

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Human Learning and Decision-Making, and Their Applications

Permalink

<https://escholarship.org/uc/item/18f6r7t3>

Author

Ma, Ning

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Human Learning and Decision-Making, and Their Applications

A Dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Electrical Engineering (Intelligent Systems, Robotics, and Control)

by

Ning Ma

Committee in charge:

Professor Angela Yu, Chair
Professor Ken Kreutz-Delgado, Co-Chair
Professor Vikash Gilja
Professor Tara Javidi
Professor John Serences
Professor Nuno Vasconcelos

2017

Copyright
Ning Ma, 2017
All rights reserved.

The Dissertation of Ning Ma is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2017

DEDICATION

To my dear son — Mason Milo Ma

EPIGRAPH

The value of achievement lies in the achieving.

—Albert Einstein

TABLE OF CONTENTS

Signature Page		iii
Dedication		iv
Epigraph		v
Table of Contents		vi
List of Figures		viii
List of Tables		x
Acknowledgements		xi
Vita		xiii
Abstract of the Dissertation		xiv
Chapter 1	Introduction	1
Chapter 2	Statistical Learning and Adaptive Decision-Making underlie Human Response Time Variability in Inhibitory Control	4
	2.1 Introduction	5
	2.2 Materials and Methods	8
	2.2.1 Experiment	8
	2.2.2 Model	9
	2.3 Results	17
	2.4 Discussion	23
Chapter 3	Inseparability of Go and Stop in Inhibitory Control: Go Stimulus Discriminability Affects Stopping Behavior	33
	3.1 Introduction	34
	3.2 Materials and Methods	36
	3.2.1 Experiment	36
	3.2.2 Model	39
	3.3 Results	46
	3.3.1 Model Predictions	46
	3.3.2 Human Behavioral Data	48
	3.4 Discussion	51

Chapter 4	Stop Paying Attention: The Need for Explicit Stopping in Inhibitory Control	57
	4.1 Introduction	58
	4.2 The Model	61
	4.2.1 Monitoring Process as Bayesian Statistical Inference	61
	4.2.2 Decision Process as Optimal Stochastic Control . .	63
	4.3 Results	66
	4.3.1 Model Simulations	66
	4.3.2 Model Comparison to Behavioral Data	67
	4.3.3 Neural Representation of Action Value	68
	4.4 Discussion	69
Chapter 5	Reduced Neural Recruitment for Bayesian Adjustment of Inhibitory Control in Methamphetamine Dependence	76
	5.1 Introduction	77
	5.2 Methods and Materials	79
	5.2.1 Participants	79
	5.2.2 Stop-Signal Task	80
	5.2.3 Bayesian Model of Probabilistic Prediction	81
	5.2.4 Behavioral Statistical Analyses	83
	5.2.5 fMRI Analyses	84
	5.3 Results	86
	5.3.1 Subject Characteristics	86
	5.3.2 Behavioral Performance and Model-Based Behavioral Adjustment	87
	5.3.3 fMRI Analyses	89
	5.4 Discussion	92
Bibliography	103

LIST OF FIGURES

Figure 2.1:	Schematic illustration of our stop signal task in (A) Go trials and (B) Stop trials.	27
Figure 2.2:	Within-trial sensory processing and decision-making. (A) Bayesian generative model of sensory processing. (B) Decision making as optimal stochastic control.	28
Figure 2.3:	Bayesian sequential inference model for learning $P(stop)$ and $\mathbb{E}[SSD]$. (A)-(B) Graphical model and stimulated dynamics of Dynamic Belief Model. (C)-(D) Graphical model and simulated dynamics of Kalman filter.	29
Figure 2.4:	Sequential effects in human data. (A) and (C) Go RT increases with frequency and recency of stop trials, and experienced SSD. (B) Predicted $P(stop)$ increases with frequency and recency of stop trials	30
Figure 2.5:	Model prediction of Go RT versus $P(stop)$ and $\mathbb{E}[SSD]$	31
Figure 2.6:	Human Go RT versus model-estimated $P(stop)$ and SSD.	32
Figure 3.1:	Schematic illustration of stop signal task in (A) Go trials and (B) Stop trials	38
Figure 3.2:	(A) The classical race model. (B) Bayesian generative model of sensory processing. (C) Decision making as optimal stochastic control.	40
Figure 3.3:	Dynamic evolution of the optimal policy map at time step (A) 30, (B) 46, and (C) 48	46
Figure 3.4:	Simulated behavioral predictions by the optimal decision-making model versus go stimulus coherence. (A) mean Go RT; (B) Discrimination error; (C) Omission error; (D) Stop error; (E) Inhibition function; (F) Stop signal reaction time (SSRT).	49
Figure 3.5:	Human behavior data versus go stimulus coherence. (A) mean Go RT; (B) Discrimination error; (C) Omission error; (D) Stop error; (E) Inhibition function; (F) Stop signal reaction time (SSRT).	52
Figure 4.1:	Models for inhibitory control in the stop-signal task. (A) The classical race model. (B) Bayesian graphical model of sensory processing. (C) Decision making as optimal stochastic control.	59
Figure 4.2:	Mean Belief state and Q-factors. Evolution of the average belief states (A) and cost (B) for different trials.	73
Figure 4.3:	Influence of reward/motivation on stopping behavior. (A) Inhibition function; (B) Go RT; (C) stop error; (D) Stop RT.	74
Figure 4.4:	Neural representation of action values. (A) Average firing rate of a fixation neuron in the frontal eye field (FEF) of a monkey. (B) Movement neurons in the FEF diverge. (C) - (D) model simulation resembles the neural activity shown in (A) and (B).	75

Figure 5.1:	(A) Stop signal task. (B) Bayesian hidden Markov model.	97
Figure 5.2:	Behavioral comparison between comparison subjects (CSs) and methamphetamine-dependent individuals (MDI).	98
Figure 5.3:	Group difference in sensitivity to surprising trials, that is, in activation to Bayesian unsigned prediction error (UPE), in the left posterior caudate.	99
Figure 5.4:	Group difference in sensitivity to surprising trials, that is, in neural activation to a Bayesian unsigned prediction error (UPE) in the left middle frontal gyrus (Brodmann area 11).	100
Figure 5.5:	Group difference in the modulation of neural activation correlated with P(stop) by inhibitory success (right thalamus).	101
Figure 5.6:	Group difference in the modulation of neural activation correlated with P(stop) by inhibitory success (right inferior parietal lobule (IPL); Brodmann area 40).	102

LIST OF TABLES

Table 5.1:	Participants' Characteristics as a Function of Group Status ($n = 96$).	87
Table 5.2:	BOLD Activation Foci for Group by P(Stop)-Modulated Trial Type (Go vs. Stop) Interaction.	90
Table 5.3:	BOLD Activation Foci for Group by P(Stop)-Modulated Stop Trial Type (SS vs. SE) Interaction.	91

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Professor Angela J. Yu, who is extremely supportive and encouraging during my Ph.D study. Angela is more than my academic advisor. She teaches me not only how to conduct rigorous research but also how to be successful in the professional world. Her mentorship has shaped my life academically, personally, socially and beyond.

I would like to thank my committee members: Professor Ken Kreutz-Delgado, Professor Vikash Gilja, Professor Tara Javidi, Professor John Serences, and Professor Nuno Vasconcelos, for their stimulating feedback over my research problems, and for helping me improve the thesis.

I would like to thank my collaborators: Martin Paulus and Katia Harle who provide their expertise in fMRI and substance dependence study.

I would also like to thank my lab members and friends at UCSD for making my Ph.D. study a colorful journey.

Finally, I would like to give my special thanks to my wife, Qian Yao, and my parents for their love and encouragement. They are always there to give me strongest support when I feel frustrated and depressed. I would not be able to finish this Ph.D. study without their support.

Chapter 2, in full, is a reprint of the material from Ning Ma and Angela J. Yu, Variability in Human Response Time Reflects Statistical Learning and Adaptive Decision-Making, Proceedings of the 37th Annual Conference of the Cognitive Science Society, 2015; Ning Ma and Angela J. Yu, Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control, *Frontiers in Psychology*, 6(1046), 2015. The dissertation author was the primary investigator and

author of this paper.

Chapter 3, in full, is a reprint of the material from Ning Ma and Angela J. Yu, Inseparability of Go and Stop in Inhibitory Control: Go Stimulus Discriminability Affects Stopping Behavior , *Frontiers in Neuroscience*, 10(54), 2016. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of the material from Ning Ma and Angela J. Yu, Stop paying attention: the need for explicit stopping in inhibitory control, *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 2016. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in full, is a reprint of the material from Katia M. Harl, Shunan Zhang, Ning Ma, Angela J. Yu, and Martin P. Paulus, Reduced Neural Recruitment for Bayesian Adjustment of Inhibitory Control in Methamphetamine Dependence, *Biological Psychiatry:Cognitive Neuroscience and Neuroimaging*, 1:448-459, 2016. The dissertation author was the investigator and co-author of this paper.

VITA

2006-2010	B. S. in Physics, University of Science and Technology of China
2010-2012	M. S. in Physics, University of Pittsburgh
2012-2016	M. S. in Statistics, University of California, San Diego
2012-2017	Ph. D. in Electrical Engineering (Intelligent Systems, Robotics, and Control), University of California, San Diego

PUBLICATIONS

Ning Ma, Angela J. Yu, “Variability in Human Response Time Reflects Statistical Learning and Adaptive Decision-Making”, *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, 2015.

Ning Ma, Angela J. Yu, “Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control”, *Frontiers in Psychology*, 6(1046), 2015.

Ning Ma, Angela J. Yu, “Inseparability of Go and Stop in Inhibitory Control: Go Stimulus Discriminability Affects Stopping Behavior”, *Frontiers in Neuroscience*, 10(54), 2016.

Ning Ma, Angela J. Yu, “Stop paying attention: the need for explicit stopping in inhibitory control”, *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 2016.

Katia M. Harl, Shunan Zhang, **Ning Ma**, Angela J. Yu, and Martin P. Paulus, “Reduced Neural Recruitment for Bayesian Adjustment of Inhibitory Control in Methamphetamine Dependence”, *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 1:448-459, 2016.

ABSTRACT OF THE DISSERTATION

Human Learning and Decision-Making, and Their Applications

by

Ning Ma

Doctor of Philosophy in Electrical Engineering (Intelligent Systems, Robotics, and Control)

University of California, San Diego, 2017

Professor Angela Yu, Chair
Professor Ken Kreutz-Delgado, Co-Chair

Intelligent agents often need to make actions with uncertain consequences under changing environment, and to modify those actions adaptively according to ongoing sensory processing and changing task demands. The ability to cancel or modify planned actions according to changing task conditions is known as inhibitory control, and thought to be an important aspect of human cognitive function. Inhibitory control has been studied extensively using the stop-signal experiment. Although a few models and experiments attempted to explain the subject's behavioral result, much

work was still needed to understand the underlying mechanism.

Using Bayesian inference, hidden Markov model and stochastic control theory, this dissertation proposes new model and experimental investigations to attain a more comprehensive understanding of the underlying mechanism of human decision making process in inhibitory control. We demonstrate how human's reaction time, previously thought of as a random quantity, is highly correlated with model simulated predictive belief state. More specifically, the model and data enable us to provide strong evidence that Go process and Stop process are highly dependent, in contrast to being independent, as previously proposed. Our new proposed model can not only cover the behavior data but also the neural data. Finally, by applying our model to the clinical data, we discover the behavior and neural difference between methamphetamine-dependent individuals and comparison group, indicating that the model simulated quantity could be served as a biomarker to predict substance dependent user.

Chapter 1

Introduction

Inhibitory control, the ability to stop or modify preplanned actions under changing task conditions, is an important component of cognitive functions. Intelligent agents often need to make optimal actions with uncertain consequences under changing environment, and to modify those actions adaptively according to ongoing sensory processing and changing task demands. For example, when you are driving a car and approaching a traffic intersection, you need to decide whether you want to pass through the traffic intersection or hit the brake, based on the perceived traffic information. Different actions are associated with different costs. If you decide to keep going when the yellow light is on, you might get a ticket or run into a crash. However, if you decide to stop your car, you have to wait a few more minutes. You want to choose the action which you think will minimize the overall cost. In Psychology, inhibitory control has been studied extensively using the stop-signal experiment. A few models and experiments have been proposed to explain the subject's behavioral and neural result. However, a comprehensive model and experiments are still lacking to fully understand the underlying mechanism. Using the Bayesian inference, hidden

Markov and stochastic control theory, this dissertation proposes new novel model and experimental study to give a more comprehensive understanding of the underlying mechanism of human decision making process in stop signal task, and show how these findings can be applied in clinical drug dependent user study.

In Chapter 2, we combine computational modeling and psychophysics to examine the hypothesis that fluctuations in the noisy measure of go reaction time (Go RT) reflect dynamic computations in human statistical learning and corresponding cognitive adjustments. We model across-trial learning of stop signal frequency, $P(\text{stop})$, and stop-signal onset time, SSD (stop-signal delay), with a Bayesian hidden Markov model, and within-trial decision-making with an optimal stochastic control model. We show that human reaction time is related to the estimated $P(\text{stop})$ and SSD. The results demonstrate that humans indeed readily internalize environmental statistics and adjust their cognitive/behavioral strategy accordingly, and that the reaction time, which was thought as a random variable, can serve as a valuable tool for validating models of statistical learning and decision-making.

In Chapter 3, we then present novel behavioral data and simulation result to valid a key assumption-that is, the Go process and Stop process are dependent with each other. The results thus favor a fundamentally inseparable account of go and stop processing, in a manner consistent with the optimal model, and contradicting the independence assumption of the race model. More broadly, our findings contribute to the growing evidence that the computations underlying inhibitory control are systematically modulated by cognitive influences in a Bayes-optimal manner, thus opening new avenues for interpreting neural responses underlying inhibitory control.

The old rational decision-making model for stopping suggested the observer

makes a repeated Go versus Wait choice at each instant, so that a Stop response is realized by repeatedly choosing to Wait. In Chapter 4, we propose an alternative stochastic control model 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.

In Chapter 5, we use a Bayesian computational approach to examine potential neural deficiencies in the dynamic predictive processing underlying inhibitory function among recently abstinent methamphetamine-dependent individuals (MDIs), a population at high risk of relapse. We show that, relative to comparison subjects, MDIs were more likely to make stop errors on difficult trials and had attenuated slowing following stop errors. MDIs further exhibited reduced sensitivity as measured by the neural tracking of a Bayesian measure of surprise (unsigned prediction error), which was evident across all trials in the left posterior caudate and orbitofrontal cortex (Brodmann area 11), and selectively on stop error trials in the right thalamus and inferior parietal lobule.

Chapter 2

Statistical Learning and Adaptive Decision-Making underlie Human Response Time Variability in Inhibitory Control

Response time (RT) is an oft-reported behavioral measure in psychological and neurocognitive experiments, but the high level of observed trial-to-trial variability in this measure has often limited its usefulness. Here, we combine computational modeling and psychophysics to examine the hypothesis that fluctuations in this noisy measure reflect dynamic computations in human statistical learning and corresponding cognitive adjustments. We present data from the stop-signal task, in which subjects respond to a go stimulus on each trial, unless instructed not to by a subsequent, infrequently presented stop signal. We model across-trial learning of stop signal frequency, $P(\text{stop})$, and stop-signal onset time, SSD (stop-signal delay), with a Bayesian hidden Markov

model, and within-trial decision-making with an optimal stochastic control model. The combined model predicts that RT should increase with both expected P(stop) and SSD. The human behavioral data (n=20) bear out this prediction, showing P(stop) and SSD both to be significant, independent predictors of RT, with P(stop) being a more prominent predictor in 75% of the subjects, and SSD being more prominent in the remaining 25%. The results demonstrate that humans indeed readily internalize environmental statistics and adjust their cognitive/behavioral strategy accordingly, and that subtle patterns in RT variability can serve as a valuable tool for validating models of statistical learning and decision-making. More broadly, the modeling tools presented in this work can be generalized to a large body of behavioral paradigms, in order to extract insights about cognitive and neural processing from apparently quite noisy behavioral measures. We also discuss how this behaviorally validated model can then be used to conduct model-based analysis of neural data, in order to help identify specific brain areas for representing and encoding key computational quantities in learning and decision-making.

2.1 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 [Lam68, Luc86, Smi95, RR98, GS02, BBM⁺06]. 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) [LC84], 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 ($P(\text{stop})$) and onset asynchrony (SSD), and a rational decision-making control policy for within-trial processing, which combines prior beliefs and sensory data to produce behavioral outputs under task-specific constraints/objectives.

This work builds on several previous lines of modeling research. The new model combines a *within-trial* rational decision-making model for stopping behavior [SY11] and an *across-trial* statistical learning model (Dynamic Belief Model; DBM) that sequentially updates beliefs about $P(\text{stop})$ [YC09, SRY10]; it also incorporates a novel across-trial learning component, a simple version of 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 [SY11], which separately penalizes stop error, go (discrimination and omission) error, and response delay, can account for both classical effects in the SST [LC84], such as increasingly frequency of stop errors as a function of SSD and faster stop-error responses than correct go responses, as well as some recently discovered, subtle influences of contextual factors on stopping behavior, such as motivation/reward [LW09] and the baseline frequency of stop trials [EBB⁺07]. We also showed that the across-trial learning model, DBM, can account for sequential adjustment effects not only in SST [SRY10, ISYL13], but also more broadly in simple 2AFC (2-alternative forced choice) perceptual decision-making tasks [YC09] and a visual search task [YH14].

The primary contribution of the current work is to extend a Bayesian model of trial-by-trial learning of $P(\text{stop})$ [SRY10] to also account for learning about the temporal distribution SSD, and to quantify how much of RT variability can be accounted for by each of these learning components. Moreover, we expect that this extended model will be quite useful in identifying brain regions in encoding key computational variables in learning and decision-making.

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

2.2 Materials and Methods

2.2.1 Experiment

The stop signal task consists of a two alternative forced-choice (2AFC) perceptual discrimination task, augmented with an occasional stop signal. Figure 3.1 schematically illustrates our version of the stop-signal task: subject responds to a default go stimulus on each trial within 1100 ms, unless instructed to withhold the response by an infrequent auditory stop signal. The go task is either a random-dot coherent motion task (8%, 15%, or 85% coherence), or a more classical square versus circle discrimination task. On a small fraction of trials, an additional *stop* signal occurs at some time (known as the stop-signal delay, or SSD) after the go stimulus onset, and the subject is instructed to withhold the *go* response. The trials without stop signals are called *go* trials. The SSD is randomly and uniformly sampled on each trial from 100 ms, 200 ms, 300 ms, 400 ms, 500 ms, and 600 ms.

Twenty subjects (13 females) participated in the stop signal task where, on approximately 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. Each subject participated in 12 blocks, 3 block for each stimulus type, and each block containing 75 trials. Two days before the main experiment session, subjects participated in a training session, which contained only 2AFC discrimination and no stop trials. In the training session, there were 10 blocks, 3 blocks for each random dot stimulus coherence and one block for shape discrimination. Subjects were given the same maximal amount of time to respond on the training session trials (1 sec) as in the main experiment. The purpose of the training session is

to allow subjects to familiarize themselves with the task and to allow their perceptual discrimination performance to stabilize. Only data from the main experimental session are analyzed and presented here.

We say that the subject makes a discrimination error when he/she incorrectly responds to the stimulus in go trials, i.e. choosing the opposite motion direction or incorrect shape. The subject makes an omission error if he/she fails to make a go response prior to the response deadline on a go trial. The trials having stop signal are called stop trials; trials without stop signal are go trials. When the subject withhold the response until the response deadline on a stop trial, the trial is considered a stop success (SS) trial; otherwise, it is considered a stop error (SE) trial. Each trial is terminated when the subject makes a response, or at the response deadline itself if no response has been recorded. To incentivize the subjects to be engaged in the task, and to standardize the relative costs of the different kind of errors across individuals, subjects are compensated proportional to points they earn in the task, whereby they lose 50 points for a go discrimination or omission error, 50 points for a stop error, and 3 points for each 100ms of response delay (so maximally 33 points for a trial that terminates with no response, and less if the subject makes a response prior to the response deadline).

This study protocol was approved by the University of California San Diego Human Subjects Review Board, and all subjects gave written informed consent.

2.2.2 Model

In this section, we give a brief description of the computational model we use to capture both within-trial sensory processing and decision-making, and across-trial

learning of $P(\text{stop})$ and SSD. The model for within-trial processing is essentially identical to that in our previous work [SRY10, SY11], while the model for across-trial processing is an augmentation of a previous model [ISYL13] by taking into account not only $P(\text{stop})$ but also SSD.

Within-Trial Processing

Within-trial processing is modeled as a combination of Bayesian sensory processing, which consists of iterative statistical inference about the identity of the go stimulus and the presence of the stop signal, and optimal stochastic control, which chooses whether to Wait or Go (and if so, which Go response) at each instant, based on the accumulating sensory information (Bayesian belief state) and general behavioral objectives (an objective function consisting of parameterized costs for response delay, go discrimination error, go omission error, and stop error). We briefly summarize the model here; a more detailed description can be found elsewhere [SY11].

Sensory processing as Bayesian statistical inference. Figure 2.2A graphically illustrates the Bayesian generative model for how iid noisy sensory data are assumed to be generated by the (true) hidden stimulus states. The two hidden variables d and s correspond respectively to the identity of the go stimulus, $d \in \{0, 1\}$ (0 for left, 1 for right), and whether or not this trial is a stop trial, $s \in \{0, 1\}$. Conditioned on the go stimulus identity d , a sequence of iid sensory inputs, representing the cue of go stimulus, are generated on each trial, x^1, \dots, x^t, \dots , where t indexes time steps *within a trial*. The likelihood functions of d generating the sensory inputs 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 parameters q_d and $1 - q_d$. 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 θ , where θ represents stop signal delay SSD. For simplicity, we assume that θ , also known as the stop-signal delay (SSD), follows a geometric distribution: $P(\theta = t | s = 1) = q(1 - q)^t$. The expected value of θ is $1/q$, which is the expected SSD, $\mathbb{E}[SSD]$, within a trial. Conditioned on z^t , each observation y^t is independently generated and indicates the cue of stop signal. For simplicity, we assume the likelihood functions, $p(y^t | z^t = 0) = g_0(y^t)$ and $p(y^t | z^t = 1) = g_1(y^t)$, are Bernoulli distributions with respective rate parameters q_s and $1 - q_s$.

In the statistically optimal recognition model, Bayes' Rule is applied in the usual iterative manner to compute the iterative posterior probability associated with go stimulus identity, $p_d^t := P(d = 1 | \mathbf{x}^t)$, and the presence of the stop signal, $p_s^t := P(s = 1 | \mathbf{y}^t)$, where $\mathbf{x}^t = \{x^1, x^2, \dots, x^t\}$ and $\mathbf{y}^t = \{y^1, y^2, \dots, y^t\}$ denotes all the data observed so far. 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 time step to time step via Bayes' Rule, by inverting the generative model (Figure 2.2).

Decision making as optimal stochastic control. Figure 2.2B graphically illustrates the sequential decision-making process used to model how an observer chooses whether to Go, when to do so, and which Go response to select on each trial. The decision policy is optimized with respect to the Bayesian belief state and a behaviorally defined cost function that captures the cost and penalty structure of SST, based on which the observer decides at each moment in time whether to Go (and if so, which Go response) or Wait at least one more time step.

On each trial, if the Go action is taken by the response deadline D , it is recorded as a Go response (correct on Go trials, error on Stop trials); otherwise the

trial is terminated by the response deadline and a Stop response is recorded (omission error on Go trials, correct on Stop trials). Let r denote the trial termination time, so that $r = D$ if no response is made before the deadline D , and $r < D$ if a Go action is chosen. $\delta \in \{0, 1\}$ represents the possible binary Go choices produced by making a Go response. We assume there is a cost c incurred per unit time in response delay (corresponding to time-dependent costs, such as time, effort, opportunity, or attention), a stop error penalty of c_s for responding on a Stop trial, and a unit cost for making a discrimination error or commission error on a Go trial – since the cost function is invariant with respect to scaling, we normalize all cost parameters relative to the Go error cost without loss of generality. Thus, the cost function is:

$$l(r, \delta; d, s, \theta, D) = cr + c_s \mathbf{1}_{\{r < D, s=1\}} + \mathbf{1}_{\{r < D, \delta \neq d, s=0\}} + \mathbf{1}_{\{r=D, s=0\}} .$$

The optimal decision policy minimizes the expected (average) loss, $L_\pi = \mathbb{E} [l(r, \delta; d, s, \theta, D)]$,

$$L_\pi = c\mathbb{E} [\tau] + c_s r P(\tau < D | s=1) + (1-r)P(\tau < D, \delta \neq d | s=0) + (1-r)P(\tau = D | s=0)$$

which is an expectation taken over hidden variables, observations, and actions, and generally computationally intractable to minimize directly. Fortunately, having formulated the problem in terms of a belief state Markov decision process, we can effectively use standard dynamic programming [Bel52], or backward induction, to compute the optimal policy and action, via a recursive relationship between the value function and the Q-factors. The value function $V^t(\mathbf{b}^t)$ denotes the expected cost of taking the optimal policy henceforth when starting out in the belief state \mathbf{b}^t . The Q-factors,

$Q_g^t(\mathbf{b}^t)$ and $Q_w^t(\mathbf{b}^t)$, denote the minimal costs associated with taking the action Go or Wait, respectively, when starting out with the belief state \mathbf{b}^t , and subsequently adopting the optimal policy. The Bellman dynamic programming principle, applied to our problem, states:

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

whereby the optimal policy in state \mathbf{b}^t is to choose between Go and Wait depending on which one has the smaller expected cost. Note that 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). Since the observer can no longer update the belief state nor take any action at the deadline, the value function at $t = D$ can be computed explicitly, without recursion, as $V^t(\mathbf{b}^D) = 1 - P_s^D$. Bellman's equation then allows us compute the value functions and Q factors exactly, up to discretization of the belief state space, backwards in time from $t = D - 1$ to $t = 1$. In practice, we discretize the belief state space, (p_d^t, p_s^t) , into 200×200 bins.

The optimal policy partitions the belief state into three discrete action regions: two symmetric *Go regions* for extreme values of p_d and relatively small values of p_s (i.e. where the subject believes the probability of a stop trial is small and the confidence about whether the go stimulus requires a left/right response is high), where the optimal action is to Go, and a large central *Wait region*, where the value of p_d is

close to 0.5 (subject is unsure of go stimulus identity) and/or the value of p_s is large (subject is fairly sure of this being a stop trial), and the optimal action is to Wait.

Across-Trial Processing

Across-trial processing is modeled as Bayesian iterative inference about trial type, $P(\text{stop})$, and the temporal onset of the stop signal, SSD.

Dynamic Belief Model for $P(\text{stop})$. We originally proposed the Dynamic Belief Model (DBM) to explain sequential effects in RT and accuracy in 2AFC tasks, as a function of experienced trial history [YC09], in particular predicting the relative probability of observing a repetition (identical stimulus as last trial) or alternation (different stimulus than last trial) on each trial. Here, as we did earlier [ISYL13], we adapt DBM to model the prior probability of observing a Stop trial (as opposed to Go trial) based on trial history (see Figure 2.3A for a graphical illustration of the generative model, and Figure 2.3B for simulated dynamics of DBM given a sequence of sample observations). We briefly describe the model here; more details can be found elsewhere [YC09, ISYL13].

We assume that γ_k is the probability that trial k is a stop trial, and it 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 generic prior distribution $p_0(\gamma_k)$. The observation s_k is assumed to be drawn from a Bernoulli distribution with a rate parameter γ_k . The iterative posterior and prior of γ_k can be updated via Bayes' Rule by:

$$p(\gamma_k | \mathbf{s}_k) \propto p(\gamma_k | \mathbf{s}_{k-1}) p(s_k | \gamma_k)$$

$$p(\gamma_k | \mathbf{s}_{k-1}) = \alpha p(\gamma_{k-1} = \gamma | \mathbf{s}_{k-1}) + (1 - \alpha) p_0(\gamma_k = \gamma) .$$

Note that the predicted value of γ_k , what we call P(stop), is the mean of the predictive prior distribution: $P(s_k = 1|\mathbf{s}_{k-1}) = \mathbb{E}[\gamma_k|\mathbf{s}_{k-1}] = \int \gamma p(\gamma|\mathbf{s}_{k-1})d\gamma$. Under this model, $P(s_k = 1|\mathbf{s}_{k-1})$ specifies the prior probability of seeing a stop signal for within-trial sensory processing in Section 2.2.2.

Kalman Filter Model for Learning Expected SSD. We use a simple linear-Gaussian dynamical systems model, also known as a Kalman filter [Kal60], to model the trial-by-trial estimation of the mean and variance of the posterior and predictive prior distribution of SSD in the stop-signal task. When the prior distribution of the hidden dynamic variable is Gaussian, the dynamics is linear and corrupted by Gaussian noise, and the observations are a linear function of the hidden variable corrupted by Gaussian noise, the posterior distribution of the hidden variable after each observation, as well as the predictive prior before the next observation, are both Gaussian as well. The Kalman filter describes the statistically optimal (Bayesian) equations for updating the posterior and prior distributions.

As shown in Figure 2.3C, we assume that the observed SSD on (stop) trial k , z_k , is generated from a Gaussian distribution with "true" (hidden) mean h_k and variance R , whereby h_k evolves from (one stop) trial to (another stop) trial under Gaussian noise, with mean 0 and variance Q . We also assume that the prior distribution over h_1 is Gaussian, $p(h_1) = \mathcal{N}(h_0, P_0)$. Then the predictive prior distribution $p(h_k|z_1, \dots, z_{k-1}) = \mathcal{N}(\hat{h}_k^-, P_k^-)$, can be updated using iterative applications of Bayes' Rule (and consistent with Kalman filter equations) as follows:

$$\hat{h}_k^- = \hat{h}_{k-1}$$

$$P_k^- = P_{k-1} + Q$$

and the the posterior distribution, $p(h_k|z_1, \dots, z_k) = \mathcal{N}(\hat{h}_k, P_k)$ can be updated as:

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

where K_k is known as the Kalman gain, which depends on the relative magnitude of state uncertainty P_k^- and the observation noise R . Note that the new posterior is a linear compromise between the predictive prior and observed data, parameterized by K_k (see Figure 2.3D for simulated dynamics of the Kalman filter given a sequence of sample observations). This constitutes a particularly simple case of the Kalman filter, as both the hidden and observed variables are scalar-valued, both the hidden dynamics (how h_k depends on h_{k-1}) and the emission transformation (how z_k depends on h_k) are trivial, and the observer does not actively control the system. The only caveat is that on trials without a stop signal (Go trials), there is no observation for z_k ; we assume on these trials the predictive prior updates as usual and the posterior distribution is identical to the prior (i.e. the inference model is allowed to diffuse as normal, but there is no observation-based correction step). An alternative implementation is to assume that the Kalman filter is not updated on Go trials (no SSD observations). We choose to allow the Kalman filter to diffuse on Go trials, because preliminary analysis indicates that the influence of recently experienced SSD diminishes with increasing number of recent Go trials. Using this model, the prior mean \hat{h}_k^- specifies the mean of the prior distribution for SSD for within-trial processing ($1/q$) in Section 2.2.2.

2.3 Results

Systematic patterns of sequential effects have long been observed in human 2AFC tasks, in which subjects' responses speed up (and accuracy increases) when a new stimulus confirms to a recent run of repetitions or alternations, and slow down (and accuracy decreases) when these local patterns are violated [SBH85, CNB⁺02], as though humans maintain an *expectancy* of stimulus type based on experienced trial sequences and their RT is modulated by this expectancy. Similar sequential effects have also been observed in other classical behavioral experiments used in psychology, including the stop-signal task (SST), in which subjects' Go RT increases with the preponderance of stop trials in recent history [EBB⁺07, LHY⁺08]. We first verify, using a relatively crude model-free method, that this effect is also present in our data. Figure 2.4A shows that Go RT indeed increases with the frequency of stop trials in recent history, and also with the recency of those experienced stop trials. In addition, we hypothesize that, unlike in a basic 2AFC task, subjects may maintain evolving statistical information about stimulus onset time (stop-signal delay, SSD) across trials as well. Figure 2.4C shows that subjects' Go RT indeed increases with the mean SSD of the two most recently experienced stop trials. The strong correlation between Go RT and SSD is also consistent with recent work on decomposing decision components in the Stop-signal Task [WCM⁺14].

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, as a function of observed data, perceived statistical structure of the environment, and behaviorally defined objectives. We can then compare model predictions with human data to see whether our assumptions about the underlying computational processes and objectives hold.

There are two key components to the model (details in Section 3.2.2): (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). These two components are generally referred to as the observation and response models [DOP⁺10]. In the context of modeling behavior, where that behavior is itself modeled under ideal Bayesian assumptions, the observer model constitutes the subject's generative model of how observations are caused, while the response model maps from the implicit beliefs to observed responses. Because we assume subjects' belief updating (Bayesian inference) and response selection are both ideal, given environmental statistics (specified by the Bayesian generative model) and behavioral objectives (specified by the objective/cost function in the stochastic control model), there are no free parameters in either the observation and response models. Furthermore, as we demonstrate through simulations, the ideal mapping between the belief state (obtained using the observation model) and the response time is essentially linear, resulting in a particularly simple parameterization of the response model.

For the first component, we separately model the evolution of subjects' beliefs about the frequency of stop trials, $P(\text{stop})$, using a Bayesian hidden Markov model known as the Dynamic Belief Model (DBM), and their beliefs about the temporal onset of the stop signal, SSD, using a Kalman filter model (Section 2.2.2). We previously proposed DBM to explain sequential effects in 2AFC tasks [YC09], and later adapted it to explain sequential effects in the SST [SRY10, ISYL13] – see Figure 2.3A for a graphical illustration of the generative model, and Figure 2.3B for simulated dynamics of DBM given a sequence of sample observations. To model sequential effects in SSD, we use a simple variant of the Kalman filter [Kal60]. which 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_{k-1} (a Martingale process), whereas DBM assumes that the new hidden variable s_k is either identical to s_{k-1} , 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 undergo large, discrete jumps, which are unlikely in the Kalman filter. In a preliminary analysis (results not shown), we used both the Kalman filter and a modified version of DBM (which takes continuously valued inputs instead of binary ones) to model subjects' beliefs about $\mathbb{E}[SSD]$, and found that the Kalman filter does a significantly better job of accounting for trial-by-trial variability in RT than does DBM.

For the second component, we use a Bayesian inference and optimal decision-making model [SRY10, SY11] to predict when and whether the subject produces a Go response on each trial, as a function of prior beliefs about $P(\text{stop})$ and SSD. The model chooses, in each moment in time, between Go and Wait, depending on ongoing sensory data and the expected costs associated with making a go (discrimination or omission) error, a stop error (not stopping on a stop trial), and response delay (details in Section 2.2.2). Our earlier work showed that this model can explain a range of behavioral and neural data in the SST [SRY10, SY11, ISYL13, HSS⁺14],

We first simulate the within-trial sensory processing/decision-making model to demonstrate how the model predicts Go RT ought to vary as a function of prior beliefs about $P(\text{stop})$ and SSD. Intuitively, we would expect that Go RT ought to increase with the prior belief $P(\text{stop})$, since a 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 induce the observer to respond earlier. Simulations (Figure 2.5) show that Go RT indeed increases monotonically with both $P(\text{stop})$ and $\mathbb{E}[SSD]$, and does so linearly. Note that $P(\text{stop})$ and $\mathbb{E}[SSD]$ are explicitly and naturally specified in the statistical model here (details in the Models section), so we only need to change these parameters and observe their normative consequences by simulating the model, without tuning any free parameters. We also tried uniform distribution for stop signal delay and found that geometric distribution outperformed uniform distribution. The possible reason is that the optimal model, though gives best prediction results, may not be identical to the true underlying learning and inference process of human being. We think it is fine to not make the optimal model be identical to every assumption made in real data analysis, where stop signal delay is uniformly and randomly sampled. As in many instances of (meta) Bayesian modeling of subject behavior, we therefore find that it is sufficient to explain responses in terms of subject-specific prior beliefs. In other words, there is no single Bayes-optimal response valid across all trials, because individuals are equipped with their own priors, continually learned and dynamically evolving according to their individual experiences, and which in turn determine how each observed outcome is assimilated into posterior beliefs and how those beliefs drive observable responses.

Given the strong linear relationship the model predicts to exist between Go RT and both $P(\text{stop})$ and $\mathbb{E}[SSD]$, we expect that the same would be true for human data if the across-trial learning model (Section ??) appropriately models subjects'

prior beliefs about $P(\text{stop})$ and SSD based on experienced trial history, *and* subjects modify their internal sensory processing and decision-making accordingly as prescribed by the rational within-trial decision-making model (Section ??).

As a strong correlation between the two model predictors, $P(\text{stop})$ and $\mathbb{E}[SSD]$, would complicate any analysis and interpretation, we first verify that they are sufficiently decorrelated from each other (as we expect them to be, since SSD on each stop trial is chosen independently in the experimental design). We find that the average correlation between the two, across all subjects, is only 0.019 (std = 0.073), and so treat them as independent variables for the remainder of the paper.

We apply the across-trial learning model to a subject’s experienced sequence of go/stop trials and SSD to estimate their prior beliefs on each trial, and then plot how Go RT varies with the model-based estimates of $P(\text{stop})$ and SSD. Figure 2.6 shows that the subjects’ Go RT increases approximately linearly with prior $P(\text{stop})$ and SSD, as predicted by the model (Figure 2.5). For individual subjects, a linear regression of Go RT versus binned values of $P(\text{stop})$ and $\mathbb{E}[SSD]$, using the same binning procedure as for the group data analysis in Figure 2.6C, is significant in 90% (18/20) of the subjects ($p < 0.05$), with $p = 0.09$ and $p = 0.14$ for the two remaining subjects. On average (across all subjects), we see that variability in $P(\text{stop})$ can explain 34.5% of the variability in the binned RT data (std = 25.0%), while the combined $P(\text{stop})$ and $\mathbb{E}[SSD]$ model can account for 47.2% of the variability in the binned RT data (std=18.9%). RT variability explained by $P(\text{stop})$, on average, accounts for 68.3% of the variability explained by the combined model (std = 34.8%). In other words, $P(\text{stop})$ is a slightly more prominent predictor of RT variability, although we do see that in 25% of the subjects (5/20), $\mathbb{E}[SSD]$ was a stronger predictor of RT variability

than $P(\text{stop})$, i.e. $P(\text{stop})$ accounted for less than 50% of the variance explained by the combined model.

These results imply that humans both continuously monitor and update internal representations about statistics related to stimulus frequency and temporal onset, and adjust their behavioral strategy rationally according to those evolving representations. Moreover, we can get some insight into implicit human assumptions about environmental statistics based on estimated model parameters. For DBM, we found that $\alpha = 0.75$ yields the best linear fit between Go RT and $P(\text{stop})$ (highest R^2 value), implying that subjects assume that the frequency of stop trials changes on average once every four trials, (the expected duration between changes is $1/(1 - \alpha)$). This is consistent with the α value previously found in a DBM account of sequential effects in a 2AFC perceptual discrimination task [YC09]. We also found through simulations (results not shown) that the model fit was not very sensitive to the precise value of a and b , the parameters of the Beta prior distribution $p_0(\gamma)$, in that different values of (a, b) yield highly correlated predictions of $P(\text{stop})$. Thus, a and b were not optimized with respect to the data but instead fixed at $a/(a + b) = 0.25$, equivalent to the empirical baseline frequency of stop trials, and $a + b = 10$. For the Kalman filter, we found that $Q = 0.03$ and $R = 0.15$ yield the best linear fit between Go RT and $\mathbb{E}[SSD]$, which implies that subjects expect on average that h_k will "diffuse" from trial to trial according to a Gaussian distribution with a standard deviation of $\sqrt{0.03} = 0.17$ sec, and that the perceived SSD is corrupted by unbiased sensory noise with a standard deviation of $\sqrt{0.15} = 0.39$ sec. The correlation between Go RT and $\mathbb{E}[SSD]$ is not very sensitive to the other Kalman filter parameters (results not shown), h_0 , and P_0 , and thus those were specified with fixed values (see caption of

Figure 2.3).

2.4 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 show that some of this variability reveals how fluctuations in experienced empirical statistics are used by human observers to continuously update their internal representation of environmental statistics and rationally adjust their behavioral strategy in response. To be sure, our model is only a partial explanation of the overall RT variability. While our model is able to account for about half of the RT variability *averaged across subjects*, there is additional RT variability not accounted for by the model, which is obscured by the averaging process. Much room remains for future work to determine additional contextual and individual-specific factors that drive variabilities in RT.

In this work, we assumed fixed model parameters, such as the critical α parameter for tracking $P(\text{stop})$ and the ratio R/Q for tracking $\mathbb{E}[SSD]$, both of which parameterize the *stability* of environmental statistics and thus determine the size of the "memory window" for using previous trials to predict the next trial. One may well ask whether human subjects in fact undergo meta-learning about these and other parameters over the time course of the experiment. The short answer is "no", as we see no statistically discernible differences in human behavior in the first and second halves of the experiment (data not shown). This is not surprising, given that in a much simpler 2AFC task (where the cognitive demands within each trial are much

lower), the first in which we successfully accounted for sequential effects as arising from tracking local statistics of the sensory environment [YC09], we found that not only did subjects not behave differently in the first and second halves of the experiment, but that from an ideal observer point of view, meta-learning of α is much *too slow* to give rise to noticeably different behavior over the time course of one experimental session.

Separately, this work makes an important contribution to 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 [Bar97, Nig00]. Stopping behavior is also known to be impaired in a number of psychiatric populations with presumed inhibitory deficits, such as attention-deficit hyperactivity disorder [ARK07], substance abuse [NWM⁺06], and obsessive-compulsive disorder [MAC⁺07]. The work present here can help elucidate the psychological and neural underpinnings of inhibitory control, by providing a quantitatively precise model for the critical computational components, and thus informing both experimental design and data analysis in future work for the identification of neural functions. Along these lines, the current work has concrete applications in the analysis of neurophysiology data. Previously, we successfully applied the P(stop)-learning model [SRY10] in a model-based analysis of fMRI data [ISYL13], and discovered that the dorsal anterior cingulate cortex (dACC) encodes a key prediction error signal related to P(stop); moreover, we found the dACC prediction-error signal is altered in young adults at risk for developing stimulant addiction [HSS⁺14], a condition known to be associated with impaired inhibitory control and specifically stopping behavior. We expect that this new, extended model should be even more powerful in capturing

human behavior and identifying neural correlates of the computations involved in proactive control, which is concerned with the preparation for inhibition in advance of sensory input. As we have *behaviorally validated*, trial-by-trial measures of underlying belief states, our model can be used to look for neural responses specifically correlated with these beliefs, allowing us to establish the functional neural anatomy of different sorts of probabilities and uncertainties.

Beyond specific implications for inhibitory control and response modeling, this work exemplifies an approach for leveraging apparently "noisy" experimental measures such as RT, to glean deep insights about cognitive representation and behavioral strategy in humans (and other animals). Even though our experiment did not explicitly manipulate the frequency of stop trials or the onset of the stop signal across the experimental session, subjects still used chance fluctuations in the local statistics of empirical observation to continuously modify their internal beliefs, and modulate their behavioral strategy accordingly. This raises the very real possibility that humans are constantly updating their internal model of the environment in any behavioral task, and the apparent "noise" in their behavioral output may often arise from an underlying monitoring and adaptation process, which can be brought to light by incorporating sophisticated normative modeling tools, such as the Bayesian statistical modeling and stochastic control methods used here. With the broadening use of these modeling tools, there should be exciting new possibilities for advancing the neural, psychological, and psychiatric study of learning, decision-making, and cognitive control.

Acknowledgments

This chapter, in full, is a reprint of the material from Ning Ma and Angela J. Yu, Variability in Human Response Time Reflects Statistical Learning and Adaptive Decision-Making, Proceedings of the 37th Annual Conference of the Cognitive Science Society, 2015; Ning Ma and Angela J. Yu, Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control, *Frontiers in Psychology*, 6(1046), 2015. The dissertation author was the primary investigator and author of this paper.

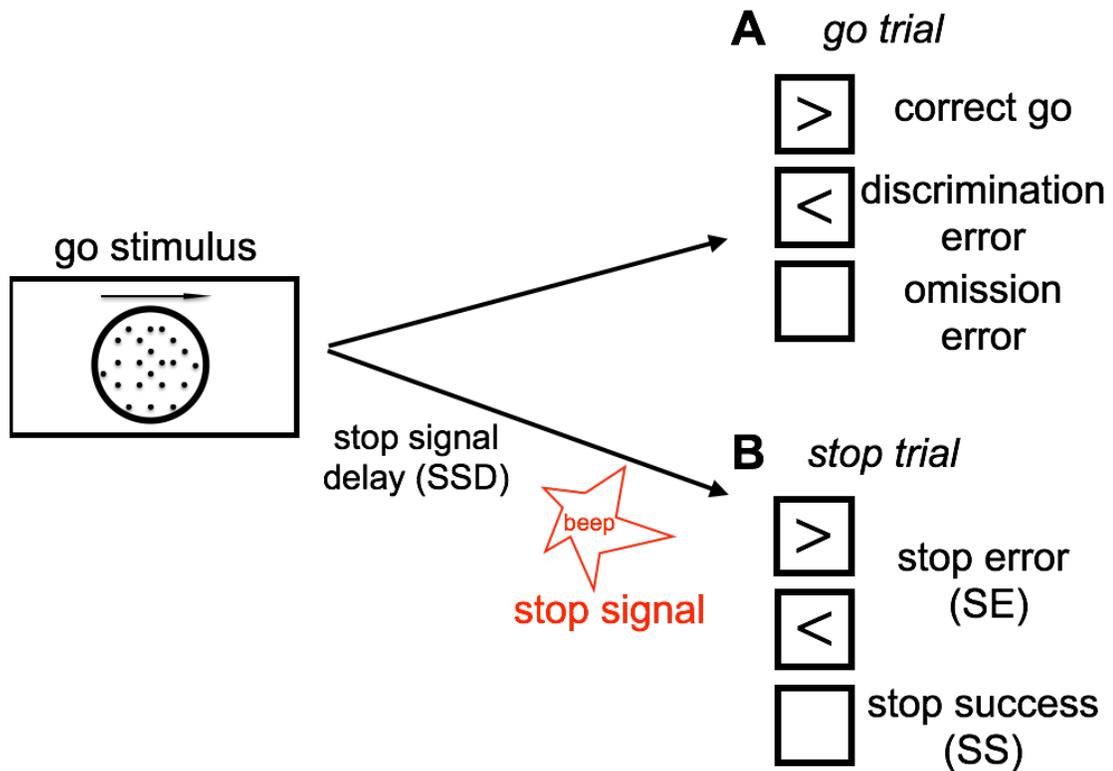


Figure 2.1: Schematic illustration of our stop signal task (A) Go trials: On go trials, subject is supposed to make a response to a default go response by pressing the left or right button, based on the coherent motion direction of random dots. The go reaction time (Go RT) is defined as the time the subject takes to make a go response since the onset of go stimulus. The subject makes a discrimination error if he/she chooses the wrong direction (hit the wrong button). (B) Stop trials: On small fraction of trails, a stop signal will appear and instruct the subject to withhold the go response. The time delay between the occurrence of stop signal task and the onset of go stimulus is called the stop-signal delay (SSD). If the subject makes a go response in a stop trial, this trial is considered a stop error (SE) trial, otherwise it is considered a stop success trial (SS).

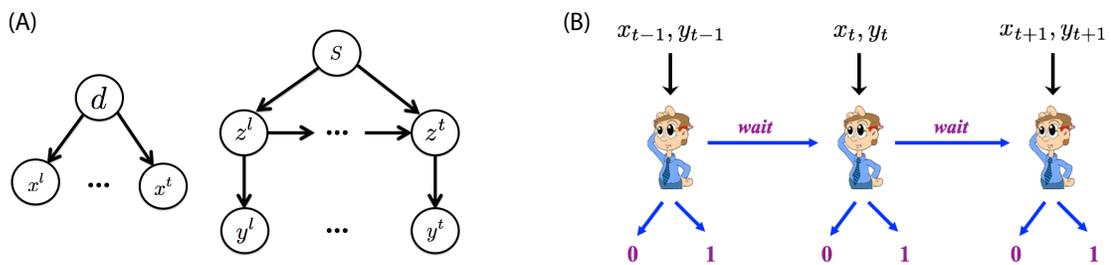


Figure 2.2: Within-trial sensory processing and decision-making. (A) Bayesian generative model of iid sampled sensory observations (x^1, \dots, x^t, \dots) conditioned on Go stimulus identity ($d = 0$ of left, $d = 1$ for right), and an independent stream of observations (y^1, \dots, y^t, \dots) conditioned on the presence ($z^t = 1$) or absence ($z^t = 0$) of the Stop signal, which has a geometrically distributed onset time when it is a stop trial ($s = 1$) and never appears on a go trial ($s = 0$). (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.

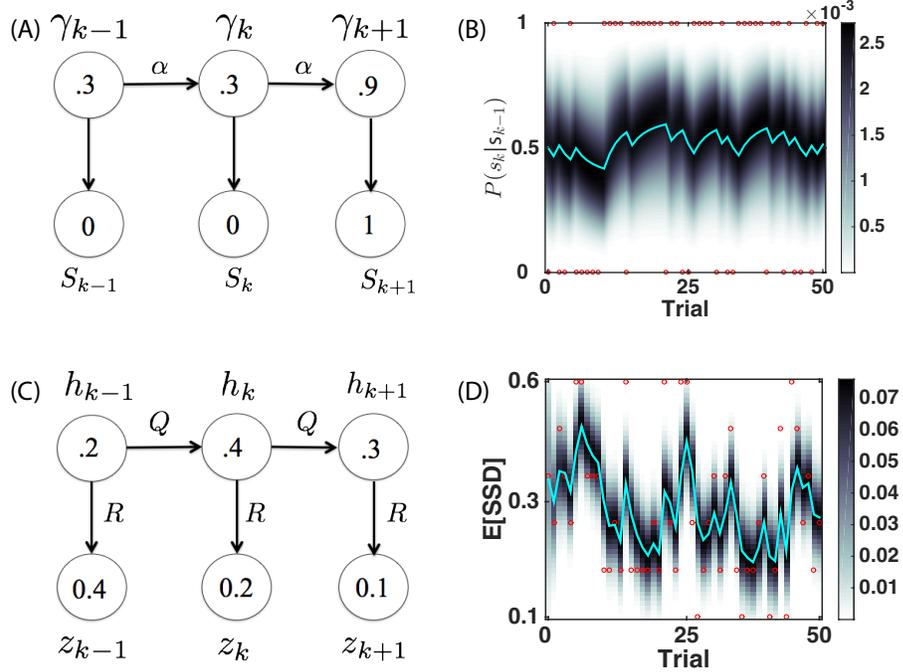


Figure 2.3: Bayesian sequential inference model for learning $P(\text{stop})$ and $\mathbb{E}[SSD]$. (A) Graphical model for DBM. $\gamma \in [0,1]$, $s_k \in \{0,1\}$. $p(\gamma_k | \gamma_{k-1}) = \alpha \delta(\gamma_k - \gamma_{k-1}) + (1 - \alpha) p_0(\gamma_k)$, where $p_0 = \text{Beta}(a, b)$. 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: $\alpha = 0.75$, $p_0 = \text{Beta}(2.5, 7.5)$. (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: $Q = 0.03$, $R = 0.15$, $h_0 = 0.35$, $P_0 = 1$. Unless otherwise stated, these parameters are used in all the subsequent simulation.

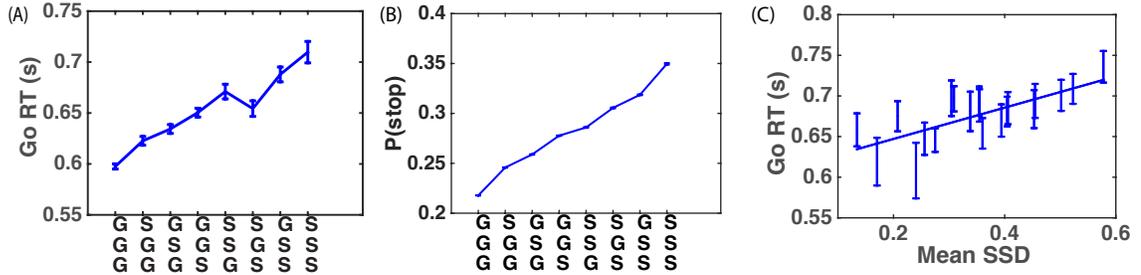


Figure 2.4: Sequential effects in human data. (A) Go RT increases with the frequency and recency of stop trials in recent trial history. Along the abscissa are all possible three-trial sub-sequences of Go and Stop trials: most recent trial is on the bottom. The Go RT of the correct go trial immediately following the sub-sequence is recorded. Go RT data are then averaged over all trials of a particular pattern for all subjects. Error bars indicate s.e.m. of Go RT in each pattern. (B) Model-predicted $P(\text{stop})$ increases with the frequency and recency of stop trials in recent trial history. Analogous to (A), the prior $P(\text{stop})$ of the trial immediately following each sub-sequence is computed using DBM. Estimates of $P(\text{stop})$ from all trials and all subjects are then averaged in each pattern. DBM parameters: $\alpha = 0.75$, $a/(a + b) = 0.25$. (C) Go RT increases with experienced SSD. Go RT is plotted against mean SSD of the two most recent stop trials. A Go trial is only included if it directly follows a Stop trial (and the response was correct), and the two previous Stop trials are separated by no more than three Go trials. These restrictions are adopted because preliminary analysis indicates that the influence of recently experienced SSD diminishes with increasing number of recent Go trials. Each bin of SSD (spaced such that there are equal number of data points in each bin) contains Go RT from all trials and subjects where $\mathbb{E}[SSD]$ fell with this bin. Both Go RT and SSD are averaged within each bin. Straight line denotes best linear fit. Error bars denote s.e.m. across subjects. $R^2 = 0.56$, $p = 0.0002$.

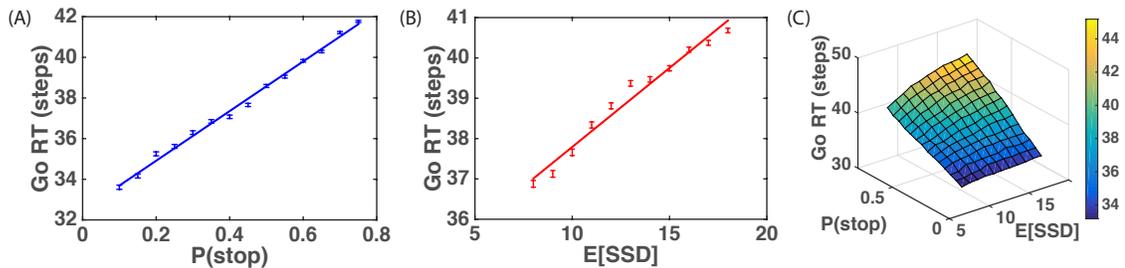


Figure 2.5: Model prediction of Go RT versus $P(\text{stop})$ and $\mathbb{E}[SSD]$. (A) Go RT vs. $P(\text{stop})$: simulated Go RT for a ranged of $P(\text{stop})$ values (.1, .15, ..., .75). Data averaged over 10000 simulated Go trials for each value of $P(\text{stop})$. Straight line denotes best linear fit. Error bars denote s.e.m. $1/q = \mathbb{E}[SSD] = 10$. (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. $P(\text{stop}) = 0.45$. (C) Go RT vs. $P(\text{stop})$ and $\mathbb{E}[SSD]$: simulated Go RT for a range of $P(\text{stop})$ and $\mathbb{E}[SSD]$ values, where $P(\text{stop})$ varies between .1 and .75, and $\mathbb{E}[SSD]$ varies between 8 and 18. Data averaged over 10000 simulated Go trials for each $(P(\text{stop}), \mathbb{E}[SSD])$. 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.

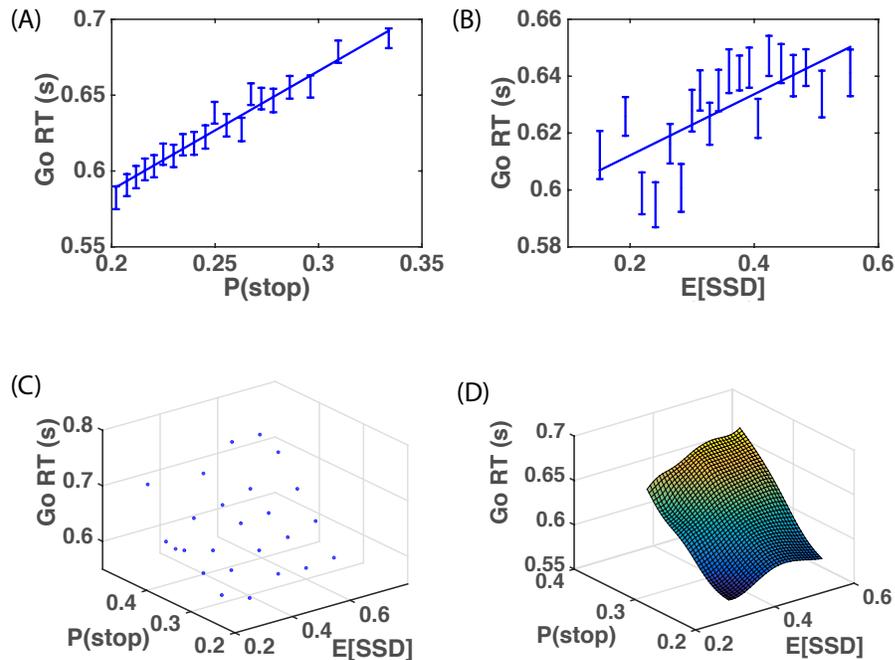


Figure 2.6: Human Go RT versus model-estimated P(stop) and SSD. (A) Go RT vs. P(stop): P(stop) on each trial estimated by DBM based on actual sequence of stop/go trials the subject experienced prior to the current trial. Binning of $\mathbb{E}[SSD]$ spaced to ensure equal number of data points in each 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: $\alpha = 0.75$, $p_0 = \text{Beta}(2.5, 7.5)$. (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: $R^2 = 0.52$, $p = 0.0003$. Error bars denote s.e.m. Kalman filter (KF) parameters: $Q = 0.03$, $R = 0.15$, $h_0 = .35$, $P_0 = 1$. (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(..., 'v4')*, a biharmonic spline interpolation method, to the data in (C).

Chapter 3

Inseparability of Go and Stop in Inhibitory Control: Go Stimulus Discriminability Affects Stopping Behavior

Inhibitory control, the ability to stop or modify preplanned actions under changing task conditions, is an important component of cognitive functions. Two lines of models of inhibitory control have previously been proposed for human response in the classical stop-signal task, in which subjects must inhibit a default *go* response upon presentation of an infrequent *stop* signal: (1) the *race model*, which posits two independent *go* and *stop* processes that race to determine the behavioral outcome, go or stop; and (2) an *optimal decision-making model*, which posits that observers decides whether and when to go based on continually (Bayesian) updated information about both the go and stop stimuli. In this work, we probe the relationship between *go*

and *stop* processing by explicitly manipulating the discrimination difficulty of the go stimulus. While the race model assumes the go and stop processes are independent, and therefore go stimulus discriminability should not affect the stop stimulus processing, we simulate the optimal model to show that it predicts harder go discrimination results in a longer go reaction time (RT), a lower stop error rate, as well as a faster stop-signal RT. We then present novel behavioral data that validate these model predictions. The results thus favor a fundamentally inseparable account of go and stop processing, in a manner consistent with the optimal model, and contradicting the independence assumption of the race model. More broadly, our findings contribute to the growing evidence that the computations underlying inhibitory control are systematically modulated by cognitive influences in a Bayes-optimal manner, thus opening new avenues for interpreting neural responses underlying inhibitory control.

3.1 Introduction

The ability to cancel or modify planned actions according to changing task conditions is known as *inhibitory control*, and thought to be an important aspect of human cognitive function. Inhibitory control has been studied extensively using the stop-signal [LC84], in which subjects typically discriminate a *go* stimulus on each trial, but occasionally encounter a *stop* signal following the go stimulus, which instructs the subject to withhold the go response (see Figure 3.1). Two major class of models have been proposed to account for the underlying computational and neural processes in the stop-signal task. The first is the classical race model and its variants [LC84, BPLS07], which posit a race between two independent go and stop processes. The model assumes essentially immutable, though noisy, termination

times for the go and stop processes, whereby the average stop process delay, known as the stop-signal reaction time (SSRT), is generally thought to be a measure of an individual's inhibitory capacity. Correspondingly, SSRT has been measured as longer in populations with presumed inhibitory deficits [ARK07, NWM⁺06, MAC⁺07], and neural activities in certain primate brain regions (e.g. frontal eye field and superior colliculus) have been interpreted to reflect components of the race model [HSP98, PH03]. However, problematic for the simple race model, various cognitive contextual factors have been shown to systematically modulate stopping behavior, such as the reward structure of the task [LW09] and the statistical frequency of stop signals [EBB⁺07]. In response to these and other observed cognitive influences, we previously proposed an alternative model of inhibitory control, a Bayes-optimal decision-making model positing that subjects choose when and whether to initiate a go response according to continually (Bayesian) updated sensory beliefs about both the go and stop stimuli, and relative to a behavioral objective function that penalize go and stop errors as well as response delay. As we previously showed, this optimal model can capture cognitive influences on stopping behavior as a function of sensory statistics at multiple timescales [ISYL13, MY15b, MY15a] and the reward structure of the task [SY11].

In this work, we explore a particular type of interaction between go and stop processing in the stop-signal task. Specifically, we consider the computational and behavioral consequences of manipulating the go stimulus discrimination difficulty. We will use simulations of the optimal model to demonstrate that, as the go stimulus becomes noisier (harder to discriminate), the go reaction time (RT) should get longer and consequently the rate of stop errors to drop, as subjects have a greater opportunity to detect the late-appearing stop signals before initiating the go response; perhaps

less obviously, SSRT also becomes shorter, which also contributes to decreasing stop errors. We will then present novel experimental data from human subjects ($n = 20$) performing a stop-signal task, in which the go task is to discriminate a random-dot coherent motion stimulus [BSNM92], and the stimulus difficulty (coherence) is varied across blocks. The key question is whether SSRT co-varies with stimulus coherence, as predicted by the optimal model, or whether stopping behavior is independent of go discriminability, as assumed by the race model. More broadly, we will examine if the spectrum of behavioral measures – SSRT, go RT, and stop error rate – changes systematically with respect to go stimulus coherence as predicted by the optimal model, as that would yield further evidence that human inhibitory control is under sophisticated, context-sensitive, and statistically optimal cognitive control.

In the following, we first describe the experimental procedure (Section 2.2.1), then model details (Section 3.2.2), followed by model simulation results and behavioral data. We conclude with some discussions and thoughts on related work and future directions (Section 5.4).

3.2 Materials and Methods

3.2.1 Experiment

The stop signal task consists of a two alternative forced-choice (2AFC) perceptual discrimination task, interspersed with an occasional stop signal. Figure 3.1 schematically illustrates our version of the stop-signal task: subject responds to a default go stimulus on each trial (go trial), unless instructed to withhold the response by an infrequent auditory stop signal (stop trial). The go task is either a random-dot

coherent motion task (8%, 15%, or 85% coherence), or a more classical square versus circle discrimination task. On a small fraction (25%) of trials, an additional auditory *stop* signal (a beep) occurs at some time after the go stimulus onset (known as the stop-signal delay, or SSD), which instructs the subject to withhold the *go* response. The SSD is randomly and uniformly sampled on each trial from 100 ms, 200 ms, 300 ms, 400 ms, 500 ms, and 600 ms.

We say that the subject makes a discrimination error when he/she incorrectly responds to the stimulus on go trials, i.e. choosing the opposite motion direction or incorrect shape. The subject makes an omission error if he/she fails to make a go response prior to the response deadline on a go trial, set to be 1100 ms in the experiment. The trials having stop signal are called stop trials; trials without stop signal are go trials. When the subject withhold the response until the response deadline on a stop trial, the trial is considered a stop success (SS) trial; otherwise, it is considered a stop error (SE) trial. Each trial is terminated when the subject makes a response, or at the response deadline itself if no response has been recorded. To incentivize the subjects to be engaged in the task, and to help standardize the relative costs of the different kind of errors across individuals, subjects are compensated proportional to points they earn in the task, whereby they lose 50 points for a go discrimination or omission error, 50 points for a stop error, and 3 points for each 100ms of response delay (so maximally 33 points for a trial that terminates with no response, and less if the subject makes a response prior to the response deadline).

Twenty subjects (13 females) participated in the experiment. Each subject performed 12 blocks, 3 block for each stimulus type, with each block containing 75 trials. Two days before the main experiment session, subjects participated in a training

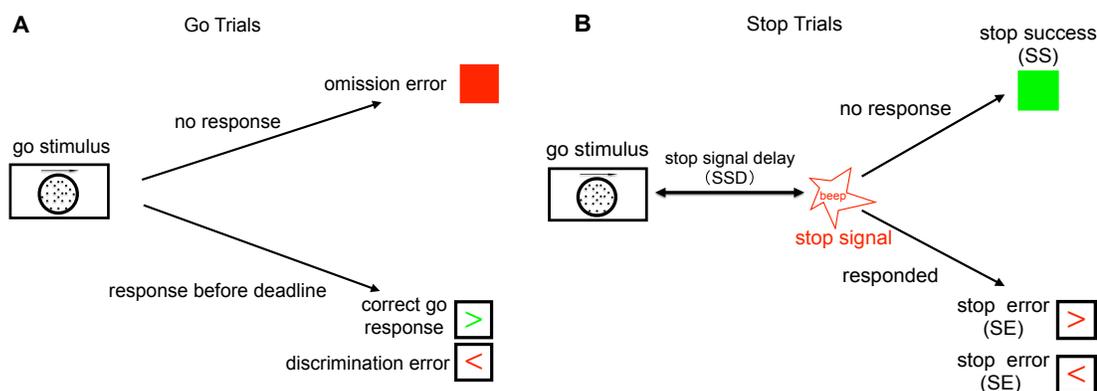


Figure 3.1: Schematic illustration of our stop signal task (A) Go trials: On go trials, subject is supposed to make a response to a default go response by pressing the left or right button, based on the coherent motion direction of random dots. The go reaction time (Go RT) is defined as the time the subject takes to respond from the onset of go stimulus. The subject makes a discrimination error if he/she chooses the wrong direction (wrong key), and an omission error if no response is recorded within the respond deadline (1100 ms). (B) Stop trials: on a small fraction of trials, a stop signal appears after the go stimulus and instructs the subject to withhold the go response. The time delay between the go stimulus onset and the stop signal onset is called the stop-signal delay (SSD). If the subject makes a go response in a stop trial, this trial is considered a stop error (SE) trial, otherwise it is considered a stop success (SS) trial.

session, which contained only the 2AFC discrimination and no stop trials. In the training session, there were 10 blocks, 3 blocks for each random dot stimulus coherence and one block for shape discrimination. Subjects were given the same maximal amount of time to respond on the training session trials (1100 ms) as in the main experiment. The purpose of the training session is to allow subjects to familiarize themselves with the task and to achieve stable perceptual discrimination performance. Only data from the main experimental session are analyzed and presented here.

This experimental protocol was approved by the University of California San Diego Human Subjects Review Board, and all subjects gave written informed consent.

3.2.2 Model

The Race Model

The classical race model for studying inhibitory control is shown in Figure 3.2A. The subject makes a stop error when the *go* response is finished processing 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 subject requires to process the stop signal and cancel the *go* response (in practice, it is often calculated as the difference between the median *go* RT and the SSD specific to each subject for achieving 50% accuracy on stop trials). SSRT characterizes the stopping ability of the subject and is highly related to the inhibitory deficits observed in various psychiatric conditions (e.g. substance abuse [NWM⁺06], attention-deficit hyperactivity disorder [ARK07], schizophrenia [BMJC02], obsessive-compulsive disorder [MAC⁺07]).

Optimal Inhibitory Control Model

Recently, we proposed a normative Bayesian Markov decision process (MDP) model for the stop-signal task [SRY10, SY11], which assumes that the subjects maintain continually evolving Bayes-optimal beliefs about the sensory environment (see Figure 3.2B), and that they make an moment-by-moment decisions between *go* and *wait* by mapping the current belief state into the action space relative to an objective function Figure 3.2C. The behavioral objective function is assumed to take into account the costs associated with *go* errors, stop errors, and response delay.

Figure 3.2B illustrates the Bayesian generative model for how iid noisy sensory data are generated by the (true) hidden stimulus states. The two hidden variables d and s correspond respectively to the identity of the *go* stimulus, $d \in \{0, 1\}$ (0 for

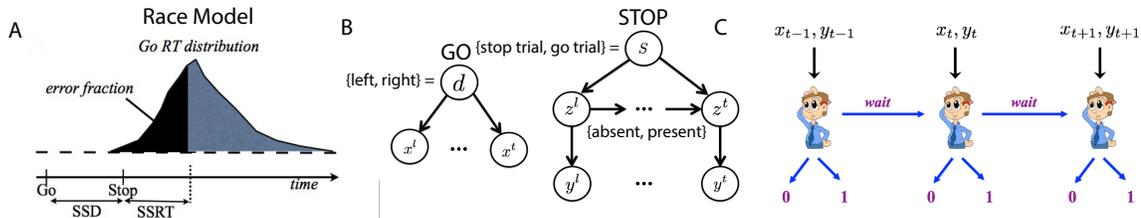


Figure 3.2: (A) The classical race mode for behavioral in the stop-signal task. The behavioral outcome (go or stop) is determined by a race between a go and a stop process. The go reaction time has a broad distribution due to noise. The stop process has an average delay known as the *stop signal reaction time* (SSRT). The stop error rate is the cumulative density of Go RT at $\text{SSD} + \text{SSRT}$. The SSRT is thus estimated from data as the difference between the median go RT and the SSD at which 50% stop error is achieved, as $\text{SSD} + \text{SSRT} = \text{median}(\text{go RT})$ implies 50% of stop trials will end in error (the rest in success). (B) Bayesian generative model of iid sampled sensory observations (x^1, \dots, x^t, \dots) conditioned on Go stimulus identity ($d = 0$ of left, $d = 1$ for right), and an independent stream of observations (y^1, \dots, y^t, \dots) conditioned on the presence ($z^t = 1$) or absence ($z^t = 0$) of the Stop signal, which has a geometrically distributed onset time when it is a stop trial $s = 1$ and never appears on a go trial ($s = 0$). (C) The decision of whether/when to Go, and which Go response to select, are modeled as sequential decision-making, where the subject chooses at each moment whether to select a Go response, or to wait at least one more time point.

left, 1 for right), and whether or not this trial is a stop trial, $s \in \{0, 1\}$. $P(s = 1)$ and $P(d = 1)$ are the prior probability of a stop trial and one of two go alternatives, respectively. Conditioned on the go stimulus identity d , a sequence of iid sensory inputs, representing the cue of go stimulus, are generated on each trial, x^1, \dots, x^t, \dots , where t indexes time steps *within a trial*. The likelihood functions of d generating the sensory inputs 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 parameters q_d and $1 - q_d$. The parameter q_d specifies stimulus signal-to-noise ratio, thus reflecting the go stimulus difficulty. The discrimination task becomes harder (lower coherence) when q_d is closer to 0.5 and easier (high coherence) when q_d is closer to 1 or 0. 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 θ , where θ represents stop signal delay SSD. For simplicity, we assume that θ , also known as the stop-signal delay (SSD), follows a geometric distribution: $P(\theta = t | s = 1) = q(1 - q)^{t-1}$. The expected value of θ is $1/q$, which is the expected SSD, $\mathbb{E}[SSD]$, within a trial. Conditioned on z^t , each observation y^t is independently generated and indicates the cue of stop signal. For simplicity, we assume the likelihood, $p(y^t | z^t = 0) = g_0(y^t)$ and $p(y^t | z^t = 1) = g_1(y^t)$, are Bernoulli distributions with rate parameters q_s and $1 - q_s$, respectively.

In the statistically optimal recognition model, Bayes' Rule is applied in the usual iterative manner to compute the iterative posterior probability associated with go stimulus identity, $p_d^t := P(d = 1 | \mathbf{x}^t)$, the presence of the stop signal, $p_z^t := P(\theta \leq t | \mathbf{y}^t)$, and whether the current trial is a stop trial, $p_s^t := P(s = 1 | \mathbf{y}^t)$, where $\mathbf{x}^t = \{x^1, x^2, \dots, x^t\}$ and $\mathbf{y}^t = \{y^1, y^2, \dots, y^t\}$ denotes all the data observed so far. 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 time step to time step via Bayes' Rule, by inverting the generative model (Figure 3.2 B) as the following.

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

To infer the stop signal, we first update p_z^t iteratively by

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

where $h(t)$ is the posterior probability that the stop-signal will appear in the next

time step given that it has not appeared yet,

$$h(t) = \frac{rP(\theta = t|s = 1)}{rP(\theta > t - 1|s = 1) + (1 - r)} = \frac{rq(1 - q)^{t-1}}{r(1 - q)^{t-1} + (1 - r)}$$

where $r = P(s = 1)$ is the prior probability of a stop trial. The posterior probability that the current trial is a stop trial can be computed as

$$p_s^t = P(s = 1|\mathbf{y}^t) = p_z^t + (1 - p_z^t)P(s = 1|\theta > t, \mathbf{y}^t)$$

where $P(s = 1|\theta > t, \mathbf{y}^t)$ is independent on the past observations

$$\begin{aligned} P(s = 1|\theta > t, \mathbf{y}^t) &= \frac{P(\theta > t|s = 1)P(s = 1)}{P(\theta > t|s = 1)P(s = 1) + P(\theta > t|s = 0)P(s = 0)} \\ &= \frac{(1 - q)^t r}{r(1 - q)^t + (1 - r)} \end{aligned}$$

Figure 3.2C illustrates the sequential decision-making process that determines how an observer chooses whether/when to Go, and which Go response to select. The Markov decision process is optimized with respect to the Bayesian belief state and a behaviorally defined cost function that captures the cost and penalty structure of SST, based on which the observer decides at each moment in time whether to Go (and if so, which Go response) or Wait at least one more time step.

On each trial, if the Go action is taken by the response deadline D , it is recorded as a Go response (correct on Go trials, error on Stop trials); otherwise the trial is terminated by the response deadline and a Stop response is recorded (omission error on Go trials, correct on Stop trials). 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 assume there is a cost c incurred per unit time in response delay (corresponding to time-dependent costs, such as time, effort, opportunity, or attention), a stop error penalty of c_s for responding on a Stop trial, and a unit cost for making a discrimination error or omission error on a Go trial – since the cost function is invariant with respect to scaling, we normalize all cost parameters relative to the Go error cost without loss of generality. Thus, the cost function is:

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

where r denotes the prior probability of encountering a stop trial, and $\mathbf{1}$ denotes the indicator function, which evaluates to 1 when the condition specified in the curly brackets are met, and to 0 otherwise.

The optimal decision policy minimizes the expected (average) loss, $L_\pi = \mathbb{E}[l(\tau, \delta; d, s, \theta, D)]$,

$$L_\pi = c\mathbb{E}[\tau] + c_s r P(\tau < D | s=1) + (1-r)P(\tau < D, \delta \neq d | s=0) + (1-r)P(\tau = D | s=0)$$

which is an expectation taken over hidden variables, observations, and actions, and generally computationally intractable to minimize directly. Fortunately, after formulating the problem in terms of a belief state Markov decision process, we can effectively use standard dynamic programming [Bel52] to compute the optimal policy and action via a recursive relationship between the value function and the Q-factors. The value function $V^t(\mathbf{b}^t)$ denotes the expected cost of taking the optimal policy henceforth when starting out in the belief state \mathbf{b}^t . The Q-factors, $Q_g^t(\mathbf{b}^t)$ and $Q_g^w(\mathbf{b}^t)$, denote

the minimal costs associated with taking the action Go or Wait, respectively, when starting out with the belief state \mathbf{b}^t , and subsequently adopting the optimal policy.

The Bellman dynamic programming principle, applied to our problem, implies:

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

whereby the optimal policy in state \mathbf{b}^t is to choose between Go and Wait depending on which one has the smaller expected cost. Note that a Go response terminates the current trial, while a Wait response lengthens the current trial by at least one more time step, and repeated choice of Wait until the response deadline constitutes a Stop response. Since the observer can no longer update the belief state nor take any action at the deadline, the value function at $t = D$ can be computed explicitly, without recursion, as $V^t(\mathbf{b}^D) = cD + (1 - P_s^D)$. Bellman's equation then allows us compute the value functions and Q factors exactly, backward in time from $t = D - 1$ to $t = 1$.

We note that the decision problem can also be formalized as a mathematically equivalent partially observable Markov decision process (POMDP), whereby the hidden state is the stimulus state (d, s) , the observations are iid noisy samples conditioned on that hidden state, and actions are chosen (Go or Wait) based on all previous observations as well as any prior beliefs about the hidden state. However, it is a rather trivial sort of POMDP, as not only do the actions not affect the hidden state, but the hidden state does not have any dynamics at all. Instead, we chose to formulate the problem as a (belief) Markov Decision Process, whereby the hidden state at time t is

the posterior distribution over the stimuli at time t (the initial state is just the the prior distribution), and its (non-trivial) evolution over time is governed exactly by Bayes' Rule applied to the previous posterior state and the new observation and is completely observed. The only caveat is that the belief state is a continuous variable, and thus in order to apply Bellman's dynamic programming equation, we have to discretize the belief state. In the simulations, we discretize the belief state space, (p_d^t, p_s^t) , into 200×200 bins.

Bayes rule implies 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 $\mathbb{E}[V^{t+1}]$ by averaging over all possible next observations x^{t+1}, y^{t+1} .

$$\begin{aligned} \mathbb{E}[V^{t+1}(\mathbf{b}^{t+1})|\mathbf{b}^t] &= \sum_{x^{t+1}, y^{t+1}} p(x^{t+1}, y^{t+1}|\mathbf{b}^t) V^{t+1}(\mathbf{b}^{t+1}(\mathbf{b}^t, x^{t+1}, y^{t+1})) \\ p(x^{t+1}, y^{t+1}|\mathbf{b}^t) &= p(x^{t+1}|p_d^t) p(y^{t+1}|p_s^t) \\ p(x^{t+1}|p_d^t) &= p_d^t f_1(x^{t+1}) + (1 - p_d^t) f_0(x^{t+1}) \\ p(y^{t+1}|p_s^t) &= (p_z^t + (1 - p_z^t) h(t+1)) g_1(y^{t+1}) + (1 - p_z^t) (1 - h(t+1)) g_0(y^{t+1}) \end{aligned}$$

The optimal decision policy partitions the belief state into Go and Stop regions, such the optimal decision is to go (and terminate the trial) if the belief state at time t , (p_d, p_z) , falls into a Go region (where $Q_g < Q_w$), and the optimal decision is to wait (at least one more time point, but with the possibility of going later before the deadline) if the belief state falls into a Wait region (where $Q_w < Q_g$). Figure 3.3 shows that there are typically two symmetric Go regions, where p_s is relatively small, and p_d is close to 0 or 1 (i.e. the probability of a stop trial is small and the confidence about go stimulus identity, left or right, is high), and a large central *Wait region*,

where the value of p_d is close to 0.5 (go stimulus identity highly uncertain) or the value of p_s is large (probability of stop trial high); this topology makes intuitive sense. Figure 3.3 also shows that the optimal decision policy is time-dependent, such that the Go regions grow over time. This is primarily due to the time pressure of the impending response deadline [FY08].

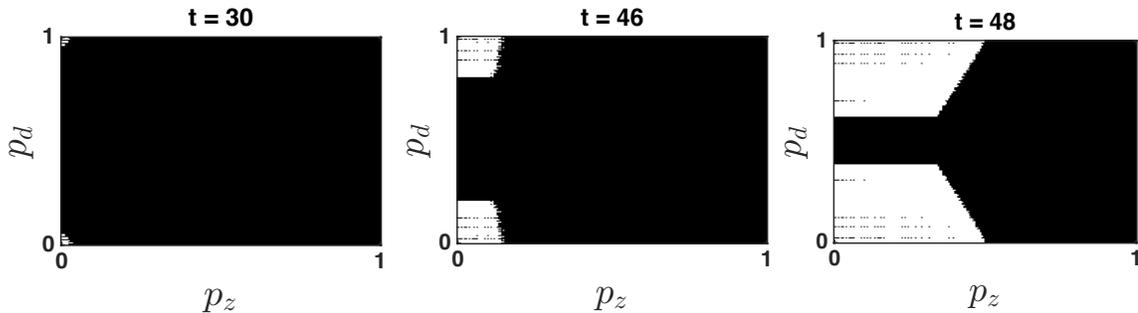


Figure 3.3: Dynamic evolution of the optimal policy map. The white and black areas denote the *Go* and *Wait* regions, respectively. Note that p_z denotes the probability that the stop signal has already occurred at time t , and is monotonically related to p_s , the probability that the current trial is a stop trial. The Bayesian update algorithm produces a belief state (p_d, p_z) at every time point t based on prior belief and all sensory data $\mathbf{x}_t, \mathbf{y}_t$ observed until time t . If the belief state falls into a Go region, a Go response is produced and the trial is terminated; otherwise, at least one more observation is obtained, and the location of the new Bayesian-updated belief state is compared to next time point’s optimal decision map. When/whether the belief state falls into a Go region determines when/whether the subject produces a response on that trial. The simulation shows that the Go regions expand over time, as the response deadline looms closer.

3.3 Results

3.3.1 Model Predictions

Classical behavioral results in the stop signal task, such as increases in stop error rate as a function of SSD and the generally faster SE RT compared to go RT,

have been shown to be natural consequences of such a rational decision-making process [SRY10, SY11], although these effects are also captured by the race model [LC84]. However, being a *computational* model in Marr’s framework of levels of analysis [Mar82], the optimal model is not only a model of the brain processes but also of the computational *task* the brain must solve – as such, it can also make *normative predictions* about how experimental manipulations of different task parameters should affect stopping behavior, since the experimental parameters are naturally represented as parameters of the Bayesian general model or of the objective function. In contrast, the race model cannot make such predictions, since it has no means of representing properties of the task itself.

Here, we specifically focus on the behavioral consequences of changing go stimulus discrimination difficulty. The race model does not represent stimulus difficulty explicitly and thus would not make any obvious predictions about behavioral consequences; moreover, since the go and stop processes are assumed to be independent, the race model would certainly not predict properties of the stop process, such as the SSRT, to change with go stimulus difficulty. On the other hand, in the optimal model, changes in the go stimulus difficulty would change the evolution of the sensory belief state (both via the empirical statistics and the Bayesian update rule, since we assume subjects to have the correct generative model), as well as the decision policy (the time-dependent mapping between the belief state and the action set), whose computation of Q_w^t involves an expectation over future belief state, which depends on the assumed likelihood function. Intuitively, we would expect a noisier go stimulus to slow down the general drift of the go stimulus posterior belief p_d^t toward 0 or 1 (depending on which true stimulus was shown), and hit the “Go region” later on a

go trial, or avoid hitting it altogether on a stop trial (Figure 3.3). Figure 3.4 show the simulated model predictions for the various behavioral measures as a consequence of changing go stimulus discrimination difficulty, parameterized by q_d in the model. We model q_d as monotonically decreasing (toward 0.5, which is a stimulus containing pure noise) for decreasing stimulus coherence. The exact values chosen for q_d in the simulations are as specified in Figure 3.4 caption. We find that the qualitative, monotonic relationships between the various predicted behavioral measures and the go stimulus coherence hold for a large range of q_d values chosen, as long as lower coherence corresponds to smaller q_d . As shown in Figure 3.4A;B;C, we expect Go RT to decrease, and Go discrimination and omission errors to decrease, as a function of increasing stimulus coherence. Correspondingly, the model predicts the stop error rate to increase as the stimulus coherence increases (Figure 3.4D), with the effect present at almost the whole range of SSD tested in the experiment (Figure 3.4E). Additionally, and perhaps more surprisingly, the model predicts the SSRT to decrease as a function of increasing go stimulus coherence. Although SSRT is not an intrinsic parameter or entity in the optimal model, one can nevertheless estimate SSRT as one does from empirical data, by identifying the SSD at which approximately 50% stop accuracy is achieved. This last prediction is particularly intriguing for differentiating the race model and the optimal model, as the race model would not predict that go stimulus difficulty should influence the speed at which the stop signal is processed.

3.3.2 Human Behavioral Data

In this section, we show that the model predictions in Section 3.4 are confirmed by the human behavioral data, where 8%, 15%, and 85% denote different coherences

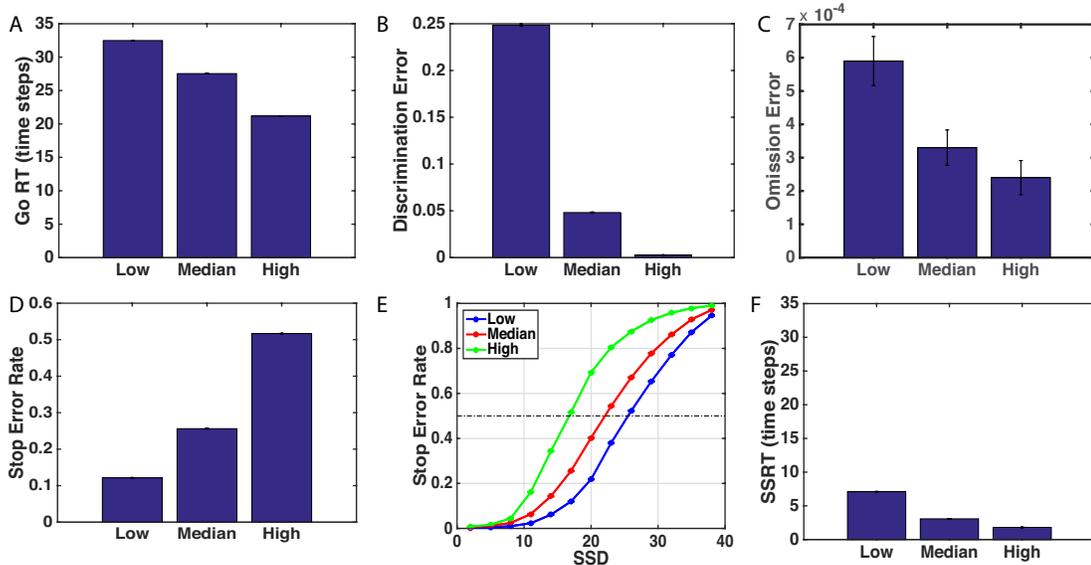


Figure 3.4: Simulated behavioral predictions by the optimal decision-making model. (A) Simulated mean Go RT decreases with higher go stimulus coherence (easier discrimination). *Low*, *median*, and *high* stimulus coherence are parameterized by different values of q_d : 0.55, 0.62, and 0.70, respectively. (B) Discrimination error rate decreases as coherence increases. (C) Omission error rate decreases with q_d . (D) Stop error rate increases with the q_d . (E) Average stop error rate as a function of SSD (known as the *inhibition function*) for different stimulus conditions. (F) SSRT decreased with q_d . Each data point is averaged over 100 simulated subjects, each performing 1000 go or stop trials. Error bars indicate standard error of the mean (sem); sem is extremely small and almost invisible for all but the omission error simulation data. Other parameters used in the model are adapted from [SRY10]: $r = 0.25$, $q_s = 0.72$, $D = 50$, $c_s = 0.4$, $c = 0.002$.

of random dot motion stimulus, while ‘X’ represents square versus circle go stimulus. Figure 3.5 shows the behavioral results from the experiment. Figure 3.5 A-C and F show that subjects’ mean Go RT, discrimination error rate, omission error rate, and SSRT decreased with coherence, as predicted in Figure 3.4. Figure 3.5 D shows that subjects’ stop error rate increases with coherence. We used one-sided paired t-test to test the significance of differences in behavioral measures across different go stimulus difficulties, e.g. \mathbf{H}_0 : mean (Go RT for 8%) = mean (Go RT for 15%), \mathbf{H}_1 : mean(Go RT for 8%) < mean (Go RT for 15%). We also conducted the more conservative

Wilcoxon rank test, which does not make the normality assumption that t-test makes, for completeness. As we detail in Supplementary Material, similar results are found using the two tests, except for the omission error, the trial type for which we have the least amount of data, since omission errors were rare. Here, we only discussed the results of the paired t-tests.

We found that mean Go RT significantly decreased as the coherence increased from 15% to 85% ($p = 0.008$, $t = 2.63$) and 8% to 85% ($p = 0.012$, $t = 2.44$), but not significant from 8% to 15% ($p = 0.263$, $t = 0.64$). In consideration of the long tail of the RT distribution (though this was ameliorated in the current study due to the response deadline), we computed the median Go RT of each subjects and then conducted paired t-test, which showed that median Go RT significantly decreased as coherence increased from 15% to 85% ($p = 0.004$, $t = 2.87$) and 8% to 85% ($p = 0.002$, $t = 3.20$), and showed a trend toward significance from 8% to 15% ($p = 0.09$, $t = 1.35$).

Paired t-tests for discrimination error rates were significant for all three cases, 8% to 15% ($p < 10^{-6}$, $t = 6.35$), 15% to 85% ($p = 0.002$, $t = 3.22$), and 8% to 85% ($p < 10^{-6}$, $t = 6.65$). The omission error rate only significantly decreased when the coherence increased from 8% to 85% ($p = 0.01$, $t = 2.5$), but not from 8% to 15% ($p = 0.16$, $t = 1.00$), not from 15% to 85% ($p = 0.22$, $t = 0.77$). Over all, the results suggest that Go RT, Discrimination Error and Omission Error decrease with coherence. In addition, the stop error rate also increased significantly when coherence increased from 8% to 85% ($p = 0.03$, $t = -1.89$), but not from 15% to 85% ($p = 0.14$, $t = -1.10$), and showed a trend toward significance from 8% to 15% ($p = 0.06$, $t = -1.61$). Altogether, the results suggest that stop error rate increases with coherence.

We used 'smoothspline' function in Matlab to fit the inhibition function of each

subject in each stimulus type and estimated the corresponding SSRT (Figure 3.5F) as the difference between median Go RT and the SSD at which 50% stop error rate is committed (Figure 3.5E). According to the paired t-test, SSRT significantly decreased as the coherence increased from 8% to 85% ($p = 0.02$, $t = 2.41$), but not from 8% to 15% ($p = 0.23$, $t = 0.79$), and showed a trend toward significance from 15% to 85% ($p = 0.09$, $t = 1.40$). Over all, the results suggest that SSRT decreases with coherence.

We also included a more classical circle versus square go discrimination task, because stop-signal tasks typically use highly discriminable go stimuli such as circle versus square. As we are among the first to use the random-dot coherent motion stimuli for the go task, as well as the first to systematically degrade the go stimulus in the stop-signal task, we wanted to make sure that the easiest condition (85% coherence) produces comparable data to the more commonly used shape discrimination task. Figure 3.5 shows that this is indeed the case across the behavioral measures we examined.

3.4 Discussion

In this work, we investigated the computational and behavioral consequences of manipulating stimulus discriminability in the stop-signal task. We simulated our previously proposed optimal decision-making model [SY11] to derive behavioral predictions, and presented novel experimental data that broadly validated these predictions. Interestingly, the SSRT, which is thought to reflect the stopping ability of the subject, is found to significantly decrease with increasing difficulty of the go stimulus discrimination task. This directly contradicts the independence assumption of the race model [LC84], as well as its more complex variants that assume a simple

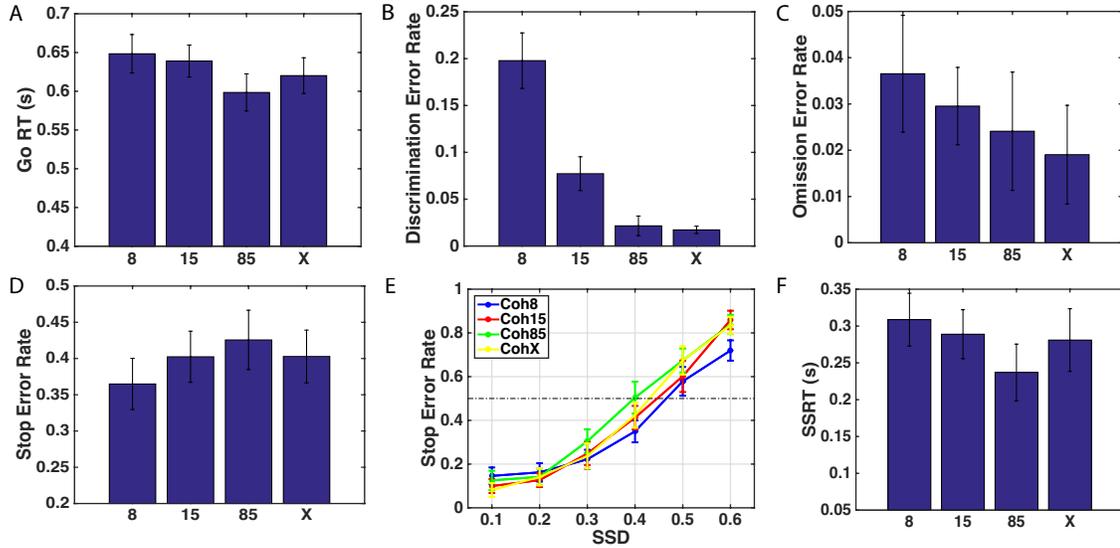


Figure 3.5: Behavioral data under varying go stimulus discrimination difficulty. 8, 15, 85 denote different coherences of random dot motion stimulus. X denote square versus circle discrimination task. (A) Mean Go RT decreased as the discrimination task became easier. Bar height indicates the mean of median Go RT (for each subject) for each condition. (B) Go discrimination error rate decreased as coherence increased. (C) Go omission error rate decreased with coherence. (D) Stop error rate increased with coherence. (E) Average inhibition function for different stimulus types. (F) SSRT significantly decreased with coherence.

inhibitory interaction between the go and stop processes that is independent of the go stimulus discriminability [BPLS07]. Together, our results imply that there exist intrinsic and complex interactions between *go* and *stop* processing, much as that postulated by our optimal decision-making model for stopping behavior [SY11]. More generally, the broad concurrence between model predictions and behavioral data demonstrate the normative predictive power of the optimal model, as well as the specific model assumptions that humans readily internalize environmental statistics and adopt decision policies that are normative and context-sensitive.

The difference between the race model and the optimal decision-making model is not only that of complexity or the nature of interactions between the stop and go

processes, but also that of levels of analysis, in the parlance of David Marr [Mar82]. The race model is primarily an algorithmic model, while the optimal model is primarily a computational model. That means for a more meaningful comparison, it would be worthwhile to consider an algorithmic description of the optimal model that is more directly comparable to the race model (or, alternatively, a computational description of the race model, which is harder to obtain). Such an analysis was done in a previous paper [SY11], which showed that while the model would not be able to *predict* changes in stopping behavior as a consequence of changes in the reward structure of the task, the parameters of a diffusion-model implementation of the race model (see [SY11] for more details) can be fit to different experimental conditions in a *post hoc* manner in order to capture qualitative changes in behavior. Likewise, one could fit parameters of a diffusion model equivalent of the race model to "capture" behavioral changes as a function of go stimulus signal-to-noise ratio, but again, it would be a *post hoc* result, not a normative predictive process as the optimal decision-making model excels in.

In the current paper, the parameter q_d , which specifies the noisiness of the sensory data related to the go stimulus, is left as a free parameter. While we kept q_d monotonically increasing as a function of increasing stimulus coherence (a rational choice), its values for different coherence conditions were somewhat arbitrarily chosen. Although the qualitative nature of the model predictions (changes in go RT, stop error rate, and SSRT as a function of go stimulus coherence) is relatively robust with respect to the precise choice of q_d (very close q_d would not result in significantly different behaviors), an even better approach would be to fit q_d for each subject in a pure 2AFC task, identical to the stop-signal task except for the total absence of the stop signal, as 2AFC behavior is fairly well captured and understood as a variant of the

sequential probability ratio test [GS02, BBM⁺06, FY08, SY12, DY12], which can be parameterized by essentially the same q_d variable. In that case, q_d would then not be a free parameter in the stop-signal task but one derived from a separate 2AFC session for each subject, and we would then be able to see whether stopping behavior really follow quantitatively from the optimal inference and decision-making process, as predicted by the model.

Another important direction of future research is a better theoretical understanding of the algorithmic aspects of the optimal model, in particular what determines SSRT in this model. While SSRT is not intrinsic to the optimal model, as it is in the race model, it is nevertheless possible to "measure" SSRT for the optimal model based on simulated trial outcomes, just as is done empirically for human data. Related to this, it is unclear *why* SSRT should decrease in the optimal model for increasing coherence, which we showed to be both predicted by the optimal model and exhibited by human data in this paper. Specifically, as the go task gets easier, subjects make more stop errors due to faster Go RT, even though SSRT decreases – it just does not decrease sufficiently to counter the faster Go RT. Future work is needed to better understand the nature SSRT in the context of the optimal model, as well as, in general, a better algorithmic understanding of the relationship between the optimal model and the race model.

Although this work was mostly focused on computational modeling and behavioral data analysis, it has implications for the neuroscientific study of inhibitory control as well. The race model has helped to advance the neuroscience of inhibitory control, by relating neural activities in various brain regions, such as the frontal eye field [HSP98] and superior colliculus [PH03], to the go and stop processes. But

as the race model does not address how different cognitive processes contribute to stopping behavior, it is also limited in its ability to anticipate or explain cognitive modulations of neural activities involved in inhibitory control. Given the ability of the computationally more sophisticated optimal model to explain a wider range of behavioral data, we can expect that it will also lead to novel and interesting interpretations of neural activation patterns related to inhibitory control, and perhaps guide future neuroscientific experimentation. Indeed, we have already used the optimal model to identify a brain region (anterior cingulate cortex) as having fMRI BOLD response consistent with encoding an unsigned prediction error ("Bayesian surprise") related to the prior belief of whether the upcoming trial will be a stop or go trial [ISYL13], and shown that this prediction signal is altered in young adults at risk for developing stimulant addiction [HSS⁺14], a condition known to be associated with impaired inhibitory control and specifically stopping behavior. Prior to this model-based fMRI study, it was thought that the anterior cingulate cortex was one of many areas generally involved in preparing or executing the "go" response. In the context of the optimal model, we now know that this area, unlike the other cortical areas, is specifically involved in reporting the surprise signal, which just happens to be greater on stop trial than go trial on average, because stop trials are generally rare. This provides just one example of how a statistically sophisticated model facilitates a richer and more theory-driven exposition of the neural basis of inhibitory control.

Acknowledgments

This chapter, in full, is a reprint of the material from Ning Ma and Angela J. Yu, Inseparability of Go and Stop in Inhibitory Control: Go Stimulus Discriminability

Affects Stopping Behavior , *Frontiers in Neuroscience*, 10(54), 2016. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Stop Paying Attention: The Need for Explicit Stopping in Inhibitory Control

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.

4.1 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 [LC84]. 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 [LC84, BPLS07], which is a mechanistic account that posits a race to threshold between independent *go* and *stop* processes (See Figure. 4.1). 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).

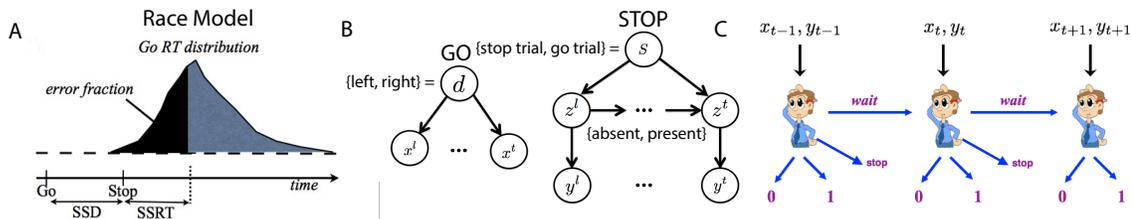


Figure 4.1: Models for inhibitory control in the stop-signal task. (A) The classical race mode posits that the behavioral outcome (go or stop) is determined by a race between two independent go and stop processes. (B) Bayesian graphical model for noisy sensory data generation in the rational decision-making 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 [SY11], 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 [SY11, ISYL13, MY15a, MY16]. 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 [ISYL13], as well as how that neural representation becomes altered in users of stimulants such as cocaine and methamphetamine [IHZ⁺15, HSS⁺14, HSZ⁺15].

One critical assumption made by the previous model [SY11] 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 [ADE⁺07] and neurophysiology data in monkeys [HSP98], 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 (Section 4.2), then show how the model captures a variety of behavioral phenomena observed in the stop signal task (Section 4.3.2), as well as neural data (Section 4.3.3) 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 (Section 4.4).

4.2 The Model

As in the earlier MDP model [SY11], this MDP model consists of two key components, a monitoring component 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.

4.2.1 Monitoring Process as Bayesian Statistical Inference

We use the same Bayesian inference model proposed in [SY11] to implement the sensory processing component, and thus provide only a short overview here.

Figure 4.1B 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 θ . For simplicity, we assume that the onset of the stop signal θ 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' 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 $\mathbf{b}^t = (p_d^t, p_s^t)$.

4.2.2 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 [SY11], 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 τ_r denote the trial termination time and τ_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:

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

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, $L_\pi = \mathbb{E}[l(\tau_r, \tau_a, \delta; d, s, \theta, D)]$,

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

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

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

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), $Q_g^t(\mathbf{b}^t)$, $Q_s^t(\mathbf{b}^t)$, and $Q_w^t(\mathbf{b}^t)$, respectively. Using the Bellman optimality equa-

tion for optimal action-value function- $Q(\mathbf{b}, a) = \mathbb{E} [c_a + c_r + V(\mathbf{b}^{t+1}) | \mathbf{b}^t = b, a_t = a]$, we can obtain the three Q-factors

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

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 future 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, δ , is 1 if $p_d^t > 0.5$ 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} .

$$\begin{aligned} \mathbb{E} [V^{t+1}(\mathbf{b}^{t+1})|\mathbf{b}^t] &= \sum_{x^{t+1}, y^{t+1}} p(x^{t+1}, y^{t+1}|\mathbf{b}^t) V^{t+1}(\mathbf{b}^{t+1}(\mathbf{b}^t, x^{t+1}, y^{t+1})) \\ p(x^{t+1}, y^{t+1}|\mathbf{b}^t) &= p(x^{t+1}|p_d^t) p(y^{t+1}|p_s^t) \\ p(x^{t+1}|p_d^t) &= p_d^t f_1(x^{t+1}) + (1 - p_d^t) f_0(x^{t+1}) \\ p(y^{t+1}|p_s^t) &= (p_z^t + (1 - p_z^t) h(t + 1)) g_1(y^{t+1}) \\ &\quad + (1 - p_z^t) (1 - h(t + 1)) g_0(y^{t+1}) \end{aligned}$$

In the simulations, we discretize the space of p_d^t and p_z^t each into 200 bins.

4.3 Results

4.3.1 Model Simulations

Figure 4.2A 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 [SY11], 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.

Figure 4.2B shows the simulated evolution of different Q-factors, or the expected cost of taking the three actions (Go, 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 .

4.3.2 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 [SY11].

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 [SY11]; Figure 4.3A 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 [LW09], a phenomenon shown to naturally arise when c_s is increased in the earlier MDP model [SY11]; Figure 4.3B-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 (Figure 4.3B) subjects can be expected to respond faster, (Figure 4.3C) make fewer errors, and (Figure 4.3D) 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 settings of c_a and c_r .

4.3.3 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 [HSP98]. 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 sub-populations 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 [HSP98].

Figure 4.4 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 Figure 4.4A, 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. Figure 4.4 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. Figure 4.4C 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 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. Figure 4.4D 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.

4.4 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 [SY11] 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 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 [LC84]. 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 [BPLS07], 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 [HSP98]. 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 as 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 and earlier, 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 [SY11], may not be fully correct in describing the cognitive and neural processes underlying inhibitory control, they exemplify a powerful new 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

This chapter, in full, is a reprint of the material from Ning Ma and Angela J. Yu, Stop paying attention: the need for explicit stopping in inhibitory control, Proceedings of the 38th Annual Conference of the Cognitive Science Society, 2016. The dissertation author was the primary investigator and author of this paper.

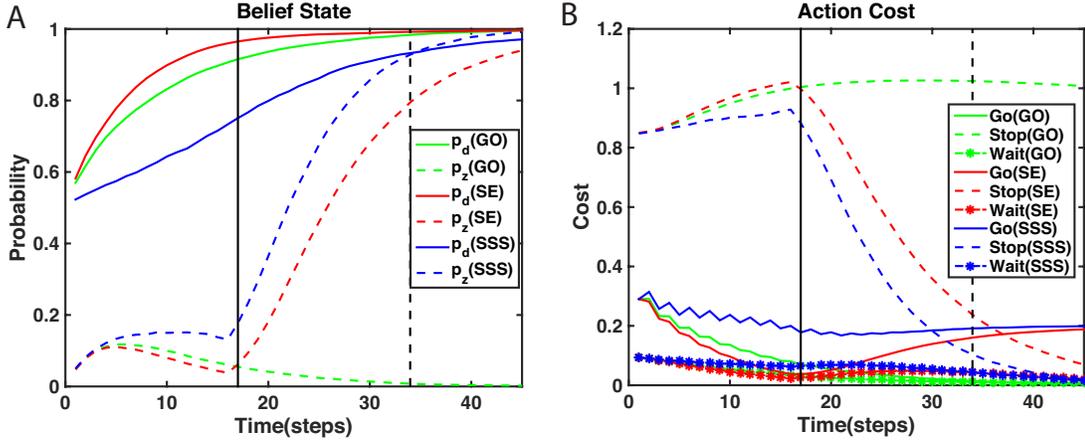


Figure 4.2: Mean Belief state and Q-factors. (A) Evolution of the average belief states p_d (solid line) and p_z (dashed line) for different trials-GO(green):go trials, SSS(blue): stop trials successfully stopped by choosing explicit stop action, SE (red): stop error trials. In SE trials, the go stimulus 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 [SY11]: $q_d = 0.68$, $q_s = 0.72$, $q = 0.2$, $r = 0.25$, $D = 50$ steps, $c_s = 0.2$, $c_r = 0.002$, $c_a = 0.002$. Unless otherwise stated, there parameters are used in all the subsequent simulations.

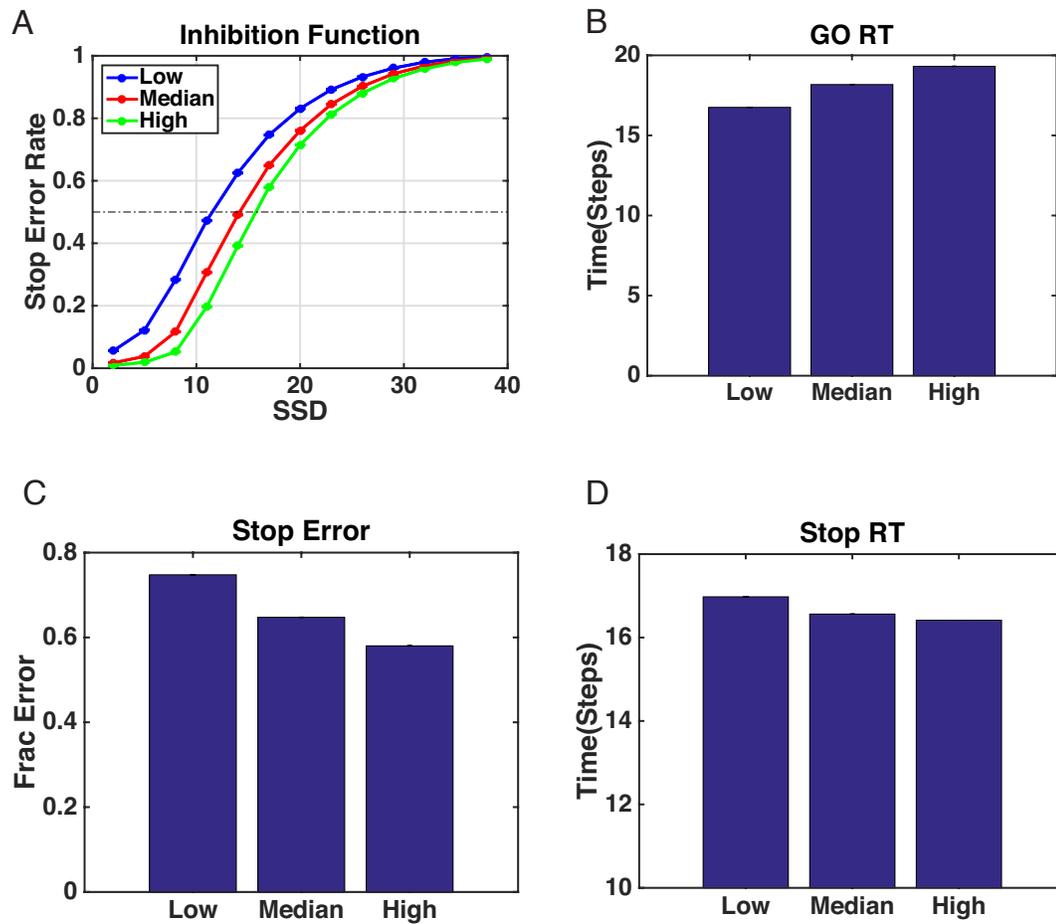


Figure 4.3: Influence of reward/motivation on stopping behavior. Results are averaged over 50 simulated subjects, each performing 10000 go and stop trials. Error bars (s.e.m.) are too small to be observed. (A) Model simulation produces an inhibition function (frequency of SE as a function of SSD) similar 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.

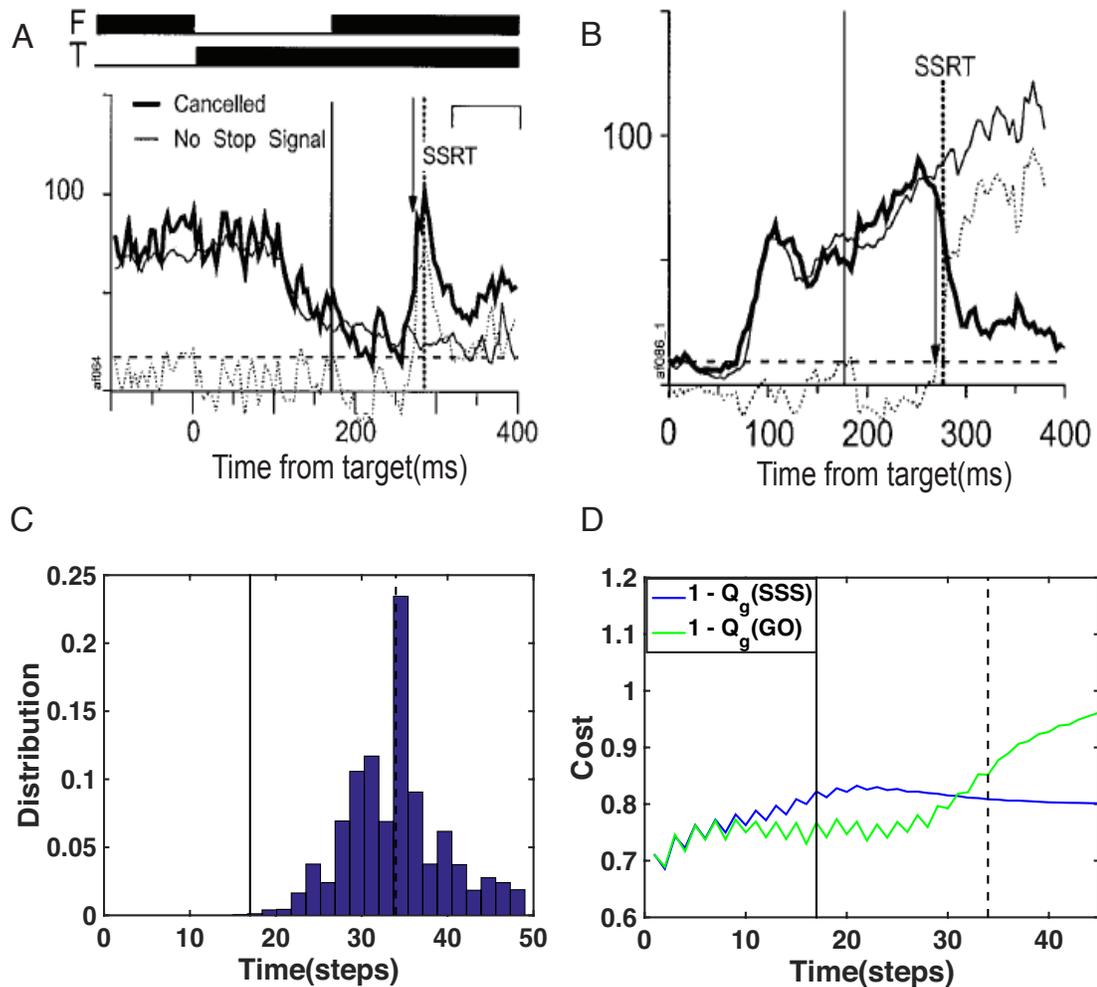


Figure 4.4: Neural representation of action values. (A) Average firing rate of a fixation neuron in the frontal eye field (FEF) of a monkey performing a saccade version of stop-signal task peaks at the time of the behaviorally measured SSRT (Hanes et al, 1998), suggesting fixation neurons encode a Stop action signal. (B) Movement neurons in the FEF diverge, between stop-success trials (Cancelled) 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 ($1 -$ expected cost) of taking the Go action diverge around SSRT on SS/Go trials, much as movement neurons do.

Chapter 5

Reduced Neural Recruitment for Bayesian Adjustment of Inhibitory Control in Methamphetamine Dependence

Delineating the processes that contribute to the progression and maintenance of substance dependence is critical to understanding and preventing addiction. Several previous studies have shown inhibitory control deficits in individuals with stimulant use disorder. We used a Bayesian computational approach to examine potential neural deficiencies in the dynamic predictive processing underlying inhibitory function among recently abstinent methamphetamine-dependent individuals (MDIs), a population at high risk of relapse. Sixty-two MDIs were recruited from a 28-day inpatient treatment program at the San Diego Veterans Affairs Medical Center and compared with 34 healthy comparison subjects. They completed a stop-signal task during functional

magnetic resonance imaging. A Bayesian ideal observer model was used to predict individuals' trial-to-trial probabilistic expectations of inhibitory response, $P(\text{stop})$, to identify group differences specific to Bayesian expectation and prediction error computation. Relative to comparison subjects, MDIs were more likely to make stop errors on difficult trials and had attenuated slowing following stop errors. MDIs further exhibited reduced sensitivity as measured by the neural tracking of a Bayesian measure of surprise (unsigned prediction error), which was evident across all trials in the left posterior caudate and orbitofrontal cortex (Brodmann area 11), and selectively on stop error trials in the right thalamus and inferior parietal lobule. MDIs are less sensitive to surprising task events, both across trials and upon making commission errors, which may help explain why these individuals may not engage in switching strategy when the environment changes, leading to adverse consequences.

5.1 Introduction

Amphetamine-type stimulants, which include methamphetamine, are the fastest rising drug of abuse worldwide and have become the second most widely used class of illicit drugs worldwide [LBM⁺10, WRT⁺13]. Moreover, methamphetamine dependence (MD) is associated with high likelihood of relapse [MD14]. Whereas addiction research has heavily focused on reward processing, executive deficits have been consistently observed in stimulant abusers and implicated in the progression of abuse to dependence [JKJ⁺14, VGR⁺12, MSM05]. Identifying precise neurocognitive processes of such deficits may therefore not only improve our understanding of how neurochemical changes in MD affect decision making, but also help identify robust neural predictors of relapse and treatment response.

Pharmacological, lesion, and neuroimaging studies suggest that neural alterations within the frontostriatal pathways, which appear to persist even during abstinence periods, may underlie the cognitive deficits observed in methamphetamine-dependent individuals (MDIs) [KE07, MT03, JJ99]. MDIs exhibit reduced integrity of dopaminergic and serotonergic neurons in dopamine-rich regions, including the anterior cingulate cortex [TRY⁺05, RTY⁺07], striatum [LDTN07], and limbic areas [PKS⁺04], and lower glucose metabolism in the striatum [YYM⁺02, NLG⁺01, LDJE11] and frontocingulate areas, including anterior cingulate, orbitofrontal, and dorsolateral prefrontal cortices [YYY⁺03, YSD⁺09, ESS⁺04]. Such neural patterns have been linked to inhibitory and impulse-control deficits on standard interference/Stroop tasks [SAX⁺10, GJK⁺11] and in delay discounting [WMR⁺06, KN04, MDJM07]. MDIs further show difficulties in detecting trends and integrating new information to predict future outcomes during decision making [RBA⁺99, MNB⁺02], which may be particularly hindering within changing environments, in which one has to constantly monitor information to know what to expect. Such learning impairment could contribute to MDIs' deficits in inhibitory control and other types of dynamic decision making [JM07]. However, the cognitive processes underlying the relationship between neural damage and substance use in MDIs remain poorly understood.

In recent work, we showed that healthy individuals [JPAC13] and nondependent occasional stimulant users [KPJ⁺14] continuously alter their response strategy in a standard inhibitory paradigm (stop-signal task), such that dynamic fluctuations in their reaction time (RT) and error rate are consistent with a particular Bayesian belief updating [AJ08] and decision strategy [PJ11]. Here, we use the same Bayesian approach combined with event-related functional magnetic resonance imaging (fMRI)

to model individuals' real-time expectations of response inhibition need in the stop-signal task. This strategy allows us to identify any difference between MDIs and healthy comparison subjects (CSs) in their neural representation of trialwise expectations of inhibitory response and Bayesian prediction errors needed for updating those expectations.

Based on the reduced functional metabolism in prefrontal, anterior cingulate, and striatal areas observed in MDIs, and given the consistent involvement of these brain regions in encoding action-related expectations, value, and prediction errors [KPJ⁺14, STJ11, LTW06, MJI⁺07], we hypothesized that inhibitory dysfunction in MDIs would be characterized by attenuated neural representation of the expectation of an inhibitory signal, as well as altered prediction error signals, coding the discrepancy between predicted and actual outcomes, which is critical for adjusting expectations and adaptive behavior to potentially adverse consequences.

5.2 Methods and Materials

5.2.1 Participants

The University of California, San Diego Human Research Protections Program approved the study protocol and all participants gave written informed consent. A total of 62 (21% female) recently abstinent (i.e., within the last month) MDIs were recruited from a 28-day inpatient Alcohol and Drug Treatment Program at the Veterans Affairs San Diego Healthcare System and Scripps Green Hospital.¹ In addition, 34 healthy CSs (30% female) were recruited via flyers, internet ads (e.g., Craigslist), and local university newspapers. CSs were selected to be matched in age and IQ with

MDIs. All subjects completed a clinical interview and a fMRI session during which they completed the stop-signal task (between the third and fourth week of treatment for MDIs). Lifetime DSM-IV Axes I and II diagnoses, including substance abuse and dependence [APA13], were assessed by the Semi-structured Assessment for the Genetics of Alcoholism II [KRC⁺94]. Diagnoses were based on consensus meetings with a clinician specializing in substance use disorders (MPP).

5.2.2 Stop-Signal Task

Participants completed six blocks of a stop-signal task (75% go trials in each block of 48 trials) while undergoing fMRI. Trial order was pseudo-randomized throughout the task and counterbalanced. On go trials ($n = 216$, or 36/block), they had to press as quickly as possible the left button when an “X” appeared and the right button when an “O” appeared. On stop trials (i.e., whenever they heard a tone during a trial, at some time subsequent to the presentation of the go stimulus; $n = 72$), they were instructed not to press either button (see Figure 5.1 A). Each trial lasted ≈ 1300 ms or until the participant responded, with a 200-ms interstimulus interval. Prior to scanning, participants’ mean go reaction time (i.e., average response latency [ARL]) from stimulus onset was determined to compute six levels of stop-signal delay (SSD), providing an individually customized range of difficulty. Such individual measures were used to determine the stop signal delay for the six different stop trial types, providing a subject-dependent jittered reference function. Specifically, stop signals were delivered at 0 (ARL-0), 100 (ARL-100), 200 (ARL-200), 300 (ARL-300), 400 (ARL-400), or 500 (ARL-500) ms less than the mean reaction time after the beginning of the trial, thus providing a range of difficulty level; ARL ranged from 504 to 925 ms

(mean = 664 ms, SD = 112 ms) [for more details see [KPJ⁺14, SAEM05]].

5.2.3 Bayesian Model of Probabilistic Prediction

In recent work [JPAC13, PJ11, PRA11], sequential effects in the stop-signal paradigms, where recently experienced stop trials tend to increase RTs on a subsequent go trial and decrease error rate on a subsequent stop trial, have been shown to be well captured by a Bayes-optimal decision-making model. This Bayesian hidden Markov model adapted from the dynamic belief model [AJ08] (see Figure 5.1 B) assumes that an individual updates the previous probability of encountering stop trials, $P(\text{stop})$, on a trial-by-trial basis based on trial history and adjusts decision policy as a function of $P(\text{stop})$, with systematic consequences for go RT and stop accuracy in the upcoming trial. A higher predicted $P(\text{stop})$ is associated with a slower go RT and a higher likelihood of correctly stopping on a stop trial in healthy subjects [JPAC13, KPJ⁺14]. Briefly, the model assumes that the stop-signal frequency r_k on trial k has probability α of being the same as r_{k-1} and probability $1 - \alpha$ of being resampled from a previous beta distribution $p_0(r)$. The probability of trial k being a stop trial, $P_k(\text{stop}) = P(s_k = 1 | \mathbf{S}_{k-1})$, where $\mathbf{S}_k = (s_1, \dots, s_k)$ is 1 on stop trials and 0 on go trials, can be computed as follows:

$$\begin{aligned} p(s_k = 1 | \mathbf{s}_{k-1}) &= \int P(s_k = 1 | r_k) p(r_k | \mathbf{S}_{k-1}) dr_k \\ &= \int r_k p(r_k | \mathbf{S}_{k-1}) dr_k = \langle r_k | \mathbf{S}_{k-1} \rangle \end{aligned}$$

The predictive probability of seeing a stop trial, $P_k(\text{stop})$, is the mean of the predictive distribution $p(r_k|\mathbf{S}_{k-1})$, which is a mixture of the previous posterior distribution and a fixed previous distribution, with α and $1 - \alpha$ acting as the mixing coefficients, respectively:

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

with the posterior distribution being updated according to Bayes's rule:

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

Thus, to capture the persistent sequential effects, the model assumes that subjects continually update $P(\text{stop})$ with effectively the same learning rate, because they believe that the true rate of stop trials is undergoing changes in the environment: formally, with probability α , it stays the same as last trial, and with probability $1 - \alpha$, it is redrawn randomly from the generic previous distribution p_0 . A larger α corresponds to a belief in a less volatile environment, and therefore a longer time window during which previous trials can affect future $P(\text{stop})$ calculations, which results in smaller sequential effects due to the most recently experienced go/stop trial (relative to a longer relevant time history) [AJ08, PRA11].

In the present study, parameters for the beta distribution $p_0(r)$ and α were kept constant across all subjects and were based on simulations that sought to optimize behavioral fit at the group level—that is, maximizing R^2 for the regression of $P(\text{stop})$ on RT. Such optimal parameters were $p_0 = \text{beta}(a = 2.5, b = 7.5; s = a + b = 10; \text{mean} = 0.25)$ and $\alpha = 0.5$. The fixed values were optimal for both groups. Given these

parameters and sequence of observed stop/go trials experienced by participants, we computed the corresponding sequence of subjective $P(\text{stop})$. In subsequent fMRI analyses, the trial-by-trial estimation of $P_k(\text{stop}) = P(s_k = 1 | \mathbf{S}_{k-1}) = \langle r_k \rangle$ (i.e., most up-to-date estimate of stop trial likelihood based on all previous trials) was used as a parametric regressor.

In previous work [AJ08], we showed that the Bayesian belief updating model for $P(\text{stop})$ can be approximated by a linear-exponential filter of past observations, such that whether the previous trial was a stop trial ($s_k = 1$) or a go trial ($s_k = 0$) linearly contributes to the estimated $P(\text{stop})$ for the current trial, and the coefficient for each trial decays exponentially into the past. This model is also equivalent to a version of a delta rule, where $P(\text{stop})$ on each trial is an appropriately tuned linear combination of $P(\text{stop})$ on the last trial, how the most recent observation ($s_k = 1$ or 0) differs from the last $P(\text{stop})$, in other words, a signed prediction error ($SPE_k = s_k - P_k(\text{stop})$), and a constant bias term, which may be realistically implemented at the neural level [AJ08].

5.2.4 Behavioral Statistical Analyses

We applied hierarchical generalized mixed-effect linear models to participants' trial accuracy (stop-success [SS] vs. stop-error [SE]) and RTs (dependent variables), treating subject as a random effect (with varying intercepts and slopes, unstructured variance/covariance assumed) and other independent variables as fixed effects [RDD08]. The first set of models for go RTs used a linear mixture of Bayesian model-based estimate of $P(\text{stop})$, group, previous trial type, and previous trial SSD. Go trials with a reaction time >1300 ms were automatically counted as go errors and were not

included in those analyses. The second set of models for error data used a logit link function [TF08] in terms of linear mixtures of SSD, group, and $P(\text{stop})$. We report change in log-likelihood ratio (following a chi-squared distribution) and regression coefficients (when applicable) with associated t test and p values.

5.2.5 fMRI Analyses

Using a fast event-related fMRI design, six T2*-weighted echo planar imaging functional runs were collected for each participant, along with one T1-weighted anatomical image (see the Supplement for image acquisition and preprocessing details). Preprocessing and subsequent fMRI analyses were conducted using Analysis of Functional NeuroImages (AFNI) software [RW96].

First-Level Analyses. Three types of trials were distinguished (go, SS, and SE; go error trials were scarce and not included in these analyses), included as predictors in a general linear model (GLM), and convolved with a canonical hemodynamic response function. They were entered as both categorical linear regressors (multiplied by the mean of the computed $P(\text{stop})$ probabilities across all trials) and parametric regressors (modulated by $P(\text{stop})$) [KPJ⁺14, CAGK98] in order to isolate neural activations associated with $P(\text{stop})$ independently of categorical trial type neural coding (i.e., categorical regressors). This model therefore included six task regressors (three categorical: go, SS, SE; and three model-based parametric: Go x $P_k(\text{stop})$, SS $P_k(\text{stop})$, SE x $P_k(\text{stop})$). To assess group differences in the updating processes related to $P(\text{stop})$, we created a second GLM with trialwise Bayesian signed prediction error [i.e., SPE: (outcome - $P(\text{stop})$)] and unsigned prediction error [i.e., UPE: |outcome - $P(\text{stop})$]| as parametric regressors of interest. This second model also included a

parametric regressor modeling trial error (0 = correct or 1 = error) to control for performance error-related activity [JPAC13]. Given the fixed parameter setting and pseudorandomized sequence of trials, the P(stop) and Bayesian UPE/SPE values were the same for all participants in these GLMs. Both GLMs included a baseline regressor (consisting of intertrial intervals), instruction phases, linear drift, and three motion regressors [pitch, yaw, roll [SAEM05]], go RTs, and SSD as parametric regressors of no interest. Images were spatially filtered (Gaussian full width half maximum 4 mm) to account for individual anatomical differences. Anatomical and functional images were manually transformed into Talairach space.

Second-Level Analyses. At the between-subject level, the coefficients of our first-level GLM were modeled with voxel-wise mixed-effects linear models analyses, performed with R statistical software (R Foundation for Statistical Computing, Vienna, Austria) [JDSD11], in terms of a linear mixture of subject-level effects and group (CS, MDI). Specifically, we tested for second-level effects of group and its interaction with P(stop) under each trial type (Go x $P_k(\text{stop})$, SS x $P_k(\text{stop})$, SE x $P_k(\text{stop})$), with subject treated as a random effect.³ In the first analysis, we isolated P(stop)-modulated activations for go versus stop trials (SS and SE were averaged). Whole brain statistical maps were obtained for the group main effect (reflecting areas tracking previous P(stop) values irrespective of trial type or accuracy) and the group x P(stop) modulated trial type interaction. We then conducted an additional contrast on stop trials only, comparing P(stop) modulated activation for SS versus SE trials and obtained statistical maps for the group x P(stop) trial type (SE vs. SS) interaction. To correct for multiple comparisons, we used a cluster threshold adjustment based on Monte Carlo simulations (generated with AFNI's 3dClustSim program); AFNI,

National Institutes of Mental Health, Bethesda, MD), based on whole-brain voxel size and 4-mm smoothness. A minimum cluster volume of 768 muL was used, with a minimum voxelwise significance of $p < .005$, corrected for multiple comparisons at familywise error rate = $p < .01$.

Group Difference in Bayesian Prediction Errors. Based on previous work [JPAC13, KPJ⁺14], we further selected from areas identified in the interaction contrast of the above-mentioned analyses [interaction of group and P(stop)-modulated trial type, i.e., Go x P(stop), Stop x P(stop)] those that were consistent with a group difference in either type of Bayesian prediction errors (UPE and SPE). Specifically, we identified those regions with nonzero P(stop) activations in CSs of opposite signs and same signs across go and stop trials, reflecting UPE and SPE activations respectively.

5.3 Results

5.3.1 Subject Characteristics

MDIs did not differ from CSs in ethnicity, sex, age, and verbal IQ ($p > .05$). On average, CSs had a higher level of education ($p < .001$). MDIs endorsed greater cocaine and cannabis intake ($p < .001$), and they used alcohol and nicotine more frequently and in larger quantities than did CSs ($p < .054$) (see Table 5.1).

Table 5.1: Participants' Characteristics as a Function of Group Status ($n = 96$). IQ, intelligence quotient; N/A, not applicable; WTAR, Wechsler Test of Adult Reading. ^at test computed using natural log transformed + 0.5 values (due to nonnormal distributions) replicated results for raw data.

	Methamphetamine-Dependent Individuals ($n = 62$)		Comparison Subjects ($n = 34$)		t Test, p Value
	Mean	SD	Mean	SD	
Demographics					
Age, years	38.0	10.4	36.1	11.1	.39
Education, years	13.0	1.7	14.8	1.6	< .001
Verbal IQ, WTAR	109.1	8.7	111.6	9.7	< .26
Alcohol, typical drinks/week	22.1	38.6	3.9	4.5	< .05
Alcohol, typical days/week	3.1	3.0	1.9	2.2	< .05
Nicotine, typical cigarettes/day	14.0	8.6	1.5	4.0	< .05
Nicotine, typical days/week	6.2	2.2	1.2	2.5	< .05
Lifetime Drug Use					
Methamphetamine	14,267.8	29,028.0	0.0	0.0	N/A
Cocaine	2560.5	6064.4	1.3	4.6	<.001 ^a
Prescription stimulant	56.7	417.5	0.0	0.2	.25 ^a
Cannabis	8853.2	25,738.0	40.0	168.9	< .001 ^a

5.3.2 Behavioral Performance and Model-Based Behavioral Adjustment

Reaction Times. Consistent with our model's assumptions [JPAC13, PJ11, PRA11], a positive linear relationship between go RT and P(stop) was observed across both groups ($B = 274$ ms, $t_{94} = 4.6$, $p < .001$, model omnibus test: $\chi_1^2 = 19.4$, $p < .001$; mean Pearson correlation coefficient: $r = .15$, adjusted $R^2 = .05$). The group main effect on go RT was not statistically significant ($\chi_1^2 = 0.48$, $p = .49$;

meanRT:CS = 624ms; MDI = 590ms) (see Figure 5.2 A for go RT distributions). The P(stop) x group interaction was marginally significant ($\chi_1^2 = 3.8, p = .05$), showing a trend for smaller positive slope for RT as a function of P(stop) in CSs. Whereas the regression slope was steeper in MDIs, we surmise that this pattern relates to the wider RT range (more short RTs) and the larger sample size in MDIs, rather than reflecting a meaningful group difference in model fit and related neural processes. A positive linear relationship between go RT and P(stop) was observed within each group (CS:B = 105, $p < .01$; $\chi_1^2 = 7.7, p < .01$; MDI: B = 365, $p < .001$; $\chi_1^2 = 144, p < .001$). For illustration of the linear trends, Figure 5.2 B shows data collapsed across all subjects for MDIs and CSs separately, where go trials were binned by P(stop) and average RT calculated for each bin separately. MDIs and CSs did not differ in stop signal reaction time (mean CS = 239 ms; mean MDI = 231 ms, $t_{94} = .49, p = .61$).

Post-stop Slowing. An overall slowing in go RT was observed following stop relative to go trials ($\chi_1^2 = 124, p < .001$), which was true for both SS ($B = +17, p < .001$) and SE ($B = +35, p < .001$) trials. Moreover, there was a group by previous trial type interaction ($\chi_1^2 = 41, p < .001$). Specifically, CSs had similar RTs following SS and go trials ($B = 6, t_{33} = -1.3, p = .18$) but exhibited significant slowing following SE trials ($B = 16, t_{32} = 2.9, p < .01$). In contrast, MDIs were overall slower following SS relative to go trials ($B = 31, t_{62} = 5.4, p < .001$) but did not exhibit additional slowing after SE relative to SS trials ($B = 10, t_{62} = -1.5, p = .19$) (see Figure 5.2 C).

Performance Accuracy. As expected, participants had a higher likelihood of error on trials with longer SSD (odds ratio = 2.6, Wald $z = 39, p < .001$; omnibus test: $\chi_1^2 = 250, p < .001$). The group difference in stop error rates did not reach

statistical significance (group main effect: $\chi_1^2 = 2.3$, $p = .13$; mean error rates: CS = 0.44; MDI = 0.49); however, the group by SSD interaction was significant ($\chi_1^2 = 4.8$, $p < .05$) with higher error likelihood in MDI for longer SSD (i.e., more difficult SSD) (see Figure 5.2 D). Moreover, as predicted by our model [JPAC13, PJ11, PRA11], we found a negative relationship between error likelihood and P(stop), with higher P(stop) overall prompting a smaller likelihood of error (odds ratio = 0.10, Wald $z = -3.38$, $p < .001$; omnibus test: $\chi_1^2 = 11.4$, $p < .001$). Other interactions (i.e., P(stop) x Group, P(stop) x SSD, P(stop) x SSD x Group) did not reach statistical significance ($p > .05$).

5.3.3 fMRI Analyses

Bayesian Prediction of Inhibitory Response (P(Stop)). Testing for any group differences in brain activation associated with P(stop), after regressing out any variance correlated with actual stimulus outcome (stop vs. go), we found no areas consistent with such neural pattern.

Modulation of Bayesian Prediction of Inhibitory Response (P(Stop)) by Trial Type (Stop vs. Go). Seven regions were associated with a significant interaction between groups (CS vs. MDI) and P(stop)-modulated trial type [Stop x P(stop) vs. Go x P(stop)]. These regions and their coordinates are listed in Table 5.2. In five of those regions (right hemisphere regions and left caudate), MDIs showed a positive activation associated with P(stop) on stop trials ($p < .05$), but no significant P(stop) activation on go trials ($p > .05$). CSs showed no significant P(stop) activations to go or stop trials in these regions ($p > .05$).

Table 5.2: BOLD Activation Foci for Group by P(Stop)-Modulated Trial Type (Go vs. Stop) Interaction. Whole brain random effect analysis; corrected for clusterwise significance: $p < .01$; minimum voxel significance is $p < .005$ and minimum cluster size is 12 voxels/768 μL . BOLD, blood oxygen-level dependent; BA, Brodmann area.

Region	Peak Voxel Talairach Coordinates (x,y,z)			Peak Voxel Z Statistics (p Value)	Cluster Size (Voxels)
Right Anterior Cingulate Cortex (BA 32)	20	36	12	4.71 (.00001)	215
Left Frontal Caudate	-16	22	11	3.81 (.00014)	53
Right Dorsolateral Prefrontal Cortex (BA 9)	44	17	39	4.21 (.00003)	34
Right Inferior Frontal Gyrus (BA 44)	53	16	18	3.31 (.00093)	17
Left Posterior Caudate	-15	-14	26	3.57 (.00036)	16
Left Orbitofrontal Gyrus (BA 11)	-38	47	-11	3.42 (.00063)	14
Right Middle Temporal Gyrus (BA 22)	51	-33	0	2.90 (.00373)	13

In the other two regions (left posterior caudate and orbito-frontal cortex [OFC]/Brodmann area 11 [BA 11]) (see Figures 5.3 A and 5.4 A), CSs showed a negative correlation with P(stop) on stop trials and positive correlation with P(stop) on go trials, consistent with a positive UPE ($\text{UPE} = |\text{outcome} - \text{P}(\text{stop})|$). However, unlike CSs, MDIs failed to show a differential P(stop) activation to go versus stop trials in those regions (see Figures 5.3 B and 5.4B). Importantly, based on supplemental analyses, CSs showed activation positively correlated with a UPE in these areas, whereas MDIs had significantly attenuated UPE activations, which was not statistically different from zero (Cohen's $d = 0.45$ and 0.42 for left posterior caudate and OFC/BA 11, respectively) (see Figure 5.3 C and 5.3 C).

Modulation of Bayesian Prediction of Inhibitory Response (P(Stop))

by Stop Accuracy (SS vs. SE). Activation in several neural regions were associated with a significant interaction between group and P(stop) modulated stop accuracy (SS vs. SE trials) (see Table 5.3 for coordinates). In all regions, CSs exhibited a significant negative correlation to P(stop) on SE trials and a positive correlation to P(stop) on SS trials ($p < .05$), whereas MDI showed no significant P(stop) activation on either SS or SE trials ($p > .05$).

Table 5.3: BOLD Activation Foci for Group by P(Stop)-Modulated Stop Trial Type (SS vs. SE) Interaction. Whole brain random effect analysis; corrected for clusterwise significance: $p < .01$; minimum voxel significance is $p < .005$ and minimum cluster size is 12 voxels/768 μ L BA, Brodmann area; BOLD, blood oxygen level?dependent; SE, stop error; SS, stop success.

Region	Peak Voxel Talairach Coordinates (x,y,z)			Peak Voxel Z Statistics (p Value)	Cluster Size (Voxels)
Left Precentral Gyrus	-26	4	33	3.71 (.00021)	19
Right Caudate (BA 25)	3	19	5	3.18 (.00147)	15
Right Thalamus	11	-18	19	2.99 (.00279)	14
Left Postcentral Gyrus (BA 2)	-54	-18	32	3.25 (.00115)	13
Left Inferior Semilunar Lobule	-18	-77	-36	3.74 (.00018)	12
Right Inferior Parietal Lobule (BA 40)	54	-51	41	3.39 (.00070)	12

To further assess whether these activation patterns may be linked to group difference in the coding of Bayesian predication errors, we extracted activation associated with both UPE and SPE in those five regions. Only two of these areas (right thalamus and right inferior parietal lobule [IPL]/BA 40) (see Figures 5.5 A and 5.6 A) were associated with activation patterns consistent with a selective encoding of a positive UPE on SE trials among CSs. In contrast, MDIs showed no significant activations

to UPE on SE or any other trials (Cohen’s $d = 0.41$ and 0.62 in right thalamus and IPL, respectively) (see Figures 5.5 B-C and 5.6 B-C). Correlational analyses further showed that such UPE activation on SE trials among CSs (not significantly different from zero in MDIs) was associated with lower error rates for difficult trials ($r = -.47$, $p < .05$; we note that a significant correlation was maintained with removal of outlier with high UPE value, i.e., $r = -0.36$, $p < .05$; scatter for MDIs not shown because UPE activation was not significantly different from 0; SSD, stop signal delay.) (see Figure 5.6 D).

5.4 Discussion

The goal of this investigation was to better delineate which processes are dysfunctional in a group of recently abstinent MDIs. We applied a Bayesian model to evaluate the probability of a stop signal and used this probability to quantify the degree of neural processing of “the need to stop,” and of the discrepancy between observations and actual outcomes—that is, sensitivity to surprising outcomes based on associated prediction errors. MDIs exhibited a pattern of reduced neural activation associated with trial-level Bayesian prediction error signals (i.e., more surprise) in orbitofrontal frontal, parietal, and subcortical areas (caudate and thalamus). Consistent with evidence of impaired inhibitory function in this population [SAX⁺10, GJK⁺11, JAX⁺05], MDIs compared with CSs were less accurate during more difficult trials (i.e., longer SSD). In addition, MDIs did not show evidence of strategic adjustment after a stop error; in other words, they did not show response time slowing after an error. Nevertheless, both CSs and MDIs maintain and use an internal representation of stop trial probability to make anticipatory adjustments for inhibitory control (e.g., in their reaction times)

based on trial history, which supports our modeling approach and is consistent with previous work [JPAC13, KPJ⁺14].

Relative to CSs, MDIs had attenuated neural activation associated with a Bayesian model-based UPE in the left OFC and posterior caudate. Whereas both the OFC and caudate have been implicated in stimulus-reward learning and value-based decision making (45,46), these areas are also critical for prediction and processing of performance feedback during decision making [MNB⁺02, RCR97, RGR99], including signaling expectation violation [AJCM99] and prediction errors [MJI⁺07, NRBR04, MJDE08, MM06, AMC11]. Importantly, failure to recruit the OFC has been linked to impulsivity and poor inhibitory function (54), as well as to impaired learning of stimulus-outcome contingencies in MDIs [MNB⁺02]. A group difference in unsigned Bayesian prediction error, as opposed to an action-dependent SPE or nonmodel-based error activity, may further suggest specific deficits among MDIs in tracking the overall inconsistency between environmental demands and their internal belief/prediction model (i.e., a “goodness-of-fit” estimate), rather than action-specific inconsistency, with negative consequences on inhibitory performance. Thus, rather than promoting a failure to recruit specific frontostriatal regions to predict appropriate actions, MD may impair neural tracking of model-based expectancy violation, which may be useful more generally in preparing individuals for switching strategy in response to significant changes in the environment [AJ08, DMM⁺14] (e.g., sudden change in reward rate), rather than in model updating (i.e., P(stop) estimation). The absence of group difference in overall P(stop) activations across trial type indeed suggests that groups did not differ in their mean estimates of P(stop), which lends credence to this interpretation.

Interestingly, MDIs further exhibited poorer neural encoding of stopping expectations [i.e., $P(\text{stop})$] on SE trials in a set of parietal and subcortical regions, consistent with their failure to slow down following SE trials. Specifically, CSs but not MDIs activated the right thalamus and IPL proportionally to a positive UPE on SE trials relative to other trials (go and SS). Such UPE activations were associated with lower error rates for difficult trials among CSs, suggesting that this selective UPE encoding (attenuated or absent in MDIs) may be particularly important to adjust and improve inhibitory performance. Such results are congruent with evidence of prediction error signaling in the thalamus [AIS⁺00, HSJ06] and the IPL, with functional connections to the dorsal caudate [AAD⁺08] and premotor areas [MTH06], which has also been implicated in predictive coding [BG07, VKM⁺11], and most recently in instrumental (action-reward) contingency learning [MEJB11] and tracking the degree of divergence/volatility in instrumental action probability distributions [MSJJ13].

This study has several limitations. First, we recognize that our computations of prediction errors are not inherently Bayesian, and they overlap with Bayesian and Information Theory estimates of precision, volatility [TMMM07], and surprise [RST⁺08]. Brain processes associated with UPE/SPE are, however, an important first step and provide good approximations of the optimal Bayesian model updating [AJ08, DMM⁺14]. Nevertheless, the present study did not explicitly measure volatility and model accuracy or precision, and thus cannot address whether the observed neural difference in UPE may reflect a difference in Bayesian estimates of volatility (i.e., degree of expectancy violation) or in precision (inversely related to the variance around mean expectation). Second, our model could be refined by including a more direct prediction of behavior such as RTs, as we have done behaviorally in previous work

[PJ11]. Such approach, however, would require significantly larger data collection owing to additional model parameters, which was not feasible for this fMRI study. Finally, we note that participants' interpretation of the stop-signal task instructions, which include incentives to limit post-stop error slowing, may lend some level of confound in the interpretation of our results, for example, reduced post-stop slowing could relate to more rigid instructions following in MDIs.

In summary, the present findings suggest that both cortical and subcortical structures in MDIs fail to adequately track the changes in environmental characteristics that would help to predict the need for increased inhibition. Importantly, using the same parameter values across groups allowed us to assess group differences in an objective measure of the “need to stop” and the “surprise” associated with each trial (i.e., participants' sensitivity to surprising or informative trials) in terms of their behavioral and neuronal responses. Our results further suggest this group difference may reflect MDIs' poorer tracking of expectancy violation (i.e., weaker sensitivity to surprising task events), rather than real-time mean predictions per se. Although we did observe a relationship between UPE activations in the IPL and error rates, suggesting a negative impact of such weaker activations in MDIs on their behavioral performance, future studies are needed to determine how such neural alterations may specifically relate to Bayesian estimates of volatility and precision, and how they may directly impact performance. This study highlights the utility of Bayesian learning models for investigating subtle cognitive alterations guiding goal-directed actions in addiction and other psychiatric populations, which can be used to develop precise addiction medicine approaches for better diagnosis and treatment of this disorder.

Acknowledgments

This chapter, in full, is a reprint of the material from Katia M. Harl, Shunan Zhang, Ning Ma, Angela J. Yu, and Martin P. Paulus, Reduced Neural Recruitment for Bayesian Adjustment of Inhibitory Control in Methamphetamine Dependence, *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 1:448-459, 2016. The dissertation author was the investigator and co-author of this paper.

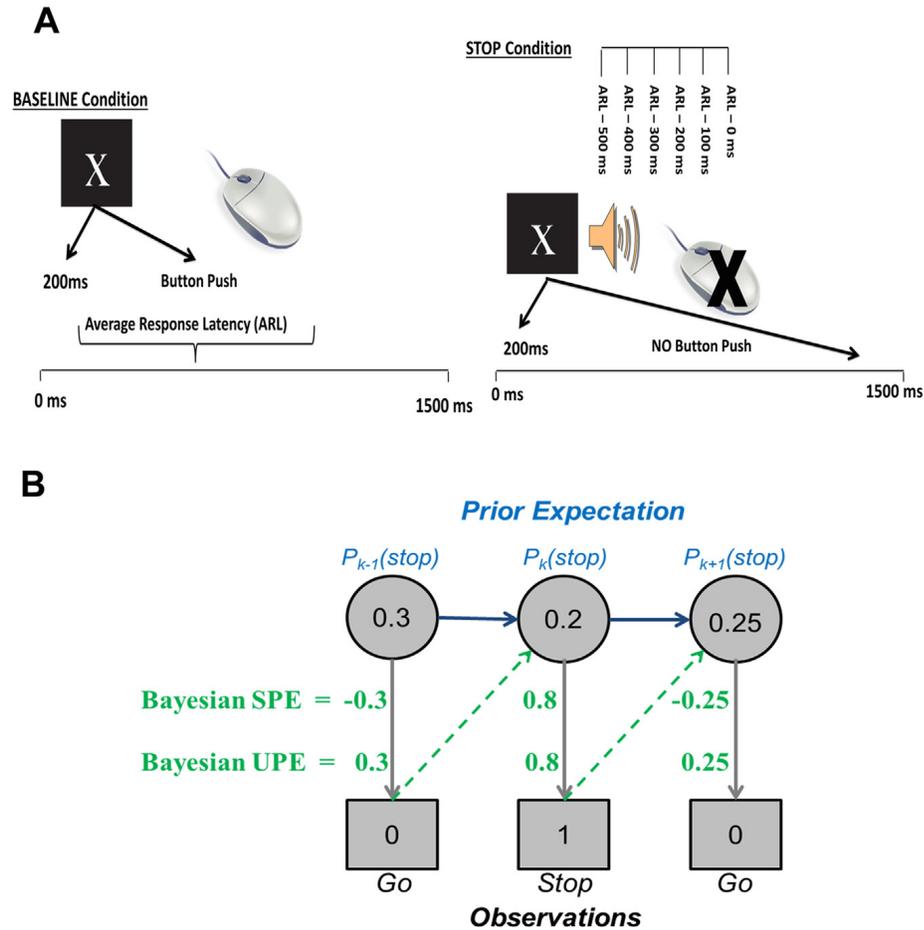


Figure 5.1: (A) Stop signal task. Participants completed a total of 288 trials, including 216 go trials (baseline condition), and 72 stop trials. Each trial lasted 1300 ms and trials were separated by 200-ms interstimulus intervals (blank screen). Participants were instructed to press as quickly as possible the left button when an ‘X’ appears, or the right button when an ‘O’ appears. They were also instructed not to press either button whenever they heard a tone during a trial (stop condition) (B) Bayesian hidden Markov model used, a version of the dynamic belief model [AJ08], which computes trialwise sequential predictions about the frequency of stop trials. The previous probability of encountering a stop signal on trial k , $P_k(\text{stop})$, is compared with the actual trial outcome (0 = go; 1 = stop) to produce a signed prediction error ([SPE]; i.e., $\text{outcome} - P_k(\text{stop})$), which is combined with the prior to produce a new updated prior for the next trial $k + 1$. UPE, unsigned prediction error (i.e., $|\text{outcome} - P(\text{stop})|$).

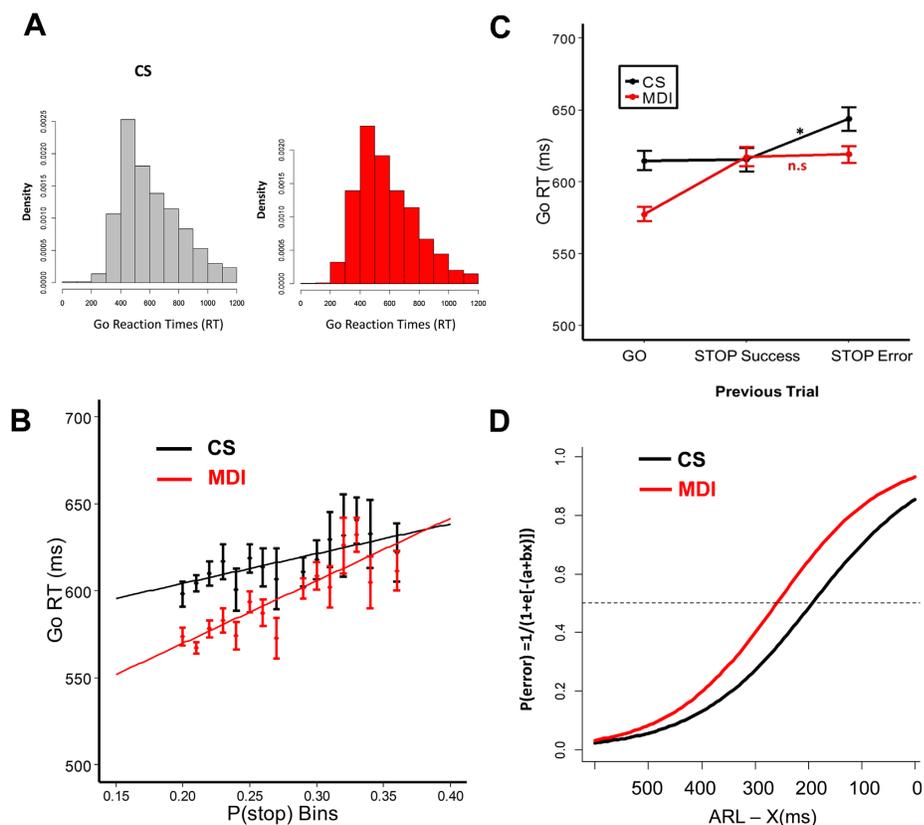


Figure 5.2: (A) Histograms of go reaction times (RT) for both comparison subjects ([CSs]; gray; $n = 36$) and methamphetamine-dependent individuals ([MDIs]; red; $n = 63$). (B) Bayesian model prediction and behavioral data presented for each group: red for MDI, black for CS. As predicted by our Bayes optimal decision-making model, a significant positive relationship was observed between individuals' go RT and trialwise $P(\text{stop})$ model estimates in each group. CS (black) and MDI (red) model lines represent best linear regression fit to mean go RT. Error bars are SEM for $P(\text{stop})$ bins. (C) Go RTs on trials following a go, successful stop (stop success), or failed stop trial (stop error). CSs had similar RTs following go and successful stop trials, but exhibited slower reaction times following stop error trials ($p < .05$). MDIs were generally slower on stop trials relative to go trials, but they did not exhibit a similar slowing after stop error trials relative to stop success trials ($p > .05$). Error bars are within-group SEM. (D) Fitted logistic inhibitory functions by group representing the likelihood of error on stop trials as a function of the stop signal delay, that is, individual mean go RT (i.e., average response latency [ARL]) - X with X ranging from 500 ms to 0 ms. n.s., not significant.).

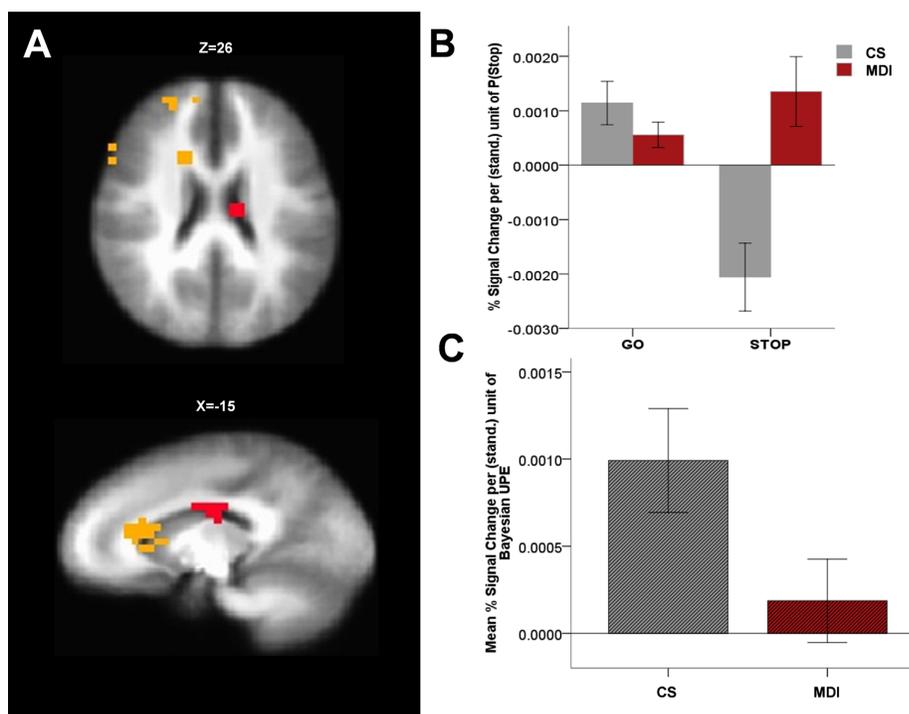


Figure 5.3: Group difference in sensitivity to surprising trials, that is, in activation to Bayesian unsigned prediction error (UPE), in the left posterior caudate. (A) Blood oxygen level-dependent signal in the posterior caudate showing group difference in percentage signal change modulation by UPE (red cluster). Yellow areas are other regions surviving whole brain analysis for a significant interaction between group and P(stop)-modulated trial type, but in which activation patterns are not consistent with a neural response correlated with a Bayesian UPE or signed prediction error. (B) Bar graph displays average P(stop) modulation of percentage signal change by trial type (go vs. stop) and group (comparison subjects [CSs]: $n = 34$; methamphetamine-dependent individuals [MDIs]: $n = 62$; error bars indicate ± 1 SEM). In this area, CSs (gray bars) demonstrated a neural response consistent with a positive UPE ($|\text{outcome} - P(\text{stop})|$), in other words, a positive correlation between percentage signal change and P(stop) on go trials and a negative correlation on stop trials, whereas MDIs (maroon bars) failed to show such differential P(stop)-dependent activation. (C) Average percentage signal change correlation with a positive Bayesian UPE ($|\text{outcome} - P(\text{stop})|$) for each group (error bars: ± 1 SEM). Relative to CSs (gray, black stripes), MDIs (maroon, black stripes) showed attenuated UPE-dependent activation (Cohen's $d = 0.45$). β was not statistically different from 0 in the MDI group ($p > .05$).

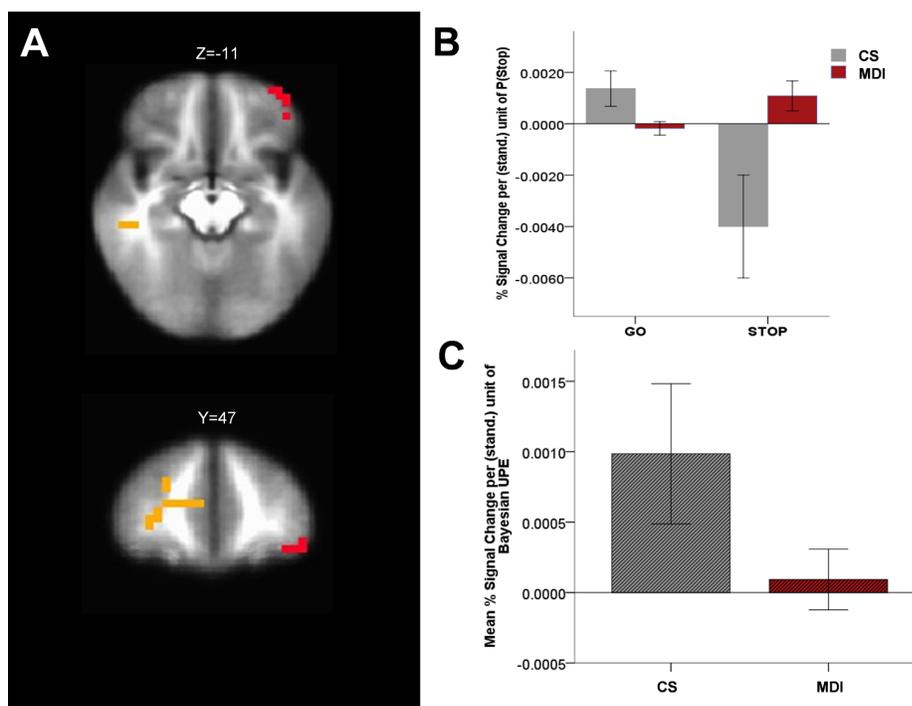


Figure 5.4: Group difference in sensitivity to surprising trials, that is, in neural activation to a Bayesian unsigned prediction error (UPE) in the left middle frontal gyrus (Brodmann area 11). (A) Blood oxygen level-dependent signal in the left middle frontal gyrus showing a group difference in percentage signal change modulation by UPE (red cluster). Yellow areas are other regions surviving whole brain analysis for a significant interaction between group and P(stop)-modulated trial type, but in which activation patterns are not consistent with a neural response correlated with a Bayesian UPE or signed prediction error. (B) Bar graph displays average P(stop) modulation of percentage signal change by trial type (go vs. stop) and group (comparison subjects [CSs]: $n = 34$; methamphetamine-dependent individuals [MDIs]: $n = 62$; error bars indicate ± 1 SEM). In this area, CSs (gray bars) demonstrated a neural response consistent with a positive UPE ($|\text{outcome} - P(\text{stop})|$), in other words, a positive correlation between percentage signal change and P(stop) on go trials and a negative correlation on stop trials, whereas MDIs (maroon bars) failed to show such differential P(stop)-dependent activation. (C) Average percentage signal change correlation with a positive Bayesian UPE ($|\text{outcome} - P(\text{stop})|$) for each group (error bars: ± 1 SEM). Relative to CSs (gray, black stripes), MDIs (maroon, black stripes) showed attenuated UPE-dependent activation (Cohen's $d = 0.42$). β was not statistically different from 0 in the MDI group ($p > .05$).

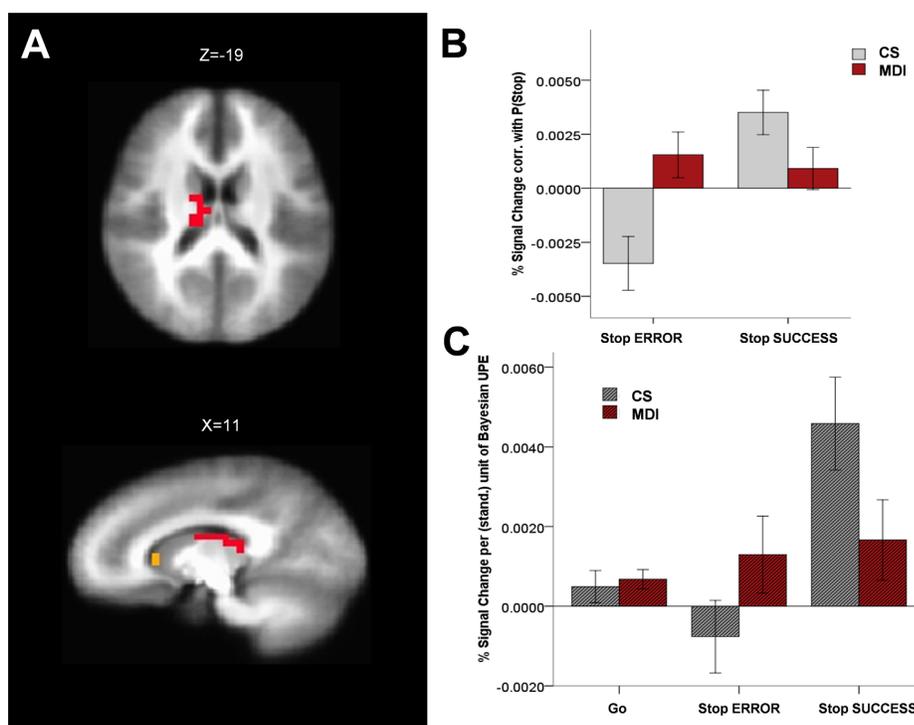


Figure 5.5: Group difference in the modulation of neural activation correlated with $P(\text{stop})$ by inhibitory success (right thalamus). (A) Blood oxygen level-dependent signal regions representing a significant interaction between group and $P(\text{stop})$ -modulated activation for stop success (SS) vs. stop error (SE) trials (yellow and red clusters). The right thalamus (red cluster), one of these regions, further showed an activation pattern consistent with a group difference in percentage signal change correlated with unsigned prediction error (UPE) on stop error trials. (B) Bar graphs represent average percentage signal change for parametric regressors SE \times $P(\text{stop})$ and SS \times $P(\text{stop})$ in comparison subjects ([CSs]; $n = 34$) and methamphetamine-dependent individuals ([MDIs]; $n = 62$). Percentage signal change in CSs (gray bars) was negatively correlated with $P(\text{stop})$ on SE trials and positively correlated with $P(\text{stop})$ on SS trials. MDIs (maroon bars) had significantly lower $P(\text{stop})$ -modulated activation on both SS and SE trials, which was not significantly different from 0 ($p > .05$). (C) Region of interest analysis (right thalamus). In this region, percentage signal change was selectively positively correlated with a Bayesian UPE, in other words, the amount of surprise or expectancy violation, on SE trials among CSs (gray striped bars), and not significantly correlated with a UPE on successful go and stop trials. In contrast, no statistically significant UPE-dependent activation was observed in MDIs for any type of trial (successful or errors; $p > .05$; group difference on SE: Cohen's $d = 0.41$); error bars indicate ± 1 SEM.

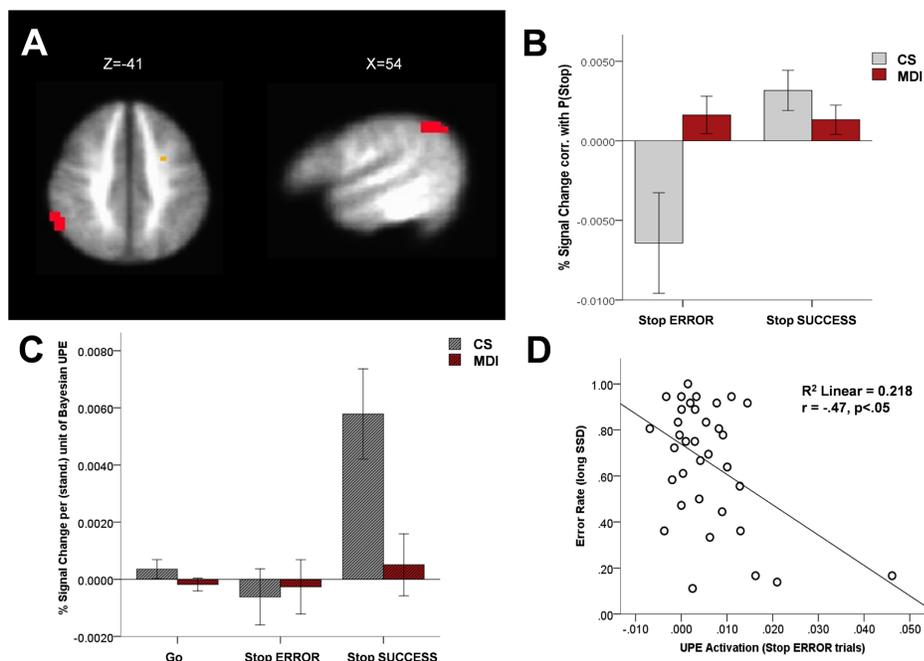


Figure 5.6: Group difference in the modulation of neural activation correlated with $P(\text{stop})$ by inhibitory success (right inferior parietal lobule [IPL]; Brodmann area 40). (A) Blood oxygen level-dependent signal regions representing a significant interaction between group and $P(\text{stop})$ -modulated activation for stop success (SS) vs. stop error (SE). The IPL (red cluster), one of these regions, further showed an activation pattern consistent with a group difference in percentage signal change correlated with unsigned prediction error (UPE) on SE trials. (B) Bar graphs represent average percentage signal change for parametric regressors SE \times $P(\text{stop})$ and SS \times $P(\text{stop})$ in comparison subjects ([CSs]; $n = 34$) and methamphetamine-dependent individuals ([MDIs]; $n = 62$). Percentage signal change in CSs (gray bars) was negatively correlated with $P(\text{stop})$ on SE trials and positively correlated with $P(\text{stop})$ on SS trials. MDIs (maroon bars) had significantly lower $P(\text{stop})$ -modulated activation on both SS and SE trials, which was not significantly different from 0 ($p > .05$). (C) Region of interest analysis. In this region (right IPL), percentage signal change was selectively positively correlated with a Bayesian UPE, that is, the amount of surprise or expectancy violation, on SE trials among CSs (gray striped bars), and not significantly correlated with a UPE on successful go and stop trials. In contrast, no statistically significant UPE-dependent activation was observed in MDI for any type of trial (successful or errors; $p > .05$); group difference on SE trials: Cohen's $d = 0.62$; error bars indicate ± 1 SEM. (D) This UPE activation on SE trials in the right IPL among CSs was further associated with lower error rates on difficult trials (stop signal delays . mean reaction time 200 ms). Graph shows the scatter plot for this significant negative correlation ($r = -0.47, p < .05$).

Bibliography

- [AAD⁺08] Di Martino A, Scheres A, Margulies D, Kelly AM, Uddin LQ, and Shehzad Z. Functional connectivity of human striatum: A resting state fmri study. *Cereb Cortex*, 18:2735–2747, 2008.
- [ADE⁺07] A R Aron, S Durston, D M Eagle, G D Logan, C M Stinear, and V Stuphorn. Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *Journal of Neuroscience*, 27(44):11860–4, 2007.
- [AIS⁺00] Ploghaus A, Tracey I, Clare S, Gati JS, Rawlins JN, and Matthews PM. Learning about pain: The neural substrate of the prediction error for aversive events. *Proc Natl Acad Sci*, 97:9281–9286, 2000.
- [AJ08] Yu AJ and Cohen JD. Sequential effects: Superstition or rational behavior. *Adv Neural Inf Process Syst*, 21:1873–1880, 2008.
- [AJCM99] Nobre AC, Coull JT, Frith CD, and Mesulam MM. Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nat Neurosci*, 2:11–12, 1999.
- [AMC11] Mattfeld AT, Gluck MA, and Stark CE. Functional specialization within the striatum along both the dorsal/ventral and anterior/posterior axes during associative learning via reward and punishment. *Learn Mem*, 18:703–711, 2011.
- [APA13] APA. *Diagnostic and Statistical Manual of Mental Disorders., 5th ed.* American Psychiatric Publishing, Washington, DC, 2013.
- [ARK07] R M Alderson, M D Rapport, and M J Kofler. 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, 2007.
- [Bar97] R A Barkley. Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological bulletin*, 121(1):65–94, 1997.

- [BBM⁺06] R Bogacz, E Brown, J Moehlis, P Hu, P Holmes, and J D Cohen. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review*, 113(4):700–65, 2006.
- [Bel52] R Bellman. On the theory of dynamic programming. *PNAS*, 38(8):716–719, 1952.
- [BG07] Knutson B and Wimmer GE. Splitting the difference: How does the brain code reward episodes? *Ann N Y Acad Sci*, 1104:54–69, 2007.
- [BMJC02] J C Badcock, P T Michie, L Johnson, and J Combrinck. Acts of control in schizophrenia: Dissociating the components of inhibition. *Psychological Medicine*, 32(2):287–297, 2002.
- [BPLS07] L Boucher, T J Palmeri, G D Logan, and J D Schall. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychological Review*, 114(2):376–97, 2007.
- [BSNM92] K H Britten, M N Shadlen, W T Newsome, and J A Movshon. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.*, 12:4745–65, 1992.
- [CAGK98] Bchel C, Holmes AP, Rees G, and Friston KJ. Characterizing stimulus-response functions using nonlinear regressors in parametric fmri experiments. *Neuroimage*, 8:140–148, 1998.
- [CNB⁺02] R Y Cho, L E Nystrom, E T Brown, Jones A D, Braver T S, P J Holmes, and J D Cohen. Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced choice task. *Cognitive, Affective, & Behavioral Neuroscience*, 2:283–99, 2002.
- [DMM⁺14] Tervo DG, Proskurin M, Manakov M, Kabra M, Vollmer A, Branson K, and Karpova AY. Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell*, 159:21–32, 2014.
- [DOP⁺10] J Daunizeau, H E.M.d Ouden, M Pessiglione, S J Kiebel, K E Stephan, and K J Friston. Observing the observer (i): Meta-bayesian models of learning and decision-making. *PLoS One*, 2010.
- [DY12] S Dayanik and A J Yu. Reward-rate maximization in sequential identification under a stochastic deadline. *SIAM Journal on Control and Optimization*, 51(4):2922–2948, 2012.
- [EBB⁺07] E E Emeric, J W Brown, L Boucher, R H S Carpenter, D P Hanes, R Harris, G D Logan, R N Mashru, M Paré, P Pouget, and V Stuphorn. Influence of history on saccade countermanding performance in humans and macaque monkeys. *Vision Research*, 47(1):35–49, 2007.

- [ESS⁺04] London ED, Simon SL, Berman SM, Mandelkern MA, Lichtman AM, and Bramen J. Mood disturbances and regional cerebral metabolic abnormalities in recently abstinent methamphetamine abusers. *Arch Gen Psychiatry*, 61:73–84, 2004.
- [FY08] P Frazier and A J Yu. Sequential hypothesis testing under stochastic deadlines. *Advances in Neural Information Processing Systems*, 20, 2008.
- [GJK⁺11] Tabibnia G, Monterosso JR, Baicy K, Aron AR, Poldrack RA, and Chakrapani S. Different forms of self-control share a neurocognitive substrate. *J Neurosci*, 31:4805–4810, 2011.
- [GS02] J I Gold and M N Shadlen. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36:299–308, 2002.
- [HSJ06] Kim H, Shimojo S, and O’Doherty JP. Is avoiding an aversive outcome rewarding? neural substrates of avoidance learning in the human brain. *PLoS Biol*, 4:e233, 2006.
- [HSP98] D P Hanes, J Schall, and W F Patterson. The role of frontal eye field in countermanding saccades: visual, movement and fixation activity. *J. Neurophysiol.*, 79:817–34, 1998.
- [HSS⁺14] K M Harlé, P Shenoy, J L Steward, S Tapert, A J Yu*, and M P Paulus*. Altered neural processing of the need to stop in young adults at risk for stimulant dependence. *Journal of Neuroscience*, 34:4567–4580, 2014. *Yu and Paulus are co-senior authors.
- [HSZ⁺15] K M Harlé, J L Steward, S Zhang, S Tapert, A J Yu*, and M P Paulus*. Bayesian neural adjustment of inhibitory control predicts emergence of problem stimulant use. *Brain*, 2015. *Yu and Paulus are co-senior authors.
- [IHZ⁺15] JS Ide, S Hu, S Zhang, A J Yu, and C-S R Li. Impaired bayesian learning for cognitive control in cocaine dependence. *Drug and Alcohol Dependence*, 151:220–227, 2015.
- [ISYL13] J S Ide, P Shenoy, A J Yu*, and C-R Li*. Bayesian prediction and evaluation in the anterior cingulate cortex. *Journal of Neuroscience*, 33:2039–2047, 2013. *Yu and Li contributed equally as senior authors.
- [JAX⁺05] Monterosso JR, Aron AR, Cordova X, Xu J, and London ED. Deficits in response inhibition associated with chronic methamphetamine abuse. *Drug Alcohol Depend*, 79:273–277, 2005.
- [JDSD11] Pinheiro J J, Bates D, DebRoy S, and Sarkar D. The r development core team 2011 nlme: Linear and nonlinear mixed effects models. 2011.

- [JJ99] Jentsch JD and Taylor JR. Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology (Berl)*, 146:373–390, 1999.
- [JKJ⁺14] Gowin JL, Harl KM, Stewart JL, Wittmann M, Tapert SF, and Paulus MP. Attenuated insular processing during risk predicts relapse in early abstinent methamphetamine-dependent individuals. *Neuropsychopharmacology*, 39:1379–1387, 2014.
- [JM07] Aron JL and Paulus MP. Location, location: Using functional magnetic resonance imaging to pinpoint brain differences relevant to stimulant use. *Addiction*, 102:33–43, 2007.
- [JPAC13] Ide JS, Shenoy P, Yu AJ, and Li CS. Bayesian prediction and evaluation in the anterior cingulate cortex. *J Neurosci*, 33:2039–2047, 2013.
- [Kal60] R E Kalman. A new approach to linear filtering and prediction problems. *J Basic Engineering*, 82(1):35–45, 1960.
- [KE07] Baicy K and London ED. Corticolimbic dysregulation and chronic methamphetamine abuse. *Addiction*, 102:5–15, 2007.
- [KN04] Kirby KN and Petry NM. Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction*, 99:461–471, 2004.
- [KPJ⁺14] Harl KM, Shenoy P, Stewart JL, Tapert SF, Angela JY, and Paulus MP. Altered neural processing of the need to stop in young adults at risk for stimulant dependence. *J Neurosci*, 34:4567–4580, 2014.
- [KRC⁺94] Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, and Nurnberger JI Jr. A new, semi-structured psychiatric interview for use in genetic linkage studies: A report on the reliability of the ssaga. *J Stud Alcohol*, 55:149–158, 1994.
- [Lam68] D R J Laming. *Information Theory of Choice-Reaction Times*. Academic Press, London, 1968.
- [LBM⁺10] Degenhardt L, Mathers B, Guarinieri M, Panda S, Phillips B, and Strathdee SA. Meth/amphetamine use and associated hiv: Implications for global policy and public health. *Int J Drug Policy*, 21:347–358, 2010.
- [LC84] G Logan and W Cowan. On the ability to inhibit thought and action: A theory of an act of control. *Psych. Rev.*, 91:295–327, 1984.

- [LDJE11] Nestor LJ, Ghahremani DG, Monterosso J, and London ED. Prefrontal hypoactivation during cognitive control in early abstinent methamphetamine-dependent subjects. *Psychiatry Res*, 194:287–295, 2011.
- [LDTN07] Chang L, Alicata D, Ernst T, and Volkow N. Structural and metabolic brain changes in the striatum associated with methamphetamine abuse. *Addiction*, 102:16–32, 2007.
- [LHY+08] C S Li, C Huang, P Yan, P Paliwal, R T Constable, and R Sinha. Neural correlates of post-error slowing during a stop signal task: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, 20(6):1021–9, 2008.
- [LTW06] Somerville LH, Heatherton TF, and Kelley WM. Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nat Neurosci*, 9:1007–1008, 2006.
- [Luc86] R D Luce. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, New York, 1986.
- [LW09] L A Leotti and T D Wager. Motivational influences on response inhibition measures. *J. Exp. Psychol. Hum. Percept. Perform.*, 36(2):430–447, 2009.
- [MAC+07] L Menzies, S Achard, S R Chamberlain, N Fineberg, C H Chen, N del Campo, B J Sahakian, T W Robbins, and E Bullmore. Neurocognitive endophenotypes of obsessive-compulsive disorder. *Brain*, 130(12):3223–36, 2007.
- [Mar82] D Marr. *Vision*. Freeman, San Francisco, 1982.
- [MD14] Brecht ML and Herbeck D. Time to relapse following treatment for methamphetamine use: A long-term perspective on patterns and predictors. *Drug Alcohol Depend*, 139:18–25, 2014.
- [MDJM07] Wittmann M, Leland DS, Churan J, and Paulus MP. Impaired time perception and motor timing in stimulant-dependent subjects. *Drug Alcohol Depend*, 90:183–192, 2007.
- [MEJB11] Liljeholm M, Tricomi E, O’Doherty JP, and Balleine BW. Neural correlates of instrumental contingency learning: Differential effects of action-reward conjunction and disjunction. *J Neurosci*, 31:2474–2480, 2011.
- [MJDE08] Delgado MR, Li J, Schiller D, and Phelps EA. The role of the striatum in aversive learning and aversive prediction errors. *Philos Trans R Soc Lond B: Biol Sc*, 363:3787–3800, 2008.

- [MJI⁺07] Menon M, Jensen J, Vitcu I, Graff-Guerrero A, Crawley A, Smith MA, and Kapur S. Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: Effects of dopaminergic modulation. *Biol Psychiatry*, 62:765–772, 2007.
- [MM06] Haruno M and Kawato M. Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J Neurophysiol*, 95:948–959, 2006.
- [MNB⁺02] Paulus MP, Hozack NE, Zauscher BE, Frank L, Brown GG, Braff DL, and Schuckit MA. Behavioral and functional neuroimaging evidence for prefrontal dysfunction in methamphetamine-dependent subjects. *Neuropsychopharmacology*, 26:53–63, 2002.
- [MSJJ13] Liljeholm M, Wang S, Zhang J, and O’Doherty JP. Neural correlates of the divergence of instrumental probability distributions. *J Neurosci*, 33:12519–12527, 2013.
- [MSM05] Paulus MP, Tapert SF, and Schuckit MA. Neural activation patterns of methamphetamine-dependent subjects during decision making predict relapse. *Arch Gen Psychiatry*, 62:761–768, 2005.
- [MT03] Fillmore MT. Drug abuse as a problem of impaired control: current approaches and findings. *Behav and Cogn Neurosci Rev*, 2:179–197, 2003.
- [MTH06] Rushworth MF, Behrens TE, and Johansen-Berg H. Connection patterns distinguish 3 regions of human parietal cortex. *Cereb Cortex*, 16:1418–1430, 2006.
- [MY15a] N Ma and A J Yu. Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control. *Frontiers in Psychology*, 2015.
- [MY15b] N Ma and A J Yu. Variability in human response time reflects statistical learning and adaptive decision-making. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 2015.
- [MY16] N Ma and A J Yu. Inseparability of go and stop in inhibitory control: Go stimulus discriminability affects stopping behavior. *Frontiers in Neuroscience*, 10(54), 2016.
- [Nig00] J T Nigg. On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, 126(2):220–46, 2000.

- [NLG⁺01] Volkow ND, Chang L, Wang GJ, Fowler JS, Franceschi D, and Sedler MJ. Higher cortical and lower subcortical metabolism in detoxified methamphetamine abusers. *Am J Psychiatry*, 158:383–389, 2001.
- [NRBR04] Ramnani N, Elliott R, Athwal BS, and Passingham RE. Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage*, 23:777–786, 2004.
- [NWM⁺06] J T Nigg, M M Wong, M M Martel, J M Jester, L I Puttler, J M Glass, K M Adams, H E Fitzgerald, and R A Zucker. 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, 2006.
- [PH03] M Pare and D P Hanes. Controlled movement processing: Superior colliculus activity associated with countermanded saccades. *Journal of Neuroscience*, 23(16):6480–9, 2003.
- [PJ11] Shenoy P and Angela JY. Rational decision-making in inhibitory control. *Front Hum Neurosci*, 5:48, 2011.
- [PKS⁺04] Thompson PM, Hayashi KM, Simon SL, Geaga JA, Hong MS, and Sui Y. Structural abnormalities in the brains of human subjects who use methamphetamine. *J Neurosci*, 24:6028–6036, 2004.
- [PRA11] Shenoy P, Rao RPN, and Yu A. A rational decision making framework for inhibitory control. *Adv Neural Inf Process Syst*, 2011.
- [RBA⁺99] Rogers RD, Everitt B, Baldacchino A, Blackshaw A, Swainson R, and Wynne K. Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: Evidence for monoaminergic mechanisms. *Neuropsychopharmacology*, 20:322–339, 1999.
- [RCR97] Elliott R, Frith CD CD, and Dolan RJ. Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia*, 35:1395–1404, 1997.
- [RDD08] Baayen RH, Davidson DJ, and Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang*, 59:390–412, 2008.
- [RGR99] Elliott R, Rees G, and Dolan RJ. Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia*, 37:403–411, 1999.
- [RR98] R Ratcliff and J N Rouder. Modeling response times for two-choice decisions. *Psychological Science*, 9:347–56, 1998.

- [RST⁺08] Mars RB, Debener S, Gladwin TE, Harrison LM, Haggard P, Rothwell JC, and Bestmann S. Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J Neurosci*, 28:12539–12545, 2008.
- [RTY⁺07] Salo R, Nordahl TE, Natsuaki Y, Leamon MH, Galloway GP, and Waters C. Attentional control and brain metabolite levels in methamphetamine abusers. *Biol Psychiatry*, 61:1272–1280, 2007.
- [RW96] Cox RW. Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*, 29:162–173, 1996.
- [SAEM05] Matthews SC, Simmons AN, Arce E, and Paulus MP. Dissociation of inhibition from error processing using a parametric inhibitory task during functional magnetic resonance imaging. *Neuroreport*, 16:755–760, 2005.
- [SAX⁺10] Simon SL, Dean AC, Cordova X, Monterosso JR, and London ED. Methamphetamine dependence and neuropsychological functioning: Evaluating change during early abstinence. *J Stud Alcohol Drugs*, 71:335–344, 2010.
- [SBH85] E Soetens, L C Boer, and J E Hueting. Expectancy or automatic facilitation? separating sequential effects in two-choice reaction time. *J. Exp. Psychol.: Human Perception & Performance*, 11:598–616, 1985.
- [Smi95] P L Smith. Psychophysically principled models of visual simple reaction time. *Psychol. Rev.*, 10:567–93, 1995.
- [SRY10] P Shenoy, R Rao, and A J Yu. A rational decision making framework for inhibitory control. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, volume 23, pages 2146–54. MIT Press, Cambridge, MA, 2010.
- [STJ11] Kennerley SW, Behrens TE, and Wallis JD. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neurosci*, 14:1581–1589, 2011.
- [SY11] P Shenoy and A J Yu. Rational decision-making in inhibitory control. *Frontiers in Human Neuroscience*, 2011. doi: 10.3389/fnhum.2011.00048.
- [SY12] P Shenoy and A J Yu. Strategic impatience in Go/NoGo versus forced-choice decision-making. *Advances in Neural Information Processing Systems*, 25, 2012.
- [TF08] Jaeger TF. Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *J Mem Lang*, 59:434–446, 2008.

- [TMMM07] Behrens TE, Woolrich MW MW, Walton ME, and Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci*, 10:1214–1221, 2007.
- [TRY⁺05] Nordahl TE, Salo R, Natsuaki Y, Galloway GP, Waters C, and Moore CD. Methamphetamine users in sustained abstinence: A proton magnetic resonance spectroscopy study. *Arch Gen Psychiatry*, 62:444–452, 2005.
- [VGR⁺12] Clark VP, Beatty GK, Anderson RE, Kodituwakku P, Phillips JP, and Lane TDR. Reduced fmri activity predicts relapse in patients recovering from stimulant dependence. *Human Brain Mapp*, 35:414–428, 2012.
- [VKM⁺11] Spoormaker VI, Andrade KC, Schrter MS, Sturm A, Goya-Maldonado R, Smann PG, and Czisch M. The neural correlates of negative prediction error signaling in human fear conditioning. *Neuroimage*, 54:2250–2256, 2011.
- [WCM⁺14] C N White, E Congdon, A J Mumford, K H Karlsgodt, F W Sabb, N B Freimer, E D London, T D Cannon, R M Bilder, and R A Poldrack. Decomposing decision components in the stop-signal task: A model-based approach to individual differences in inhibitory control. *Journal of Cognitive Neuroscience*, 26:1601–1614, 2014.
- [WMR⁺06] Hoffman WF, Moore M, Templin R, McFarland B, Hitzemann RJ, and Mitchell SH. Neuropsychological function and delay discounting in methamphetamine-dependent individuals. *Psychopharmacology (Berl)*, 188:162–170, 2006.
- [WRT⁺13] Panenka WJ, Procyshyn RM, Lecomte T, MacEwan GW, Flynn SW, Honer WG, and Barr AM. Methamphetamine use: A comprehensive review of molecular, preclinical and clinical findings. *Drug Alcohol Depend*, 129:167–179, 2013.
- [YC09] A J Yu and J D Cohen. Sequential effects: Superstition or rational behavior? *Advances in Neural Information Processing Systems*, 21:1873–80, 2009.
- [YH14] A J Yu and H Huang. Maximizing masquerading as matching: Statistical learning and decision-making in choice behavior. *Decision*, 1(4):275–287, 2014.
- [YSD⁺09] Kim YT, Lee SW, Kwon DH, Seo JH, Ahn BC, and Lee J. Dose-dependent frontal hypometabolism on fdg-pet in methamphetamine abusers. *J Psychiatr Res*, 43:1166–1170, 2009.
- [YYM⁺02] Sekine Y, Minabe Y, Kawai M, Suzuki K, Iyo M, and Isoda H. Metabolite alterations in basal ganglia associated with methamphetamine-related

psychiatric symptoms: A proton mrs study. *Neuropsychopharmacology*, 27:453–461, 2002.

- [YYY+03] Sekine Y, Minabe Y, Ouchi Y, Takei N, Iyo M, and Nakamura K. Association of dopamine transporter loss in the orbitofrontal and dorsolateral prefrontal cortices with methamphetamine-related psychiatric symptoms. *Am J Psychiatry*, 160:1699–1701, 2003.