Gaussian noise on each time step of a trial. The stop process is modeled as a fixed-latency process with a corresponding latency parameter (*SSRT*), to closely mirror the assumptions underlying the actual estimation of SSRT in practice [5].

The basic random-walk model has the following form:

$$k(t+1) = k(t) + b + w \tag{5}$$

with drift parameter b, and Gaussian white noise $w \sim \mathcal{N}(0, \sigma^2)$. We assume that a go decision is made when $k(\tau)$ first crosses the threshold h or $-h$, whichever first (the side determines which of the two go responses is chosen). The simulated RT for a trial is $\tau + \tau_0$, where τ_0 is a temporal offset parameter denoting non-decision time [51, 53]. For $b > 0$, the response is correct if h is first crossed, and incorrect if $-h$ is first crossed; vice versa if $b < 0$.

In addition to the three parameters for the go process, SSRT constitutes a fourth parameter, determining the finishing time of the stop process (it finishes at time SSD+SSRT after the go stimulus onset). On stop trials, if $\tau + \tau_0 <$ SSD+SSRT, an error response occurs at $\tau + \tau_0$; otherwise, it is a correct stop trial and τ is assumed to take on the value ∞. To approximate optimal decision-making with this race model implementation, we choose a parameterization $\boldsymbol{\theta} \triangleq (b, h, \tau_0, \text{SSRT})$ so that the joint distribution of RT and choice based on simulation of the optimal model, $p_{opt}(\tau, \delta)$, is well-approximated by that of the race model $p(\tau, \delta|\boldsymbol{\theta})$. Specifically, we choose model parameters that minimize the KL-divergence between the two distributions (or equivalently, maximizing the likelihood of observing the empirical statistics of the optimal model by a particular setting of the race model):

$$\begin{aligned} \boldsymbol{\theta}^* &= \operatorname*{argmin}_{\boldsymbol{\theta}} \text{KL}[p_{opt}(\tau, \delta), p(\tau, \delta|\boldsymbol{\theta})] \\ &\approx \operatorname*{argmax}_{\boldsymbol{\theta}} \sum_{i=1}^{n} \log p(\tau_i, \delta_i|\boldsymbol{\theta}) \qquad \text{where } (\tau_i, \delta_i) \sim p_{opt}(\tau, \delta) \end{aligned} \tag{6}$$

The expectation of the log likelihood ratio $\langle \log p(\tau, \delta)/p(\tau, \delta|\boldsymbol{\theta}) \rangle$ under the distribution $p(\tau, \delta)$ is approximated by a finite sum based on samples from $p(\tau, \delta)$. For a given setting of θ, we compute $p(\tau, \delta|\boldsymbol{\theta})$ exactly, up to a discretization of values of $k(t)$, by convolving $p(k(t))$ with Gaussian noise and removing probability mass beyond both thresholds at the next timestep, to get $p(k(t+1))$. This gives $p(\tau, \delta|\boldsymbol{\theta})$ on go trials. On stop trials, we truncate the distribution at SSD+SSRT, which then gives us both the stop error rate and error RT distribution on stop trials, for each SSD. We evaluate eq. 6 using $n = 10000$ samples from the optimal model, $\{(\tau_i, \delta_i)\}_{i=1}^n$, and use Matlab's `fmincon` function to find the best-fitting parameter θ for different simulations of the optimal model corresponding to the stop signal frequency $r = \{0.1, 0.3.0.50.7\}$.

Acknowledgments

The authors thank Joseph Schilz for assistance in collecting behavioral data for the stop signal task.

References

[1] Nigg J (2001) Is ADHD a disinhibitory disorder? Psychological Bulletin 127: 571.

[2] Nigg JT, Wong MM, Martel MM, Jester JM, Puttler LI, et al. (2006) Poor response inhibition as a predictor of problem drinking and illicit drug use in adolescents at risk for alcoholism and other substance use disorders. Journal of Amer Academy of Child & Adolescent Psychiatry 45: 468.

[3] Menzies L, Achard S, Chamberlain SR, Fineberg N, Chen CH, et al. (2007) Neurocognitive endophenotypes of obsessive-compulsive disorder. Brain 130: 3223.

[4] Iacono WG, Malone SM, McGue M (2008) Behavioral Disinhibition and the Development of Early-Onset Addiction: Common and Specific Influences. Annual Review of Clinical Psychology 4: 325–348.

[5] Logan G, Cowan W (1984) On the ability to inhibit thought and action: A theory of an act of control. Psychological review 91: 295.

[6] Emeric EE, Brown JW, Boucher L, Carpenter RHS, Hanes DP, et al. (2007) Influence of history on saccade countermanding performance in humans and macaque monkeys. Vision research 47: 35–49.

[7] Li CR, Huang C, Yan P, Paliwal P, Constable RT, et al. (2008) Neural correlates of post-error slowing during a stop signal task: A functional magnetic resonance imaging study. Journal of cognitive neuroscience 20: 1021–1029.

[8] Verbruggen F, Logan GD (2009) Proactive adjustments of response strategies in the stop-signal paradigm. Journal of Experimental Psychology: Human Perception and Performance 35: 835–854.

[9] Chamberlain SR, Müller U, Blackwell AD, Clark L, Robbins TW, et al. (2006) Neurochemical modulation of response inhibition and probabilistic learning in humans. Science 311: 861–3.

[10] Chamberlain SR, Hampshire A, Müller U, Rubia K, Del Campo N, et al. (2009) Atomoxetine modulates right inferior frontal activation during inhibitory control: a pharmacological functional magnetic resonance imaging study. Biological psychiatry 65: 550–5.

[11] Swanson CJ, Perry KW, Koch-Krueger S, Katner J, Svensson KA, et al. (2006) Effect of the attention deficit/hyperactivity disorder drug aomoxetine on extracellular concentrations of norepinephrine and dopamine in several brain regions of the rat. Neuropharmacology 50: 755–760.

[12] Invernizzi R, Garattini S (2004) Role of presynaptic [alpha] 2-adrenoceptors in antidepressant action: recent findings from microdialysis studies. Progress in Neuro-Psychopharmacology and Biological Psychiatry 28: 819–827.

[13] Shenoy P, Yu AJ (2011) Rational decision-making in inhibitory control. Frontiers in human neuroscience 5.

[14] Leotti LA, Wager TD (2009) Motivational influences on response inhibition measures. J Exp Psychol Hum Percept Perform .

[15] Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. Neuron 46: 681–92.

[16] Dayan P, Yu AJ (2006) Phasic norepinephrine: a neural interrupt signal for unexpected events. Network (Bristol, England) 17: 335–50.

[17] Hanes D (1995) Countermanding saccades in macaque. Visual Neuroscience .

[18] Eagle DM, Robbins TW (2003) Inhibitory control in rats performing a stop-signal reaction-time task: effects of lesions of the medial striatum and d-amphetamine. Behavioral neuroscience 117: 1302–17.

[19] Yu AJ, Cohen JD (2009) Sequential effects: Superstition or rational behavior? Advances in Neural Information Processing Systems 21: 1873–1880.

[20] Lipszyc J, Schachar R (2010) Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task. Journal of the International Neuropsychological Society 1: 1–13.

[21] Hanes DP, Patterson WF, Schall JD (1998) The role of frontal eye field in countermanding saccades: Visual, movement and fixation activity. Journal of Neurophysiology 79: 817–834.

[22] Pare M, Hanes DP (2003) Controlled movement processing: superior colliculus activity associated with countermanded saccades. Journal of Neuroscience 23: 6480–6489.

[23] Boucher L, Palmeri TJ, Logan GD, Schall JD (2007) Inhibitory control in mind and brain: an interactive race model of countermanding saccades. Psychological Review 114: 376–397.

[24] Prince J (2008) Catecholamine dysfunction in attention-deficit/hyperactivity disorder: an update. Journal of clinical psychopharmacology 28: S39.

[25] Chamberlain SR, Del Campo N, Dowson J, Müller U, Clark L, et al. (2007) Atomoxetine improved response inhibition in adults with attention deficit/hyperactivity disorder. Biological psychiatry 62: 977–84.

[26] Robinson ESJ, Eagle DM, Mar AC, Bari A, Banerjee G, et al. (2008) Similar effects of the selective noradrenaline reuptake inhibitor atomoxetine on three distinct forms of impulsivity in the rat. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 33: 1028–37.

[27] Bari A, Eagle DM, Mar AC, Robinson ESJ, Robbins TW (2009) Dissociable effects of noradrenaline, dopamine, and serotonin uptake blockade on stop task performance in rats. Psychopharmacology 205: 273–83.

[28] Rajkowski J, Kubiak P, Aston-Jones P (1994) Locus coeruleus activity in monkey: phasic and tonic changes are associated with altered vigilance. Synapse 4: 162–164.

[29] Aston-Jones G, Rajkowski J, Kubiak P (1997) Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. Neuroscience 80: 697–715.

[30] Clayton EC, Rajkowski J, Cohen JD, Aston-Jones G (2004) Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. J Neurosci 24: 9914–9920.

[31] Bouret S, Sara SJ (2004) Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. Eur J Neurosci 20: 791–802.

[32] Rajkowski J, Majcznynski H, Clayton E, Aston-Jones G (2004) Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. J Neurophysiol 92: 361–371.

[33] Nandam LS, Hester R, Wagner J, Cummins TDR, Garner K, et al. (2010) Methylphenidate But Not Atomoxetine or Citalopram Modulates Inhibitory Control and Response Time Variability. Biological psychiatry : 2010–2012.

[34] Minzenberg M, Watrous A, Yoon J, Ursu S, Carter C (2008) Modafinil shifts human locus coeruleus to low-tonic, high-phasic activity during functional MRI. Science 322: 1700.

[35] Aston-Jones G, Gold J (2009) How We Say No: Norepinephrine, Inferior Frontal Gyrus, and Response Inhibition. Biological psychiatry 65: 548.

[36] Madras B, Xie Z, Lin Z, Jassen A, Panas H (2006) Modafinil occupies dopamine and norepinephrine transporters in vivo and modulates the transporters and trace amine activity in vitro. Journal of pharmacology and experimental therapeutics 319: 561–569.

[37] Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annual review of neuroscience 28: 403–50.

[38] Nieuwenhuis S, Aston-Jones G, Cohen JD (2005) Decision making, the P3, and the locus coeruleus-norepinephrine system. Psychological bulletin 131: 510–32.

[39] Stuphorn V, Schall JD (2006) Executive control of countermanding saccades by the supplementary eye field. Nature neuroscience 9: 925–31.

[40] Stuphorn V, Brown JW, Schall JD (2010) Role of supplementary eye field in saccade initiation: executive, not direct, control. Journal of neurophysiology 103: 801–16.

[41] Kerns J, Cohen J, MacDonald A, Cho R, Stenger V, et al. (2004) Anterior cingulate conflict monitoring and adjustments in control. Science 303: 1023.

[42] Brown JW, Braver TS (2005) Learned predictions of error likelihood in the anterior cingulate cortex. Science (New York, NY) 307: 1118–21.

[43] Eagle D, Bari A, Robbins T (2008) The neuropsychopharmacology of action inhibition: cross-species translation of the stop-signal and go/no-go tasks. Psychopharmacology 199: 439–456.

[44] Donders F (1969) On the speed of mental processes. Acta Psychologica 30: 412.

[45] Li C, Morgan P, Matuskey D, Abdelghany O, Luo X, et al. (2010) Biological markers of the effects of intravenous methylphenidate on improving inhibitory control in cocaine-dependent patients. Proceedings of the National Academy of Sciences 107: 14455–14459.

[46] Gold J, Shadlen M (2007) The neural basis of decision making. Annu Rev Neurosci 30: 535–574.

[47] Cho RY, Nystrom LE, Brown ET, D JA, S BT, et al. (2002) Mechanisms underlying dependencies of performance on stimulus history in a two-alternative foced choice-task. Cognitive, Affective, & Behavioral Neuroscience 2: 283–299.

[48] Bellman R (1952) On the theory of dynamic programming. Proceedings of the National Academy of Sciences of .

[49] Stone M (1960) Models for choice-reaction time. Psychometrika 25: 251–260.

[50] Laming D (1968) Information theory of choice-reaction times. Academic Press.

[51] Ratcliff R (1978) A theory of memory retrieval. Psychological Review 85: 59.

[52] Hanes D, Schall J (1996) Neural control of voluntary movement initiation. Science 274: 427.

[53] Bogacz R, Brown E, Moehlis J, Holmes P, Cohen J (2006) The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. Psychological Review 113: 700.