Response Time Variability in an Inhibitory Control Task Reflects Statistical Learning and Adaptive Decision-Making

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

Response time (RT) is an oft-used but "noisy" behavioral measure in psychology. Here, we combine modeling and psychophysics to examine the hypothesis that RT variability may reflect ongoing statistical learning and consequent adjustment of behavioral strategy. We utilize the stop-signal task, in which subjects respond to a go stimulus on each trial, unless instructed not to by a subsequent, rare stop signal. We model across-trial learning of stop signal frequency (P(stop)) and stop-signal onset time (SSD) with a Bayesian hidden Markov model, and within-trial decision-making as optimal stochastic control. The model predicts that RT should increase with expected P(stop) and SSD, a prediction borne out by our human data. Thus, it appears that humans continuously monitor environmental statistics and adjust behavioral strategy accordingly. More broadly, our approach exemplifies the use of "noisy" RT measures for extracting insights about cognitive and neural processing.

Keywords: Bayesian modeling, decision making, learning, response time, behavioral psychophysics

Introduction

Response time (RT) is an oft-reported behavioral measure in psychology and neuroscience studies. As RT can vary greatly across trials of apparently identical experimental conditions, average or median RT across many identical trials is typically used to examine how task performance or an internal speed-accuracy tradeoff might be affected by different experimental conditions. Separately, a specialized subfield of quantitative psychology has used not only the first-order statistics (e.g. mean and median) but also second-order (e.g. variance) and higher-order (e.g. skewness, kurtosis) statistics to make inferences about the cognitive or neural processes underlying behavior (Laming, 1968; Luce, 1986; Smith, 1995; Ratcliff & Rouder, 1998; Gold & Shadlen, 2002; Bogacz et al., 2006). In general, RT is considered a very noisy experimental measure, with single-trial responses yielding little useful information about the underlying mental processes.

In this work, we approach RT modeling from a different angle, attempting to capture trial-to-trial variability in RT as a consequence of statistically normative learning about environmental statistics and corresponding adaptations within an internal decision-making strategy. We focus on behavior in the stop-signal task (SST) (Logan & Cowan, 1984), a classical inhibitory control task, in which subjects respond to a go stimulus on each trial unless instructed to withhold their response by an infrequent stop signal that appears some time after the go stimulus (stop-signal delay; SSD). We model trial-by-trial behavior in SST, using a Bayesian hidden Markov model to capture across-trial learning of stop signal frequency (P(stop)) and onset asynchrony (SSD), and a rational decision-making control policy, which combines prior beliefs and sensory data to produce behavioral outputs under task-specific constraints/objectives, to model within-trial decision-making.

This work builds on several previous lines of modeling research. The new model combines a *within-trial* rational decision-making model for stopping behavior (Shenoy & Yu, 2011) and an *across-trial* statistical learning model (Dynamic Belief Model; DBM) that sequentially updates beliefs about P(stop) (Yu & Cohen, 2009; Ide, Shenoy, Yu*, & Li*, 2013); it also incorporates a novel across-trial learning component, essentially a Kalman filter, that updates beliefs about the temporal statistics of the stop-signal onset (SSD). Using this new model, we can then predict how RT on each trial *ought* to vary as a function of the sequence of stop/go trials and SSD's previously experienced by the subject, and compare it to the subject's actual RT.

Several key elements of the combined model have previously received empirical support. For example, we showed that the rational decision-making model for stopping behavior (Shenoy & Yu, 2011), which separately penalizes stop error, go (discrimination and omission) error, and response delay, can account for both classical effects in the SST (Logan & Cowan, 1984), such as an increase in rate of stop errors as a function of SSD and the faster stop-error responses (relative to correct go responses), as well as some recently discovered, subtle influences on stopping behavior by contextual factors, such as motivation/reward (Leotti & Wager, 2009) and the baseline frequency of stop trials (Emeric et al., 2007). We also showed that the across-trial learning model, DBM, can account for sequential adjustment effects not only in SST (Ide et al., 2013), but also more broadly in simple 2AFC perceptual decision-making tasks (Yu & Cohen, 2009) and a multi-target visual search task (Yu & Huang, 2014). Neurally, we have evidence from fMRI studies that a key prediction error signal related P(stop) is encoded in the brain region known as dorsal anterior cingulate cortex (dACC) (Ide et al., 2013), and that dACC response is altered in young adults at-risk for developing stimulant addiction (Harlé et al., 2014).

In the following, we first describe the experimental design, then the modeling details, followed by the results; we finally conclude with a discussion of broader implications and future directions for research.

Experiment

22 UCSD students participated in the stop signal task. On each trial, the subject was presented with a two alternative forced-choice (2AFC) perceptual discrimination task, in which he must press a left or right arrow depending on whether the stimulus was a square or circle or the direction of random dot motion(stimulus-key association was counterbalanced across subjects). On approximate 25% of trials, an auditory "stop" signal was presented some time after the go (discrimination) stimulus, indicating that the subject should withhold their response to the go stimulus. Trials containing a stop signal are stop trials; otherwise they are go trials. The delay in presentation between the go stimulus and the the stop signal is known as the stop-signal delay (SSD), which was uniformly and randomly sampled from 100 ms, 200 ms, 300 ms, 400 ms, 500 ms, and 600 ms in stop trials. Each subject participated in 12 blocks, with each block containing 75 trials.

Models

In this section, we give a brief description of the three computational models: a stochastic control model for within-trial sensory processing and decision-making, a hidden Markov model (DBM) for across-trial learning of stop signal frequency, and a Kalman filter model for across-trial learning of SSD.

Stochastic Control Model for Within-Trial Processing

We briefly summarize the rational decision-making model for stopping behavior here; a more detailed description can be found elsewhere (Shenoy & Yu, 2011).

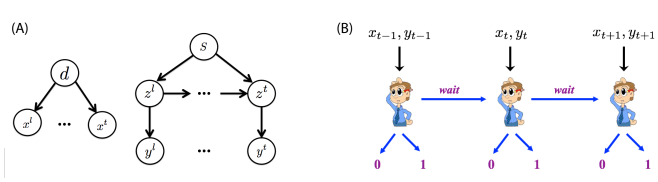

Figure 1. Within-trial sensory processing and decision-making. (A) Bayesian generative model of iid sampled sensory observations $(x^1,\ldots,x^t,\ldots)$ conditioned on Go stimulus identity ($d=0$, $d=1$), and an independent stream of observations $(y^1,\ldots,y^t,\ldots)$ conditioned on the presence ($z^t=1$) or absence ($z^t=0$) of the Stop signal. (B) The decision of whether to Go, when to do so, and which Go response to select are modeled as a sequential decision-making process, where the subject chooses at each moment in time whether to select a Go response ($\delta=0$ for square, $\delta=1$ for circle), or to wait at least one more time point.

Sensory processing as Bayes statistical inference. Figure 1A graphically illustrates the Bayesian generative model, whereby the two hidden variables correspond respectively to the identity of the go stimulus, $d\in\{0,1\}$, and whether or not this trial is stop trial, $s\in\{0,1\}$. The priors of d and s are $P(d=1)$ and $r=P(s=1)$. Conditioned on the go stimulus identity d, a sequence of iid sensory inputs are generated on each trial, x^1, ... ,x^t, ... ,where t indexes time steps *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 respective rate rate parameters q_d and $1-q_d$. The dynamic variable z^t denotes the presence/absence of the stop signal. $z^1=...=z^{\theta-1}=0$ and $z^\theta=z^{\theta+1}=...=1$ if a stop signal appears at time θ, where θ represents stop signal delay SSD. For simplicity, we assume that the onset of the stop signal θ (SSD) follows an geometric distribution: $p(\theta=t|s=1)=q(1-q)^t$. The mean of θ is equal to $\frac{1}{q}$ which is the expected SSD ($\mathbb{E}[SSD]$) within a trial. Conditioned on z^t, a stream of iid observations are generated. 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 assume that the likelihood functions, g_0 and g_1, are also Bernoulli distributions with respective rate parameters q_s and $1-q_s$.

In Bayesian statistical inference process, Bayes' Rule is applied in the usual iterative manner way to compute the sequential posterior probability associated with go stimulus identity, $p_d^t := P(d=1|\mathbf{x}^t)$, and the presence of the stop signal, $p_s^t := P(s=1|\mathbf{x}^t)$, where $\mathbf{x}^t=\{x^1,x^2,...,x^t\}$ denotes all the data observed so far. $p_z^t := P(\theta<t|\mathbf{y}^t)$ denotes the posterior probability that the stop signal is already present. The *belief state* at time t is defined to be the vector $\mathbf{b}^t=(p_d^t,p_s^t)$, which can be iteratively computed from step to step via Bayes' Rule, by inverting the generative model (Figure 1).

Decision process as optimal stochastic control. Figure 1B graphically illustrates the sequential decision-making process. On Go trials, if the Go action is taken by the response deadline D, it is recorded as a Go response (correct on Go trials, stop error on Stop trials); otherwise the trial terminates and a Stop response is recorded (omission error on Go trials, correct on Stop trials). We define a cost (loss) function to account for the cost and penalty structure of the stop-signal task. The observer is assumed to minimize the expected value of this loss function in choosing whether to Go or Wait at each tilmestep, based on the current belief state. A Go response terminates the current trial, while a Wait response lengthens the current trial by at least one more time step (unless terminated by the externally imposed response deadline).

Let τ denote the trial termination time, so that $\tau=D$ if no response is made before the deadline D, and $\tau<D$ if a Go action is chosen. $\delta\in\{0,1\}$ represents the possible binary Go choices produced by making a Go response. We also assume there is a basic cost c per unit time on each trial, a stop error penalty of c_s for choosing to respond on a stop-signal trial, and a unit cost for making a discrimination error on a go trial

(since the cost function is invariant with respect to scaling, we can normalize one of the cost parameters to 1 without loss of generality). We assume the cost function to be:

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

The optimal decision policy minimizes the expected loss, $L_\pi = \mathbb{E}[l(\tau,\delta;d,s,\theta,D)]$. It is computationally intractable to directly minimize L_π 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

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

where a ranges over all possible actions. The optimal policy is to choose the action corresponding to the smallest action cost. Using the Bellman optimality equation for optimal action-value function, we can obtain the cost functions

$$\begin{aligned} Q_g^t(b^t) &= ct + c_s p_s^t + (1-p_s^t)\min(p_d^t, 1-p_d^t) \\ Q_w^t(b^t) &= \mathbf{1}_{\{D>t+1\}}\mathbb{E}\left[V^{t+1}(b^{t+1})|b^t\right]_{b^t} + \mathbf{1}_{\{D=t+1\}}(1-P_s^t) \\ V^t(\mathbf{b}^t) &= \min(Q_g^t, Q_w^t) \end{aligned}$$

Since the observer 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 $V^t(\mathrm{b}^D) = 1 - p_s^D$. The recursive relationship between the optimal action-value and state-value functions in Bellman optimality equation allows us to the compute the optimal state-value functions and Q factors *backwards in time* from $t = D-1$ to $t = 1$. In our simulation, we discretize the space of p_d^t and p_z^t each into 200 bins.

With this model, the mean Go RT can be obtained by simulating the model for a certain parameter setting (variability in outcome arises entirely from assumed observation noise, parameterized by q_d and q_s, thus providing a way to study how the Go RT is related to expectation of environmental statistics, e.g. P(stop) and $\mathbb{E}[SSD]$. $\mathbb{E}[SSD]$ is the mean of geometrically distributed prior of SSD ($\mathbb{E}[SSD] = \frac{1}{q}$) and P(stop) (P(stop) = r) represents the expectation of the probability of seeing a stop trial.

Dynamic Belief Model for learning P(stop)

We use a previously proposed Bayesian hidden Markov model, the Dynamic Belief Model (DBM) (Yu & Cohen, 2009; Ide et al., 2013), to model trial-by-trial evolution of prior (and posterior) beliefs about P(stop). We briefly describe the model here; more details can be found elsewhere (Yu & Cohen, 2009; Ide et al., 2013).

Dynamic Belief Model (DBM) was proposed to explain *sequential effects* in reaction time and accuracy in 2AFC tasks, as a function of experienced trial history (Yu & Cohen, 2009). In DBM, γ_k is the probability the subject will see a stop trial at time step k and has a Markovian dependence on γ_{k-1}, so that with probability α, $\gamma_k = \gamma_{k-1}$, and probability $1-\alpha$, γ_k is redrawn from a fixed Beta distribution $p_0(\gamma_k)$. The observation s_k represents the occurrence of a stop trial and is assumed to be drawn from a Bernoulli distribution with parameter γ_k. The predicted value of γ_k is the mean of its posterior: $<\gamma_k|\mathbf{s}_{k-1}> = \int \gamma p(\gamma|\mathbf{s}_{t-1})d\gamma$. The posterior and iterative prior of γ_k can be updated by

$$\begin{aligned} p(\gamma|\mathbf{s}_{k-1}) &= \alpha p(\gamma_{k-1} = \gamma|\mathbf{s}_{k-1}) + (1-\alpha)p_0(\gamma_k = \gamma) \\ p(\gamma_k|\mathbf{s}_k) &\propto p(\gamma|\mathbf{s}_{k-1}) \end{aligned}$$

We adapt DBM to model the prior probability of observing a Stop trial (as opposed to Go trial) based on trial history (see Figure 2A for a graphical illustration of the generative model, and Figure 2B for simulated dynamics of DBM given a sequence of sample observations). We briefly describe the model here; more details can be found elsewhere (Yu & Cohen, 2009; Ide et al., 2013).

Kalman Filter Model for Learning Expected SSD

We use a classical linear-Gaussian dynamical systems model, otherwise known as a Kalman Filter (Welch & Bishop, 2006), to model the trial-by-trial estimation of the mean and variance of SSD in the stop-signal task. As shown in Figure 2C, Kalman Filter tries to estimate the state h of a discrete-time controlled process governed by the linear stochastic equation

$$h_k = Ah_{k-1} + Bu_{k-1} + w_{k-1}$$

with a measurement z which is

$$z_k = Hh_k + v_k$$

The random variable w_k and v_k represent the process and measurement noise, respectively. They are assumed to be independent, white and with normal distribution

$$\begin{aligned} p(w) &\sim N(0,Q) \\ p(v) &= N(0,R) \end{aligned}$$

The equation for the Kalman filter consists of two parts: time update equations and measurement update equation. The time update equation obtains a prior estimate from the next time step. $\hat{h}_k^-$ is defined as a priori estimate at step k given information before step k ,and $\hat{h}_k$ as a posteriori estimate at step k given the measurement z_k. $\hat{P}_k^-$ is defined to be a priori estimate of error covariance and $\hat{P}_k$ to be a posterior estimate.

$$\begin{aligned} \hat{h}_k^- &= A\hat{h}_{k-1} + Bu_{k-1} \\ P_k^- &= AP_{k-1}A^T + Q \end{aligned}$$

The measurement update equations incorporates a new measurement into the priori estimate to obtain an improved a posterior estimate of the state.

$$\begin{aligned} K_k &= P_k^- H^T(HP_k^- H^T + R)^{-1} \\ \hat{h}_k &= \hat{h}_k^- + K_k(z_k - H\hat{h}_k^-) \\ P_k^- &= (I - K_k H)P_k^- \end{aligned}$$

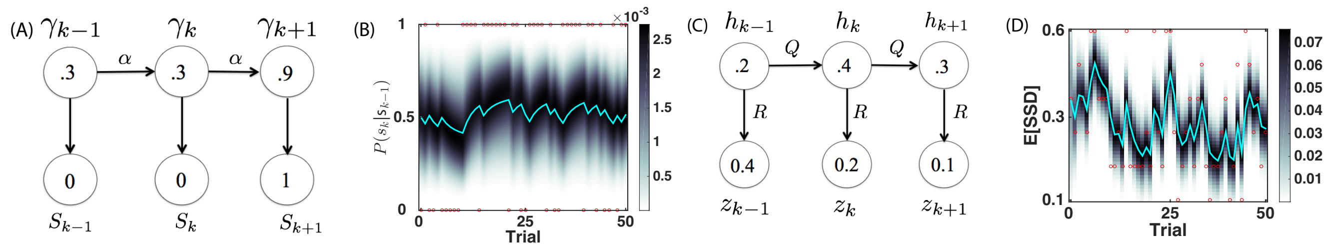

Figure 2. Bayesian sequential inference model for learning $P(stop)$ and $\mathbb{E}[SSD]$. (A) Graphical model for DBM. $\gamma \in [0,1]$, $s_k \in \{0,1\}$. $p(\gamma_k|\gamma_{k-1}) = \alpha\delta(\gamma_k - \gamma_{k-1}) + (1-\alpha)p_0(\gamma_k)$, where $p_0 = Beta(a,b)$. Numbers inside circles indicate example random variable values. (B) Evolution of predictive probability mass for DBM $p(\gamma_t|\mathbf{s}_{k-1})$ (grayscale) and its mean, the predictive probability $P(s_k = 1|\mathbf{s}_{k-1})$ (cyan), for a randomly generated sample sequence of observations (red dots valued 1 or 0). $P(s_k = 1|\mathbf{s}_{k-1})$ fluctuates with transient runs of stop (e.g. starting at trial 11) and go trials (e.g. starting at trial 6). Simulation parameters: α = 0.75, $p_0 = Beta(2.5, 7.5)$. (C) Graphical model for the Kalman filter. $p(h_k|h_{k-1}) = \mathcal{N}(h_{k-1}, Q)$, $p(z_k|h_k) = \mathcal{N}(h_k, R)$, $p(h_1) = \mathcal{N}(h_0, P_0)$. Numbers inside circles indicate example random variable values. (D) Evolution of posterior mean (cyan) and probability mass (grayscale) of SSD over time, for a randomly generated sequence of observations (red circles) with values in $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$. $\mathbb{E}[SSD]$ tends to increase when a number of large SSD have been observed (e.g. starting at trial 6) and decrease when a number of small SSD (e.g. starting at trial 11) have been observed. Simulation parameters: Q = 0.034, R = 0.15, $h_0 = 0.35$, $P_0 = 1$. Unless otherwise stated, these parameters are used in all the subsequent simulation.

In this paper, we use a simple Kalman Filter model, where $B = 0$, $A = 1$ and $H = 1$, to compute the $\mathbb{E}[SSD]$ ($\hat{h}_k^-$) by incorporating real observed SSD (z_k) to the filter. Figure 2D shows simulated dynamics of the Kalman filter given a sequence of sample observations.

Results

Our main modeling goal here is to develop a principled explanation for how Go RT *ought* to vary from trial-to-trial in the stop-signal task. We can then compare model predictions with human data to see whether our assumptions about the underlying computational processes hold. There are two key components to the model: (1) how subjects' beliefs about task statistics vary across trials as a function of previously experienced outcomes, and (2) how subjects' behavioral strategy within each trial depends on prior beliefs (learned from prior experience). For the first component, we separately model the evolution of subjects' beliefs about the frequency of stop trials, P(stop), using a Bayesian hidden Markov model known as the Dynamic Belief Model (DBM), and their beliefs about the temporal onset of the stop signal, stop-signal delay (SSD), using a Kalman Filter model. For the second component, we use an optimal stochastic control model to predict when and whether the subject produces a GO response on each trial, as a function of prior beliefs about P(stop) and SSD, dynamically evolving posterior beliefs about the type of Go stimulus and presence of stop signal, as well as the relative cost of GO now and WAIT another time-step depending on the expected costs associated with making a go (discrimination or omission) error, a stop error (not stopping on a stop trial), and response delay (Shenoy & Yu, 2011) . Details of the model can be found in the Models section.

We first simulate the stochastic control model to demonstrate how Go RT ought to vary as a function of prior beliefs about P(stop) and SSD. Intuitively, we would expect that Go RT ought to increase as prior P(stop), since the higher probability of encountering a stop signal should make the subject more willing to wait for the stop signal despite the cost associated with response delay. We also expect that Go RT ought to increase with $\mathbb{E}[SSD]$ for the prior distribution, since expectation of an earlier SSD should give confidence to the observer that no stop signal is likely to come after a shorter amount of observations and thus respond earlier. Simulations of the stochastic control model (Figure 3) shows that Go RT indeed increases monotonically with both P(stop) and $\mathbb{E}[SSD]$, and does so linearly. Note that P(stop) and $\mathbb{E}[SSD]$ are specified explicitly in the statistical model here (details in the Models section), so we only need to change these parameters and observe their normative consequences by running the stochastic control model.

Next, we want to examine how subjects' actual Go RT varies with prior beliefs about P(stop) and SSD. Since the experiment does not explicitly manipulate the baseline frequency of stop signals or SSD, we *estimate* these psychological quantities by assuming that subjects continuously modify their prior beliefs according to experienced trial history. We apply the across-trial learning models to a subject's experienced sequence of go/stop trials and SSD to estimate their priors on each trial, and then plot how Go RT varies with the model-based estimates of P(stop) and $\mathbb{E}[SSD]$ Figure 4 shows that the subjects' Go RT increases approximately linearly with prior P(stop) and $\mathbb{E}[SSD]$, just as predicted by the model (Figure 3). These result imply that subjects both continuously monitor and update internal representations about environmental statistics, and adjust their behavioral strategy rationally according to those evolving representations.

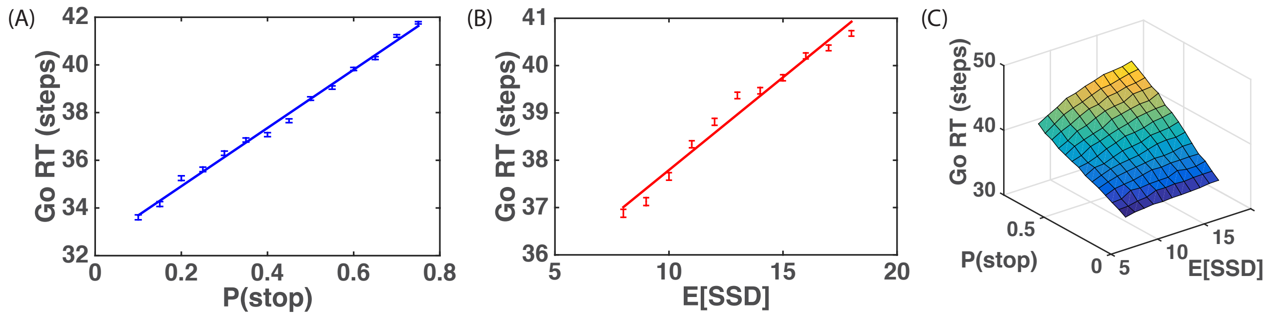

Figure 3. Model prediction of Go RT versus P(stop) and $\mathbb{E}[SSD]$. (A) Go RT vs. P(stop): simulated Go RT for a ranged of P(stop) values (.1, .15, ..., .75). Data averaged over 10000 simulated Go trials for each value of P(stop). Straight line denotes best linear fit. Error bars denote s.e.m. $1/q = \mathbb{E}[SSD] = 7$. (B) Go RT vs. $\mathbb{E}[SSD]$: simulated Go RT for a range of SSD values (8, 9, ..., 18). Data averaged over 10000 simulated Go trials for each value of $\mathbb{E}[SSD]$. Straight line denotes best linear fit. Error bars denote s.e.m. P(stop) $= 0.25$. (C) Go RT vs. P(stop) and $\mathbb{E}[SSD]$: simulated Go RT for a range of P(stop) and $\mathbb{E}[SSD]$) values, where P(stop) varies between .1 and .75, and $\mathbb{E}[SSD]$ varies between 8 and 18. Data averaged over 10000 simulated Go trials for each (P(stop), $\mathbb{E}[SSD]$). Simulation parameters for A-C: $q_d = 0.55$, $q_s = 0.72$, $D = 50$, $c_s = 0.4$, $c = 0.002$. Initial string of Go trials in each block (on average 3 trials, 1/4 time none at all) are excluded from all analyses, as subjects' initial beliefs about task statistics may vary widely and unpredictably before any stop trials are observed.

Discussion

In this paper, we presented a rational inference, learning, and decision-making model of inhibitory control, which can account for significant variability of human RT in the stop-signal task. Unlike most previous models of human response time, which assumes RT variability to be due to irreducible noise, we conclude that this variability is largely driven by fluctuations in experienced empirical statistics, which observers use to continuously update their internal representation of environmental statistics and rationally adjust their behavioral strategy as needed.

The reader may well wonder why we choose to use a different model for capturing sequential effects in SSD (Kalman Filter) than for P(stop) (DBM). The Kalman Filter primarily differs from DBM in that the hidden variable s is assumed to undergo (noisy) continuous dynamics, such that the mean of the new variable is centered at the old s (it is a Martingale process), whereas DBM assumes that the new hidden variable s is either identical to its value at the last time step, or redrawn from a generic prior distribution $p_0(s)$, which is identical on each trial. This means that hidden variables dynamics in DBM are not Martingale, and the variable s can readily undergo large, discrete jumps, which are not likely in the Kalman Filter. We used both the Kalman Filter and an adapted version of DBM to model subjects' beliefs about $\mathbb{E}[SSD]$, and found that the Kalman Filter does a significantly better of accounting for trial-by-trial variability in RT than does DBM (data not shown).

The work is also important for making a substantial contribution in advancing the understanding of inhibitory control. Inhibitory control, the ability to dynamically modify or cancel planned actions according to ongoing sensory processing and changing task demands, is considered a fundamental component of flexible cognitive control (Barkley, 1997; Nigg, 2000). Stopping behavior is also known to be impaired in a number of psychiatric populations with presumed inhibitory deficits, such as attention-deficit hyperactivity disorder (Alderson, Rapport, & Kofler, 2007), substance abuse (Nigg et al., 2006), and obsessive-compulsive disorder (Menzies et al., 2007). The work shown here elucidates the psychological and potential neural underpinnings of inhibitory control, and makes powerful predictions that can be validated using experimental methods. For example, this work can investigate how different individuals' ability to represent and respond to those expectations, e.g. P(stop) and $\mathbb{E}[SSD]$, may be correlated with eventual development/absence of abusive stimulant behavior, as we have already done in some previous work in collaboration with neuroimaging researchers (Ide et al., 2013; Harlé et al., 2014).

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 subjects' evolving beliefs about SSD.

References

Alderson, R. M., Rapport, M. D., & Kofler, M. J. (2007). Attention-deficit hyperactivity disorder and behavioral inhibition: A meta-analytic review of the stop-signal paradigm. *Journal of Abnormal Child Psychology*, *35*(5), 745–58.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological bulletin*, *121*(1), 65-94.

Bogacz, R., Brown, E., Moehlis, J., Hu, P., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in

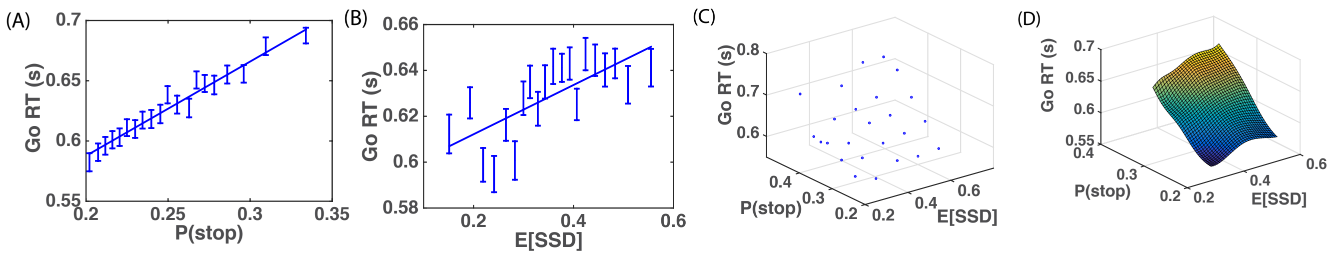

Figure 4. Human Go RT versus model-estimated P(stop) and SSD. (A) Go RT vs. P(stop): P(stop) on each trial estimated by DBM based on actual sequence of stop/go trials the subject experienced prior to the current trial. Binning of $\mathbb{E}[SSD]$ spaced to ensure equal number of data points in each bin. Straight line denotes best linear fit of average Go RT for each bin versus average P(stop) for each bin. Linear regression of group data: $R^2 = 0.97$, $p < 0.0001$. Error bars denote s.e.m. DBM parameters: $\alpha = 0.75$, $p_0 = Beta(2.5, 7.5)$. (B) Go RT vs. $\mathbb{E}[SSD]$: $\mathbb{E}[SSD]$ on each trial estimated by a Kalman filter based on actual sequence of SSD the subject experienced prior to the current trial. Binning of $\mathbb{E}[SSD]$ spaced to ensure equal number of data points in each bin. Straight line denotes best linear fit between average Go RT versus average $\mathbb{E}[SSD]$ for each bin. Linear regression of group data: $R^2 = 0.53$, $p = 0.0003$. Error bars denote s.e.m. Kalman filter (KF) parameters: $Q = 0.034$, $R = 0.15$, $h_0 = .35$, $P_0 = 1$. (C) Go RT vs. P(stop) and $\mathbb{E}[SSD]$: P(stop) and $\mathbb{E}[SSD]$ are equally discretized into 5 bins between minimum and maximum "observed" values (by applying the model to subjects' experienced sequence of trials). Each point in the grid contains RT data from all trials and all subjects where P(stop) and $\mathbb{E}[SSD]$ fell within corresponding bins. (D) Fitted surface plot of the scatter plot in (C), by applying Matlab function $griddata(\ldots,' v4')$, a biharmonic spline interpolation method, to the data in (C).

two-alternative forced choice tasks. *Psychological Review*, *113*(4), 700-65.

Emeric, E. E., Brown, J. W., Boucher, L., Carpenter, R. H. S., Hanes, D. P., Harris, R., … others (2007). Influence of history on saccade countermanding performance in humans and macaque monkeys. *Vision Research*, *47*(1), 35-49.

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

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)

Ide, J. S., Shenoy, P., Yu*, A. J., & Li*, C.-R. (2013). Bayesian prediction and evaluation in the anterior cingulate cortex. *Journal of Neuroscience*, *33*, 2039-2047. (*Yu and Li contributed equally as senior authors)

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

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.

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

Menzies, L., Achard, S., Chamberlain, S. R., Fineberg, N., Chen, C. H., del Campo, N., … Bullmore, E. (2007). Neurocognitive endophenotypes of obsessive-compulsive disorder. *Brain*, *130*(12), 3223-36.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*(2), 220-46.

Nigg, J. T., Wong, M. M., Martel, M. M., Jester, J. M., Puttler, L. I., Glass, J. M., … Zucker, R. A. (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*(4), 468-75.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347-56.

Shenoy, P., & Yu, A. J. (2011). Rational decision-making in inhibitory control. *Frontiers in Human Neuroscience*. (doi: 10.3389/fnhum.2011.00048)

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychol. Rev.*, *10*, 567-93.

Welch, G., & Bishop, G. (2006). An introduction to the kalman filter.

Yu, A. J., & Cohen, J. D. (2009). Sequential effects: Superstition or rational behavior? *Advances in Neural Information Processing Systems*, *21*, 1873-80.

Yu, A. J., & Huang, H. (2014). Maximizing masquerading as matching: Statistical learning and decision-making in choice behavior. *Decision*, *1*(4), 275-287.