

# Stop paying attention: the need for explicit stopping in inhibitory control

Ning Ma (nima@eng.ucsd.edu)

Department of Electrical and Computer Engineering, University of California San Diego  
9500 Gilman Drive, La Jolla, CA 92037 USA

Angela J. Yu (ajyu@ucsd.edu)

Department of Cognitive Science, University of California San Diego  
9500 Gilman Drive, La Jolla, CA 92037 USA

## Abstract

Inhibitory control, the ability to stop inappropriate actions, is an important cognitive function often investigated via the stop-signal task, in which an infrequent stop signal instructs the subject to stop a default go response. Previously, we proposed a rational decision-making model for stopping, suggesting the observer makes a repeated Go versus Wait choice at each instant, so that a Stop response is realized by repeatedly choosing to Wait. We propose an alternative model here that incorporates a third choice, Stop. Critically, unlike the Wait action, choosing the Stop action not only blocks a Go response at the current moment but also for the remainder of the trial – the disadvantage of losing this flexibility is balanced by the benefit of not having to pay attention anymore. We show that this new model both reproduces known behavioral effects and has internal dynamics resembling presumed Go neural activations in the brain.

**Keywords:** Bayesian model, decision-making, stochastic control theory, inhibitory control, stop-signal task, neural data

## Introduction

Humans and animals are often faced with the need to choose, under time pressure, an action among options with uncertain consequence. The ability to dynamically withhold or modify planned actions according changing task conditions is known as *inhibitory control*. In psychology and neuroscience, inhibitory control has been studied extensively using the stop-signal task (Logan & Cowan, 1984). In this task, subject performs a default go task on each trial, usually consisting of two-alternative forced choice (2AFC) discrimination between two stimuli (e.g. press "L" for square, press "R" for circle). On a small fraction of trials, an additional *stop* signal occurs at some time (known as SSD, or stop-signal delay) after the go stimulus onset, and the subject is instructed to withhold or stop the *go* response. When the subject succeeds to stop, the trial is considered a stop success (SS) trial; otherwise, it is considered a stop error (SE) trial. Typically the SSD is chosen such that subjects on average only achieve 50% accuracy on the stop trials.

The classical model for the stop signal task is the race model (Logan & Cowan, 1984; Boucher, Palmeri, Logan, & Schall, 2007), which is a mechanistic account that posits a race to threshold between independent *go* and *stop* processes (See Fig. 1). A stop trial results in SE if the *go* response is processed before the stop process. The race model also defines a subject-specific *stop signal reaction time* (SSRT), which is a measure of the average amount of time the observer requires to process the stop signal and cancel the go

response (in practice, it is often calculated as the difference between mean Go RT and the SSD specific to each subject for achieving 50% accuracy on stop trials). Although the race model provides a simple and elegant description of the basic behavioral phenomena, it is not a normative account of how the brain *ought* to treat the stop signal task according to task demands or behavioral goals. As such, it does not have a principled basis for predicting how behavior might change according to changes in task conditions or motivational factors. It also does not possess the representational richness to identify the computational functions of all the different brain areas implicated in the stop-signal task, or to explain the distinct causes of the myriad inhibitory deficits observed in various psychiatric conditions (e.g. ADHD, depression, OCD, drug addiction).

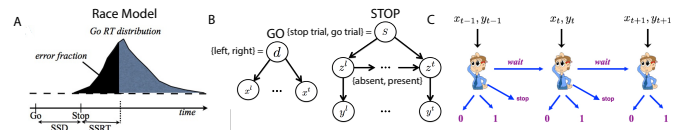


Figure 1: Models for inhibitory control in the stop-signal task. (A) The classical race mode posits that the behavioral outcome (go or stop) is determined by a race between two independent go and stop processes. (B) Bayesian graphical model for noisy sensory data generation in the rational decision-making model. (C) The decision of whether/when to Go, which Go response to select, and whether/when to Stop, are modeled as sequential decision-making, where the subject chooses at each moment whether to select a Go or Stop response, or to wait at least one more time point.

In part to overcome some of these challenges, we previously proposed a normative Bayesian Markov decision process (MDP) model for stopping (Shenoy & Yu, 2011), which assumed that subjects maintains a continually evolving, Bayes-optimal belief state about stimulus properties, and that they make an moment-by-moment optimal decision between *go* and *wait* by mapping the current belief state to optimal action. We showed that this model accounts for a range of classical and more subtle behavioral effects in the stop-signal task, includes the possibility of predicting how experimental manipulations of different cognitive factors should affect stopping behavior (Shenoy & Yu, 2011; J. S. Ide, Shenoy,

Yu\*, & Li\*, 2013; Ma & Yu, 2015, 2016). The model was also used to successfully identify brain regions involved in representing and predicting the probability of encountering a stop signal in healthy human subjects (J. S. Ide et al., 2013), as well as how that neural representation becomes altered in users of stimulants such as cocaine and methamphetamine (J. Ide, Hu, Zhang, Yu, & Li, 2015; Harlé et al., 2014, 2015).

One critical assumption made by the previous model (Shenoy & Yu, 2011) is that only two possible actions are entertained at each moment within the trial, Go or Wait. A Stop response is only realized in the model by choosing the Wait action repeatedly until the response deadline. Essentially, this model implies that SSRT is an emergent property, that there is no underlying stop process that terminates in a stop action. However, empirical evidence, including neuroimaging data in humans (Aron et al., 2007) and neurophysiology data in monkeys (Hanes, Schall, & Patterson, 1998), suggests that the brain may instead execute an explicit *stop* action on successful stop trials. However, adding a Stop action to the action set is non-trivial, because if both Wait and Stop block the Go action in the current moment in time, but Stop in addition blocks the possibility of choosing Go in the future while Wait allows that possibility, then it is always more rewarding to choose Wait over Stop in order to keep that possibility open. There must be some additional benefit to the Stop action that would make it worthwhile to execute. We hypothesize here that the extra benefit is a certain savings in attentional cost, such that choosing Stop alleviates the observer from the cost associated with attending to the sensory input and engaging with the task for the remainder of the trial.

Specifically, we formulate a novel Bayesian MDP model for inhibitory control, in which there are three explicit actions available to the decision-maker: Go, Wait, and Stop. In addition to the four kinds of behavioral costs included in the original model – the cost of making the wrong go response, the cost of not responding on a go trial, the cost of not stopping on a stop trial, and the cost of time incurred proportional to the total length of the trial – we incorporate an additional term, the cost of attending to the sensory input, which is terminated by either a Go action, a Stop action, or the expiration of the response period. Analogous to the stop-signal reaction time (SSRT) assumed by the Race Model, we define the Stop RT to be the temporal delay between the onset of stop signal and the time when the *Stop* action is chosen.

In the following, we first describe the model, then show how the model captures a variety of behavioral phenomena observed in the stop signal task, as well as neural data implicating neurons in the frontal eye field in participating in the initiation or execution of the Stop action. We conclude with some discussion of related work and thoughts on future directions.

## The Model

As in the earlier MDP model (Shenoy & Yu, 2011), this MDP model consists of two key components, a monitoring compo-

nent that formalizes sensory processing as iterative Bayesian posterior inference based on conditionally iid data, and a decision process that applies an optimal stochastic control policy. We describe the two components below, and show how the model behaves on different trial types.

## Monitoring process as Bayesian statistical inference

We use the same Bayesian inference model proposed in (Shenoy & Yu, 2011) to implement the sensory processing component, and thus provide only a short overview here. Fig. 1B shows the graphical model, whereby the two hidden variables correspond respectively to the identity of the go stimulus,  $d \in \{0, 1\}$ , and whether this trial is stop trial,  $s \in \{0, 1\}$ . The priors of  $d$  and  $s$ , in our model, are  $P(d = 1) = 0.5$  and  $r = P(s = 1) = 0.25$ , consistent with general experimental settings. Conditioned on the go response identity  $d$ , a sequence of iid sensory inputs are generated on each trial,  $x^1, \dots, x^t, \dots$ , where  $t$  indexes time step within a trial. The likelihoods of the sensory inputs given  $d$  are  $f_0(x^t) = p(x^t | d = 0)$  and  $f_1(x^t) = p(x^t | d = 1)$ , which are assumed to be Bernoulli distribution with distinct rate parameters  $q_d$  and  $1 - q_d$ , respectively. The dynamic variable  $z^t$  denotes the presence/absence of the stop signal.  $z^1 = \dots = z^{\theta-1} = 0$  and  $z^\theta = z^{\theta+1} = \dots = 1$  if a stop signal appears at time  $\theta$ . For simplicity, we assume that the onset of the stop signal  $\theta$  follows a geometric distribution:  $P(\theta = t | s = 1) = q(1 - q)^{t-1}$ . Conditioned on  $z^t$ , a stream of iid observations are generated on each trial. The likelihoods of the the sensory inputs, associated with the stop signal, are  $p(y^t | z^t = 0) = g_0(y^t)$  and  $p(y^t | z^t = 1) = g_1(y^t)$ . We still assume that the likelihood functions,  $g_0$  and  $g_1$ , are Bernoulli distributions with distinct parameters  $q_s$  and  $1 - q_s$ .

In the recognition model, Bayes' Rule is applied in the usual iterative manner way to compute the the sequential posterior probability associated with go stimulus identity,  $p_d^t = P(d = 1 | \mathbf{x}^t)$ , where  $\mathbf{x}^t = \{x^1, x^2, \dots, x^t\}$  denotes all the data observed so far. Similarly, computing the posterior probability that the stop signal is already been present,  $p_z^t = P(\theta < t | \mathbf{y}^t)$ , involves inverting the generative model, which is a simple version of a hidden Markov model.  $p_z^t$  then can be used to compute the posterior probability that the current trial is a stop trial,  $p_s^t = P(s = 1 | \mathbf{y}^t) = p_z^t + (1 - p_z^t)P(s = 1 | \theta > t, \mathbf{y}^t)$ , where  $P(s = 1 | \theta > t, \mathbf{y}^t)$  represents the probability that the stop signal will occur in the future. The *belief state* at time  $t$  is defined to be the vector  $\mathbf{b}^t = (p_d^t, p_s^t)$ .

## Decision process as optimal stochastic control

In each trial, the subject is required to make response to a go stimulus by a response deadline  $D$ , or else the trial terminates and the response is recorded as *stop*. We define a loss function that accounts for the cost and penalty structure of the stop-signal task, and assume that the observer minimizes the expected value of this loss function in choosing whether to *go*, *wait*, or *stop* at each moment in time, based on the current belief state.

Like the earlier MDP model (Shenoy & Yu, 2011), we assume that there is a basic cost  $c_r$  per unit time on each trial if the current trial is not terminated. In addition, the subject has to pay an extra attentional cost  $c_a$  per unit time if it decides to continually collect and process the new coming sensory input. The subject can explicitly choose the *stop* action to stop processing the sensory information and take the benefit of only paying the basic cost  $c_r$  in the rest of trial. Once the stop action is instantiated, the subject will no longer choose the go action again, thus will incur a unit penalty cost for missing a go option on a go trial. Making response to the *go* stimulus will terminate the current trial and save the subject the basic and attentional costs in the rest trial, but in the price of paying a unit penalty cost for making a discrimination error on a go trial or a penalty cost  $c_s$  for responding on a stop trial. The subject can also take the *wait* action to process new sensory information and update the belief state in the next time step.

Let  $\tau_r$  denote the trial termination time and  $\tau_a$  the time the subject is involved in the task, so that  $\tau_a = \tau_r < D$  if a go response is made before the deadline  $D$ ,  $\{\tau_r = D, \tau_a < D\}$  if an explicit stop action is chosen, and  $\tau_a = D$  if the subject waits until the deadline.  $\delta \in \{0, 1\}$  represents the possible binary discriminations produced by making a go response. We assume the loss function:

$$l(\tau_r, \tau_a, \delta; d, s, \theta, D) = (c_r + c_a)\tau_a \mathbf{1}_{\{\tau_r = \tau_a < D\}} + c_s \mathbf{1}_{\{\tau_r < D, s=1\}} + \mathbf{1}_{\{\tau_r < D, \delta \neq d, s=0\}} + (c_a \tau_a + c_r D) \mathbf{1}_{\{\tau_r = D\}} + \mathbf{1}_{\{\tau_r = D, s=0\}}$$

where the first three terms correspond to the cost for taking the go action and the last two denote the cost for taking stop action or waiting until the deadline. The optimal decision policy will minimize the expected loss,  $L_\pi = \mathbb{E}[l(\tau_r, \tau_a, \delta; d, s, \theta, D)]$ ,

$$L_\pi = (c_r + c_a)\mathbb{E}[\tau_a]P(\tau_a = \tau_r < D) + c_s r P(\tau_r < D | s = 1) + (1 - r)P(\tau_r < D, \sigma \neq d | s = 0) + (c_a \mathbb{E}[\tau_a] + c_r D)P(\tau_r = D) + (1 - r)P(\tau_r = D | s = 0)$$

It is computationally intractable to directly minimize  $L_\pi$  over the policy space. Fortunately, Bellman's dynamic programming principle provides an iterative relationship between the optimal state-value function and optimal action-value function. The Bellman optimality equation for optimal state-value function,  $V^t(\mathbf{b}^t)$ , is

$$V^t(\mathbf{b}^t) = c_a + c_r + \min_a \int P(\mathbf{b}^{t+1} | \mathbf{b}^t; a) V^{t+1}(\mathbf{b}^{t+1}) d\mathbf{b}^{t+1}$$

where  $a$  ranges over all possible actions. In our alternative model, the action space is  $\{go, stop, wait\}$  associated with three optimal action-value functions (also called Q-factors),  $Q'_g(\mathbf{b}^t)$ ,  $Q'_s(\mathbf{b}^t)$ , and  $Q'_w(\mathbf{b}^t)$ , respectively. Using the Bellman optimality equation for optimal action-value function- $Q(\mathbf{b}, a) = \mathbb{E}[c_a + c_r + V(\mathbf{b}^{t+1}) | \mathbf{b}^t = \mathbf{b}, a_t = a]$ , we can obtain the three Q-factors

$$\begin{aligned} Q'_g(\mathbf{b}^t) &= (1 - p'_s) \min(p'_d, 1 - p'_d) + c_s p'_s \\ Q'_a(\mathbf{b}^t) &= c_r(D - t) + (1 - p'_s) \\ Q'_w(\mathbf{b}^t) &= c_a + c_r + \mathbf{1}_{\{D > t+1\}} \mathbb{E}[V^{t+1}(\mathbf{b}^{t+1}) | \mathbf{b}^t]_{\mathbf{b}^{t+1}} \\ &\quad + \mathbf{1}_{\{D = t+1\}} (1 - p'_s) \\ V^t(\mathbf{b}^t) &= \min(Q'_g, Q'_s, Q'_w) \end{aligned}$$

Note that, in our model, the optimal state-value function and action-value functions only account for the future cost after the current time step, regardless of how much cost has been paid before, since only the expected futures costs matter in adjudicating among the action options. The optimal state-value function is the smallest of three optimal action-value functions. The optimal policy chooses the action corresponding to the smallest Q-factor at each time step. The value of discrimination response,  $\delta$ , is 1 if  $p'_d > 0.5$  and 0 otherwise. Since the subject can no longer update the belief state nor take any action at the deadline, the optimal state-value function can be initially computed at  $D$  as  $V^t(\mathbf{b}^D) = 1 - p'_s^D$ . The recursive relationship between the optimal action-value and state-value functions in Bellman optimality equation allows us to compute the optimal state-value functions and Q factors backwards in time from  $t = D - 1$  to  $t = 1$ .

In the last section, we showed that the belief state  $\mathbf{b}^{t+1}$  is a deterministic function of  $\mathbf{b}^t$  and the observations. Thus, given  $V^{t+1}$ , we can compute  $\mathbb{E}[V^{t+1}(\mathbf{b}^{t+1}) | \mathbf{b}^t]$  by averaging over all possible next observations  $x^{t+1}, y^{t+1}$ .

$$\begin{aligned} \mathbb{E}[V^{t+1}(\mathbf{b}^{t+1}) | \mathbf{b}^t] &= \sum_{x^{t+1}, y^{t+1}} p(x^{t+1}, y^{t+1} | \mathbf{b}^t) V^{t+1}(\mathbf{b}^{t+1}(\mathbf{b}^t, x^{t+1}, y^{t+1})) \\ p(x^{t+1}, y^{t+1} | \mathbf{b}^t) &= p(x^{t+1} | p'_d) p(y^{t+1} | p'_s) \\ p(x^{t+1} | p'_d) &= p'_d f_1(x^{t+1}) + (1 - p'_d) f_0(x^{t+1}) \\ p(y^{t+1} | p'_s) &= (p'_z + (1 - p'_z) h(t + 1)) g_1(y^{t+1}) \\ &\quad + (1 - p'_z) (1 - h(t + 1)) g_0(y^{t+1}) \end{aligned}$$

In the simulations, we discretize the space of  $p'_d$  and  $p'_z$  each into 200 bins.

## Results: Model Simulations

Fig. 2A shows the simulated evolution of belief state in the model for different trial types: (1) go trial (Go), where no stop signal appears, (2) stop success trial (SSS), where a stop signal is successfully processed by taking an explicit stop action, (3) stop error trial (SE), where a go response is made on a stop trial. Similar to (Shenoy & Yu, 2011), in SSS trials, the go stimulus happens to be processed slowly while the stop signal is being processed quickly, thus leading to successful stopping; conversely, on SE trials, the go stimulus happens to be processed quickly while the stop signal is being processed slowly. Note that the difference in these belief state trajectories across SSS and SE trials is solely due to sensory noise in the observation generation process.

Fig. 2B shows the simulated evolution of different Q-factors, or the expected cost of taking the three actions (Go,

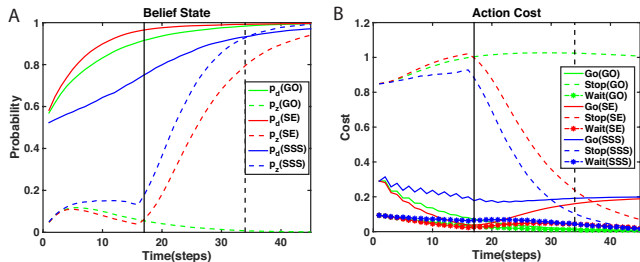


Figure 2: Mean Belief state and Q-factors. (A) Evolution of the average belief states  $p_d$  (solid line) and  $p_z$  (dashed line) for different trials-Go(green):go trials, SSS(blue): stop trials successfully stopped by choosing explicit stop action, SE (red): stop error trials. In SE trials, the go stimulus is processed faster than the stop signal, but the converse in SSS trials. We also assume that  $d = 1$  for all trials in the figure for the purpose of simplicity. The onset of stop signal is  $\theta_s = 17$  time steps(solid vertical line). The dashed vertical line represents the SSRT with current parameters. (B) Go(solid line), Stop(dashed line) and Wait(dotted line) cost for the same classification of trials. In SSS trials, the go costs significantly overpass the wait cost until the stop cost drops below the wait cost. In contrast to SSS, SE trials shows a rapidly decreasing go cost which dips below the wait cost before the stop cost decreases sufficiently, leading to stop errors. Although the average go cost never falls below the average wait cost, each individual trajectory will cross over at different time due to the stochasticity of observations. We adopt most of the parameters used in (Shenoy & Yu, 2011):  $q_d = 0.68$ ,  $q_s = 0.72$ ,  $q = 0.2$ ,  $r = 0.25$ ,  $D = 50$  steps,  $c_s = 0.2$ ,  $c_r = 0.002$ ,  $c_a = 0.002$ . Unless otherwise stated, these parameters are used in all the subsequent simulations.

Wait, Stop), over time on different trial types. In SSS trial, the *go* cost decreases slowly and never drops below the *wait* cost, while the *stop* cost drops rapidly after the onset of the stop signal and eventually below both the *wait* and *go* costs. The *go* cost in the SE trial shows the converse, dipping below the *wait* cost before the *stop* cost has decreased sufficiently ( $Q_g$  does not look like it dips below  $Q_w$  in the average trajectory, but it does do so on every individual trial, but at different moments, such that the average looks like it does not do so). In Go trials, the *stop* cost is large and continuously increasing, while the *go* cost is small and continuously decreases until it dips below  $Q_w$ .

## Results: Model Comparison to Behavioral Data

Here, we show that the model reproduces behavioral phenomena observed in relation to the stop signal task, including all those demonstrated by the earlier MDP model (Shenoy & Yu, 2011).

A classical behavioral phenomenon is that SE frequency increases in an approximately logistic fashion as a function of SSD (inhibition function), which is captured by both the

race model and the earlier MDP model (Shenoy & Yu, 2011); Fig. 3A shows that our model also captures this effect. Our model also capture the effect that the stop error decreases with the stop error penalty  $c_s$ , with the effect present at almost the whole range of SSD.

Additionally, it is known that subjects have slower Go reaction time (RT), lower SE rate, and faster SSRT when the relative of a stop error is increased via experimental design (Leotti & Wager, 2009), a phenomenon shown to naturally arise when  $c_s$  is increased in the earlier MDP model (Shenoy & Yu, 2011); Fig. 3B-D shows the new model also captures this. As  $c_s$ , the parameter specifying the stop error cost in the model, varies from low to high, our model simulation shows that (Fig. 3B) subjects can be expected to respond faster, (Fig. 3C) make fewer errors, and (Fig. 3D) have longer stopping latency (Stop RT). Note that Stop RT is analogous to SSRT in the race model, but instead of being estimated jointly from the Go RT distribution and the stop error rate as a function of SSD, Stop RT is computed directly by taking the difference between the Stop action time and the Go stimulus onset. The effects associated with Go RT and Stop error rate are generally robust for different setting of  $c_a$  and  $c_r$ .

## Neural Representation of action value

In this section, we show how internal computational components of the MDP model compares to neural responses observed in the frontal eye field (FEF) region of the monkey cortex during an oculomotor version of the stop-signal task (Hanes et al., 1998). FEF is known to be important for the planning and execution of eye movements and is under strong top-down cognitive control. It has two known subpopulations of neurons, "movement" neurons and "fixation" neurons, which are respectively more active on go and stop trials, and which have been postulated to be instantiating the go and stop processes in the race model (Hanes et al., 1998).

Fig. 4 A and B show the spike density function of fixation and movement neurons, respectively. The go (no-stop-signal) trials (thin solid lines) are latency-matched to canceled trials (thick solid lines) with saccade latencies that are long enough, e.g greater than the SSD + SSRT, such that they would have been canceled if a stop signal had been presented. During canceled stop (SS) trials Fig. 4A, the activity of fixation neuron is significantly enhanced after the onset of the stop signal and peaks around the SSRT, diverging from its weak response in go (no-stop-signal) trials. Fig. 4 B shows that the activity of movement neuron also diverge on go (no-stop-signal) trial as compared to SS (canceled) trial around SSRT. These neuron data imply that these neurons may encode computations leading to the cancellation and execution of the go response.

We hypothesize that fixation neurons may encode the explicit Stop action in our model, while movement neural activities may reflect the formation of the decision to Go. Fig. 4C shows the simulated distribution of Stop RT on successfully stopped trials in which the models takes an explicit stop action. Stop RT shows a peak right near the SSRT, which

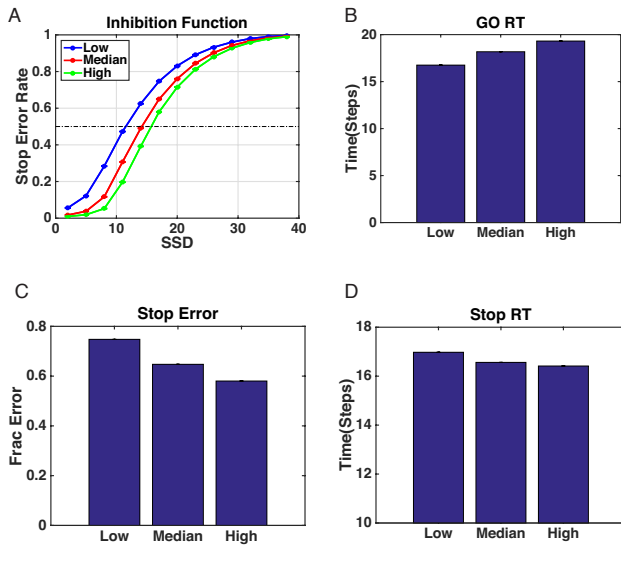


Figure 3: Influence of reward/motivation on stopping behavior. Results are averaged over 50 simulated subjects, each performing 10000 go and stop trials. Error bars (s.e.m.) are too small to be observed. (A) Model simulation produces an inhibition function (frequency of SE as a function of SSD) similar as that observed in behavioral data. Stop error (SE) decreases as the stop error penalty  $c_s$  increase (low = 0.2, median = 0.4, high = 0.6), with the effect present at almost the whole range of SSD. (B) When  $c_s$  is increased, the model responds slower in go trials, (C) make fewer stop errors (SSD = 17), and (D) exhibit shorter stopping latency (Stop RT). Stop RT is analogous to SSRT in the race model, here computed explicitly by taking the difference between the onset of stop signal and the time the decision maker chooses the *stop* action.

closely resembles the fixation neuron activity in canceled trial, implying that the fixation neurons may activate when an explicit stop action is chosen in SS trials. Fig. 4D shows the trajectories of  $1 - Q_g$ : the negative expected cost, or the expected reward, associated with the Go action in our model. The qualitative similarity between the expected Go reward and the activity of movement neurons suggests that the movement neuron may encode moment-by-moment estimate of Go action values.

## Discussion

In this work, we presented a novel Bayesian Markov Decision Process model of inhibitory control in the stop-signal task. The key difference between this model and our earlier Bayesian MDP model (Shenoy & Yu, 2011) is that the earlier model only allowed two actions, go and wait, with the stop response only implicitly realized when the observer repeatedly chooses the Wait action until the response period

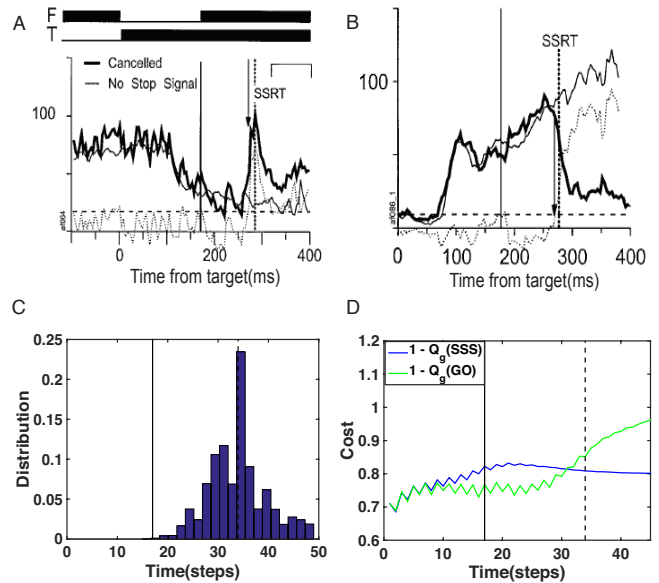


Figure 4: Neural representation of action values. (A) Average firing rate of a fixation neuron in the frontal eye field (FEF) of a monkey performing a saccade version of stop-signal task peaks at the time of the behaviorally measured SSRT (Hanes et al, 1998), suggesting fixation neurons encode a Stop action signal. (B) Movement neurons in the FEF diverge, between stop-success trials (Canceled) and go trials (No Stop Signal), at a time close to SSRT, suggesting participation in the decision process for the Go response. (C) In our model, the distribution of Stop RT in SS trials peaks near the SSRT as fixation neurons do, while (D) the expected reward ( $1 - \text{expected cost}$ ) of taking the Go action diverge around SSRT on SS/Go trials, much as movement neurons do.

expires. Here, we formally introduce an additional Stop action, the existence of which has long been postulated based on neuroimaging data in humans and neurophysiology data in monkeys. We also posited an extra attention cost associated with being engaged in the task, and which is spared when the stop action is taken. We showed that the new model can reproduce all the behavioral effects captured by the previous model. In addition, our model simulations indicate that previously observed activities of “movement” neurons in the monkey frontal eye field are consistent with their encoding the moment-by-moment valuation of the Go action, while the “fixation” neural activities may encode an explicit Stop action.

We can relate the present model to the classical race model, as the go action is typically chosen in our model when the expected go cost dips below the cost of waiting, and the stop action is typically chosen when the stop cost dips below the cost of waiting. Notably, the cost of waiting is fairly constant over the time course of the trial and stable across trial types (go, SS, SE), so that it can be thought of as the common

threshold that the go cost and stop costs race to reach first in order to determine the response outcome. However, this is not an independent race between go and stop processes, as originally envisioned in the race model (Logan & Cowan, 1984). Rather, it is closer to an interactive version of the race model that posits a late mutually inhibitory interaction between the stop and go processes (Boucher et al., 2007), except here, there is no direct antagonism between the two processes, but rather a common input (increasing sensory evidence that a stop signal is present) that drives the two processes in two opposite directions (suppressing the go process and accelerating the stop process). This leads to another notable difference between this model and the race model: the race model assumes the go process to be identical between go and stop trials, and uses that assumption to estimate the SSRT. In the present model, the go process is suppressed by the (late) onset of a stop signal, and thus the process splits into a bimodal distribution of termination times, such that there is an early mode that escapes the stop signal's suppressive influence and which ends in slightly faster average SE RT than Go RT, and a late mode that gets suppressed by the stop signal. Because of this, the true stopping latency can afford to be, and is indeed found to be, much later than the estimated SSRT. This may explain why the FEF neural response diverges between correct go and stop trials apparently too late to participate in executing the stop action (Hanes et al., 1998). This is because the race model's assumption of an unchanging go process (between go and stop trials) may be leading it to systematically under-estimate the true stopping latency. The current model would interpret FEF "fixation" neurons are signaling or relaying the decision to stop, while the FEF "movement" neurons are encoding the expected value of executing the go response.

One implication of the current work is that contextual changes in the attentional state of the observer, or the costs associated with paying attention to the task, should have systematic consequences on the observer's readiness and timing in executing a stop action. In particular, the model predicts that if the attention cost is raised, for example due to the presence of a dual task siphoning away cognitive resources, then the stop action should be chosen more readily, which would both have a behavioral consequence and be reflected in the neural dynamics. A productive line of future experimental work would be to test these predictions empirically by manipulating attentional costs. While the proposed model of inhibitory control, and the earlier MDP model that preceded it (Shenoy & Yu, 2011), may not be fully correct in describing the cognitive and neural processes underlying inhibitory control, they exemplify a powerful modeling framework for hypothesizing neural computations in the context of behaviorally defined goals and computations, which can then be tested experimentally by changing experimental conditions or behavioral objectives.

## Acknowledgments

We are grateful to Pradeep Shenoy and Joseph Schilz for helping to collect data in the stop-signal task, and to Pradeep Shenoy for preliminary discussions on how to model an explicit Stop action. Funding was provided in part by an NSF CRCNS grant (BCS- 1309346) to AJY.

## References

- Aron, A. R., Durston, S., Eagle, D. M., Logan, G. D., Stinear, C. M., & Stuphorn, V. (2007). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *Journal of Neuroscience*, *27*(44), 11860-4.
- Boucher, L., Palmeri, T. J., Logan, G. D., & Schall, J. D. (2007). Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychological Review*, *114*(2), 376-97.
- Hanes, D. P., Schall, J., & Patterson, W. F. (1998). The role of frontal eye field in countermanding saccades: visual, movement and fixation activity. *J. Neurophysiol.*, *79*, 817-34.
- Harlé, K. M., Shenoy, P., Steward, J. L., Tapert, S., Yu\*, A. J., & Paulus\*, M. P. (2014). Altered neural processing of the need to stop in young adults at risk for stimulant dependence. *Journal of Neuroscience*, *34*, 4567-4580. (\*Yu and Paulus are co-senior authors)
- Harlé, K. M., Steward, J. L., Zhang, S., Tapert, S., Yu\*, A. J., & Paulus\*, M. P. (2015). Bayesian neural adjustment of inhibitory control predicts emergence of problem stimulant use. *Brain*. (\*Yu and Paulus are co-senior authors)
- Ide, J., Hu, S., Zhang, S., Yu, A. J., & Li, C.-S. R. (2015). Impaired bayesian learning for cognitive control in cocaine dependence. *Drug and Alcohol Dependence*, *151*, 220-227.
- Ide, J. S., Shenoy, P., Yu\*, A. J., & Li\*, C.-S. R. (2013). Bayesian prediction and evaluation in the anterior cingulate cortex. *Journal of Neuroscience*, *33*, 2039-2047. (\*Co-senior authors)
- Leotti, L. A., & Wager, T. D. (2009). Motivational influences on response inhibition measures. *J. Exp. Psychol. Hum. Percept. Perform.*, *36*(2), 430-447.
- Logan, G., & Cowan, W. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psych. Rev.*, *91*, 295-327.
- Ma, N., & Yu, A. J. (2015). Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control. *Frontiers in Psychology*, *6*(1046). doi: 0.3389/fpsyg.2015.01046
- Ma, N., & Yu, A. J. (2016). Inseparability of go and stop in inhibitory control: Go stimulus discriminability affects stopping behavior. *Frontiers in Neuroscience*, *10*(54). doi: 10.3389/fnins.2016.00054
- Shenoy, P., & Yu, A. J. (2011). Rational decision-making in inhibitory control. *Frontiers in Human Neuroscience*. (doi: 10.3389/fnhum.2011.00048)