OXFORD

Cerebral Cortex, 2016; 1–10

doi: 10.1093/cercor/bhw219
Original Article

ORIGINAL ARTICLE

Dissociable Fronto-Operculum-Insula Control Signals for Anticipation and Detection of Inhibitory Sensory Cues

Weidong Cai[1], Tianwen Chen[1], Jaime S. Ide[2], Chiang-Shan R. Li[3,4], and Vinod Menon[1,5,6]

[1]Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, Stanford, CA 94305, USA, [2]Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY 11794, USA, [3]Department of Psychiatry, [4]Department of Neuroscience, Yale University School of Medicine, New Haven, CT 06519, USA, [5]Program in Neuroscience, and [6]Department of Neurology & Neurological Sciences, Stanford University School of Medicine, Stanford, CA 94305, USA

Address correspondence to Weidong Cai and Vinod Menon, Department of Psychiatry & Behavioral Sciences, 401 Quarry Rd., Stanford University, Stanford, CA 94305, USA. Email: wdcai@stanford.edu and menon@stanford.edu

Abstract

The ability to anticipate and detect behaviorally salient stimuli is important for virtually all adaptive behaviors, including inhibitory control that requires the withholding of prepotent responses when instructed by external cues. Although right fronto-operculum-insula (FOI), encompassing the anterior insular cortex (rAI) and inferior frontal cortex (rIFC), involvement in inhibitory control is well established, little is known about signaling mechanisms underlying their differential roles in detection and anticipation of salient inhibitory cues. Here we use 2 independent functional magnetic resonance imaging data sets to investigate dynamic causal interactions of the rAI and rIFC, with sensory cortex during detection and anticipation of inhibitory cues. Across 2 different experiments involving auditory and visual inhibitory cues, we demonstrate that primary sensory cortex has a stronger causal influence on rAI than on rIFC, suggesting a greater role for the rAI in detection of salient inhibitory cues. Crucially, a Bayesian prediction model of subjective trial-by-trial changes in inhibitory cue anticipation revealed that the strength of causal influences from rIFC to rAI increased significantly on trials in which participants had higher anticipation of inhibitory cues. Together, these results demonstrate the dissociable bottom-up and top-down roles of distinct FOI regions in detection and anticipation of behaviorally salient cues across multiple sensory modalities.

Key words: dynamic causal modeling, fMRI, human, prefrontal cortex, response inhibition

Introduction

Anticipation and detection of behaviorally salient events are key components of flexible cognitive control (Corbetta and Shulman 2002; Petersen and Posner 2012). The prefrontal cortex is known to play an essential role in this process and disruptions in prefrontally mediated control processes are thought to underlie many psychiatric disorders, including attention deficit hyperactivity disorder (ADHD), autism, and schizophrenia (Casey et al. 1997; Menon 2011; Palaniyappan et al. 2013; Uddin et al. 2013). Although the right anterior insula (rAI) and inferior frontal cortex (rIFC) subdivisions of fronto-operculum-insula (FOI) are among the most critical prefrontal regions for

cognitive control (Levy and Wagner 2011; Swick et al. 2011; Cai et al. 2014), little is known about dynamic causal signaling underlying detection of salient task cues from sensory inputs and how such signaling is modulated by top-down influences arising from subjective moment-by-moment changes in anticipation of sensory cues.

Much of our understanding of the mechanisms underlying flexible cognitive control has come from human functional neuroimaging studies using inhibitory control tasks which require participants to respond to prepotent cues and withhold responses to infrequent inhibitory cues (Logan et al. 1984; Verbruggen and Logan 2008). Meta-analyses of functional magnetic resonance imaging (fMRI) studies of human inhibitory control have demonstrated right hemispheric dominance and peak activations localized to both rAI and rIFC during inhibitory control (Figure S1, Levy and Wagner 2011; Swick et al. 2011; Cai et al. 2014). Lesions and disruptions to right FOI impair the ability to inhibit prepotent behavioral responses (Aron et al. 2003; Chambers et al. 2006). The right FOI also displays increased gamma-band activity immediately after the onset of inhibitory cues, suggesting its involvement in the early stages of inhibitory control (Swann et al. 2012). Thus, converging lines of evidence suggest that right FOI serves as a crucial interface between inhibitory cue detection and inhibitory control. However, despite years of research, the precise role of individual subdivisions of FOI in detection of inhibitory cues remains unknown. Crucially, no studies to date have attempted to disentangle the differential roles of rAI and rIFC in detection of inhibitory cues from sensory inputs.

Based on their distinct patterns of activation, intrinsic and task-related connectivity, and relation to behavior, Cai et al. (2014) suggested that the rAI may play a more important role in inhibitory cue detection, whereas the rIFC may be more involved in the subsequent stages of inhibitory control. However, no studies to date have directly examined whether rAI and rIFC differ in their interactions with the sensory regions from which they receive inputs. This question is challenging for lesion, transcranial stimulation, and electrocorticogram investigations of inhibitory control for several reasons. First, due to the close anatomical proximity of the rAI to rIFC, it is difficult to find patients with lesions restricted exclusively to only one region. Second, it is hard to temporally disrupt only one subdivision using transcranial stimulation without impacting other neighboring regions. Third, although electrocorticograms have the requisite anatomical and temporal resolution, limited electrode coverage and small sample sizes have precluded investigations of long-range cortical interactions. Finally, reproducibility remains a challenge for both invasive and non-invasive brain imaging techniques. In this study, we overcome these limitations using a novel neurocognitive systems approach. Accordingly, the first major goal of this study was to investigate dynamic causal interactions between sensory cortex and FOI control system associated with detection of salient inhibitory cues and to replicate findings in 2 independent data sets.

Anticipatory and adaptive adjustments in response to a dynamically changing environment are another fundamental, but much less studied, aspect of inhibitory control (Aron 2011). While most previous studies of inhibitory control have focused on the behavioral and neural bases of reactive processes associated with inhibitory cues, a few recent studies have provided novel evidence for trial-by-trial changes in inhibitory cue anticipation based on Bayesian prediction models (BPM) of accumulating sensory evidence (Shenoy and Yu 2011; Ide et al. 2013; Hu et al. 2015). Interestingly, higher inhibitory cue anticipation is associated with less stopping errors (Ide et al. 2013). We posit that when the likelihood of an inhibitory cue increases, the brain network important for stop-signal detection, such as the salience network, is signaled by the proactive control system via top-down modulation in order to facilitate cue detection. Previous studies have used contextual cues to modulate probability of stop trials and found that rIFC is part of a proactive control system (Chikazoe et al. 2009; Swann et al. 2012). Furthermore, direct electrical stimulation (Wessel et al. 2013) or transcranial direct current stimulation (Cai et al. 2016) to right FOI cortex has been shown to facilitate proactive control. However, the differential roles of rAI and rIFC and their interactions with each other and with sensory cortex during inhibitory cue anticipation are currently unknown. Therefore, the second major goal of our study was to examine dynamic causal interactions of sensory cortex and FOI control system associated with anticipation of inhibitory cues.

We hypothesized that detection of inhibitory cues would be associated with stronger interactions between rAI, than rIFC, and domain-specific sensory regions, whereas anticipation of inhibitory cues would involve top-down modulation from rIFC to the salience detection system, anchored in rAI. Here we directly test this hypothesis in a parsimonious neurocognitive model by analyzing dynamic causal interactions associated with inhibitory cue detection and anticipation. We first used dynamic causal modeling (DCM) (Friston et al. 2003) to investigate causal interactions between rAI, rIFC, and domain-specific sensory regions during detection of stop signals. To ensure reliability and generalizability, we evaluated our hypotheses in 2 independent data sets: stop-signal task 1 (SST1, Xue et al. 2008) obtained from a public fMRI database—OpenfMRI (http://openfmri.org, Texas Advanced Computing Center, the University of Texas at Austin) and stop-signal task 2 (SST2, Zhang and Li 2012), acquired at Yale University. In both SST1 and SST2, participants were required to make motor responses to frequent go signals and to inhibit their responses to infrequent stop signals (inhibitory cue) (Fig. 1*a*). Because stop signals were delivered via 2 different sensory channels, auditory in SST1 and visual in SST2, use of these 2 data sets allowed us to examine the generalizability of our findings with respect to primary auditory and visual cortices.

Next, we combined BPM (Shenoy and Yu 2011; Ide et al. 2013) with DCM to investigate dynamic causal interactions between rAI, rIFC, and domain-specific sensory regions during anticipation of inhibitory cues. BPM was used to compute p_{stop}, the subjective expectation of stop signals in each trial (Shenoy and Yu 2011; Ide et al. 2013). This analysis was only conducted on SST2 because of the small sample size and small number of trials with low p_{stop} in SST1. We predicted that proactive top-down signaling from rIFC (Aron 2011; Wessel et al. 2013) would have strong causal influences on rAI. We contrasted our findings with a model based on a purely attentional view of rIFC function, which would predict that rIFC plays a dominant role in detection, but not anticipation, of inhibitory cues via interaction with rAI and sensory cortex. We demonstrate the crucial role of rAI in detecting behaviorally salient inhibitory cues in both auditory and visual modalities and highlight the role of rIFC in top-down modulation of rAI during inhibitory cue anticipation.

Materials and Methods

Overview of Data Sets

We used 2 previously published Stop-signal task data sets (Xue et al. 2008; Zhang and Li 2012), with Stop signals presented in 2

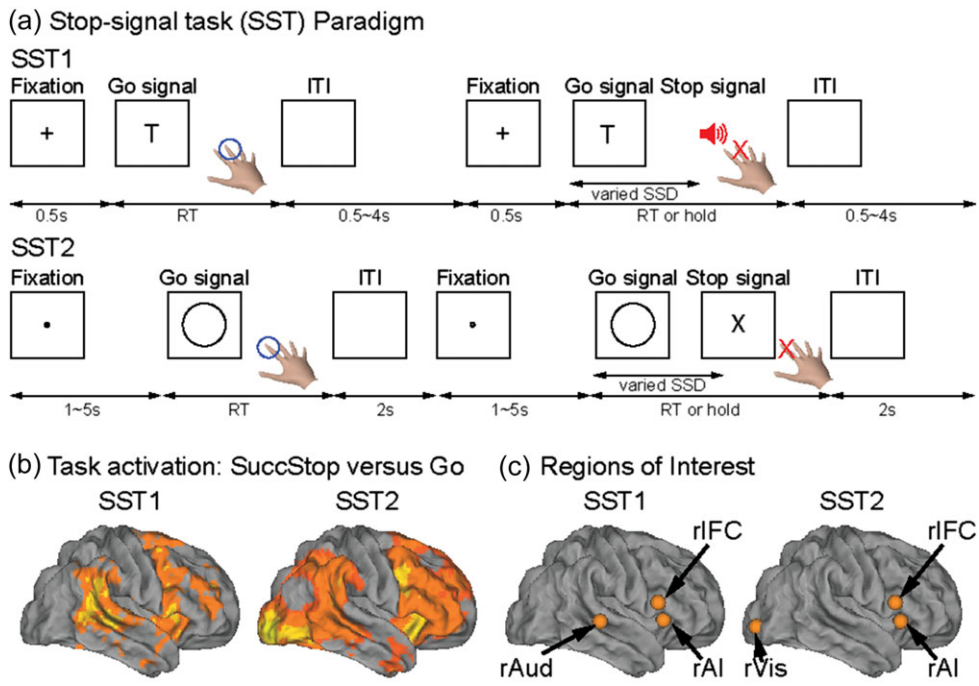

Figure 1. Task design and brain activations. (*a*) SST paradigms SST1 and SST2. The design of the 2 tasks is similar except that an auditory stop cue was used in SST1, while a visual stop cue was used in SST2. (*b*) Increased activation of FOI, sensory and other cortical regions during successful stop trials (SuccStop) compared with Go trials during SST1 and SST2. (*c*) Regions of interest (ROIs), based on SST1 and SST2 task activations, which were used in the DCM analysis.

different sensory modalities. All participants were healthy adults and they provided written consent to their local institutional review boards. Details of task design and data acquisition for the SST1 and SST2 can be found in Supplementary Information (Fig. 1).

Preprocessing of fMRI Data

SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8) was used for preprocessing, including realignment, slice-timing correction, normalization to the Montreal Neurological Institute space, and smoothing carried out using a 5-mm full-width half maximum Gaussian kernel to decrease spatial noise.

General Linear Models of the Stop-Signal Task

We used 2 different general linear models (GLMs) to investigate task-related activation and dynamic causal interactions. The first "conventional" model was used to examine causal interaction between the sensory regions and the right FOI during detection of Stop signals on successful Stop trials only. The second "anticipation" model, in which trials were grouped by participants' subjective expectation of Stop signals, allowed us to investigate causal interactions between sensory regions and right FOI during detection and anticipation of stop signals on successful Stop trials. Details of each model are described as follows.

Conventional Model

First, we examined causal interactions between the right FOI and sensory regions when a Stop signal was detected and the prepotent response was successfully canceled. To obtain the effect of interest for successful Stop trials only, a conventional task design construct of the Stop-signal task was used. The task design construct included successful Go (Go), successful Stop (SuccStop), and unsuccessful Stop (UnsuccStop) trials.

Anticipation Model—High versus Low P_{stop}

Second, we examined causal interactions between the right FOI and sensory regions modulated by anticipation of inhibitory cues p_{stop}—as described in the following section. To obtain the effect of interest for anticipation of inhibitory cues, the regressors in the conventional model were further split by the median p_{stop}, including Go with low p_{stop} (Go-low-p_{stop}), Go with high p_{stop} (Go-high-p_{stop}), SuccStop with low p_{stop} (SuccStop-low-p_{stop}), SuccStop with high p_{stop} (SuccStop-high-p_{stop}), UnsuccStop with low p_{stop} (UnsuccStop-low-p_{stop}), and UnsuccStop with high p_{stop} (UnsuccStop-high-p_{stop}). Trial-by-trial p_{stop} was estimated using the BPM (see below).

All GLM models incorporated 6 motion parameters as covariates of no interest. ROI time series were extracted for the subsequent DCM analyses using an *F* contrast to remove confounds of nuisance regressors.

BPM of P_{stop}

We used a well-validated BPM (Shenoy and Yu 2011; Ide et al. 2013) to investigate brain responses and dynamic causal interactions associated with trials with high and low inhibitory cue probability p_{stop}. Here we briefly describe the BPM. More detailed information and its validation can be found in previous studies (Shenoy and Yu 2011; Ide et al. 2013).

The BPM estimates the belief about the chance of an inhibitory cue occurring in the coming trial based on trial history (Yu and Cohen 2008). On an incoming trial *k*, subjects believe that the chance that an inhibitory cue will occur (Stop trial) is r_k and the chance that no inhibitory cue will occur (Go trial) is $1–r_k$. The model assumes that subjects believe that r_k has

a probability α of being the same as r_{k-1} (the chance that an inhibitory cue occurs in the previous trial) and a probability $1-\alpha$ of being resampled from the prior distribution $\pi(r_k)$:

$$p(r_k|s_{k-1}) = \alpha \times p(r_{k-1}|s_{k-1}) + (1-\alpha) \times \Pi(r_k),$$

where s_k refers to the true trial type of trial k ($s_k = 1$ for Stop trial, $s_k = 0$ for Go trial); $p(r_{k-1}|s_{k-1})$ refers to the posterior distribution conditional on the last observed trial $k-1$; $\pi(r_k)$ is assumed to be a β distribution with prior mean (pm) and another shape parameter (scale).

The model also assumes that subjects update their prior belief using Bayesian inference, and therefore the posterior distribution is computed based on Bayes' rule:

$$p(r_k|s_k) \propto p(s_k|r_k) \times p(r_k|s_{k-1}).$$

The probability of trial k being a Stop trial is determined by

$$\begin{aligned} P(s_k = 1|s_{k-1}) &= \int P(s_k = 1|r_k) \times p(r_k|s_{k-1}) dr_k \\ &= \int r_k \times p(r_k|s_{k-1}) dr_k = (r_k|s_{k-1}). \end{aligned}$$

In summary, the model allows us to estimate trial-by-trial anticipation of inhibitory cues p_{stop} based on subjects' trial history (Go or Stop trials). Here we used the same model parameters {α, pm, scale} as in the previous study since they have been well validated on the same data set (Ide et al. 2013).

Behavioral and Brain Analysis of High versus Low-P_{Stop} Stop Trials

We investigated how behavior was modulated by prediction of the likelihood of an inhibitory cue (stop signal) in the upcoming trial. p_{stop} was computed using the BPM as described earlier. Then, we conducted whole-brain analysis for the contrast of SuccStop-low-p_{stop} versus SuccStop-high-p_{stop} and further analysis on regions of interest.

Regions of Interest

We constructed 4 ROIs, including the rAI, rIFC, and 2 sensory regions, to study causal interactions between the right FOI and sensory regions in Stop tasks. We leveraged our previous work on functional parcellation of the right FOI (Cai et al. 2014) to construct the specific rAI and rIFC ROIs used in this study. Using the same data sets, our previous study identified 2 functional masks for the rAI and rIFC (Figure S1). In both SST1 and SST2, we first identified peak activations in successful Stop trials within the rAI and rIFC functional masks then averaged coordinates in the SST1 and SST2 for the rAI ($x = 35$, $y = 17$, $z = 3$) and rIFC ($x = 56$, $y = 13$, $z = 16$) separately. Because an auditory and visual Stop signal was used in the SST1 and SST2, we constructed 2 sensory ROIs. Peak activations were identified in the right auditory cortex (rAud, $x = 64$, $y = -30$, $z = 2$) in SST1 and in the right visual cortex (rVis, $x = 34$, $y = -90$, $z = -2$) in SST2, respectively. All ROIs are spheres of 6 mm radius (Fig. 1).

We also conducted several additional analyses to examine interactions of the left sensory cortex (left V1 and left A1) with right FOI during detection and anticipation of stop signals. We found similar patterns of causal interactions with right FOI for left and right sensory cortex (see Supplementary Information for details).

Dynamic Causal Modeling

We used the DCM module in SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) to examine causal interactions between the rAI, rIFC, and rAud in SST1 and rVis in SST2.

DCM is a quantitative approach to infer the causal architecture in a network constructed by a set of predefined brain regions in the latent state given the observed brain signals in a cognitive task (Friston et al. 2003). It is a confirmatory method where several candidates of causal models are tested and the model with the highest evidence is chosen.

We performed DCM analyses in the following steps. First, a set of DCM candidate models was determined by combinations of 3 matrices, which defined fixed/intrinsic connectivity, modulatory connectivity, and extrinsic/driving input (see below), respectively. Entries of the 3 matrices were then estimated for each DCM candidate model in each session for each participant using Bayesian approximation given the mean fMRI time series from the corresponding ROIs. Next, at the group level, families of DCM models were compared and the winning family was selected using Bayesian model selection (BMS) with random effects (Stephan et al. 2009). BMS estimated the posterior model probability, referring to the probability that a model is true given the observed data, and the exceedance probability, referring to the probability that a model candidate is more likely than other model candidates. The entries of the 3 matrices in the winning family for the group were computed using Bayesian parameter averaging across the specified DCM models belonging to the wining family (Stephan et al. 2009).

Each DCM model is determined by 3 matrices: A, B, and C. Matrix A represents fixed/intrinsic connectivity among ROIs. Matrix B represents modulatory connectivity among ROIs under a certain task condition. Matrix C represents extrinsic/driving input to the predefined ROIs. In this study, we assumed reciprocal fixed connectivity among all 3 ROIs (i.e. full connection Matrix A) and systematically varied modulatory connectivity among the 3 ROIs (Matrix B). Matrix C was set such that extrinsic inputs (i.e. Stop signal) went through sensory cortex.

Matrix B estimates how and to what extent specific cognitive events change connectivity between regions. Since our study aims to examine causal interactions between sensory cortex and right fronto-operculo-insular regions in detection of Stop signals, all possible models are grouped into 3 model families distinguished by whether there is modulatory connectivity from sensory cortex to the rAI alone, from the sensory cortex to the rIFC alone, or from the sensory cortex to both the rAI and rIFC, when a Stop signal was detected (Figure S2). The 3 families are named "rVis/rAud → rAI," "rVis/rAud → rIFC," and "rVis/rAud → rAI & rIFC." In each family, the rest of connections were systematically varied, including "rAI → rVis/rAud," "rIFC → rVis/rAud," "rAI → rIFC," and "rIFC → rAI." The full combination leads to 13 models for each family, resulting in 39 models for each session in each participant.

DCM analyses based on the 39 models were applied to the 2 task design constructs for SST1 and SST2 separately. First, using task design in the conventional model, we examined dominant causal interactions between sensory cortex and the right fronto-operculo-insular regions on SuccStop during which each participant successfully detected the Stop signal and canceled preprotent responses. We predicted that causal interactions between sensory cortex and rAI would play a dominant role in detection of Stop signals on SuccStop trials. Then, using the task design in the anticipation model, we examined whether the strength of causal interactions between the rAI and rIFC was modulated by

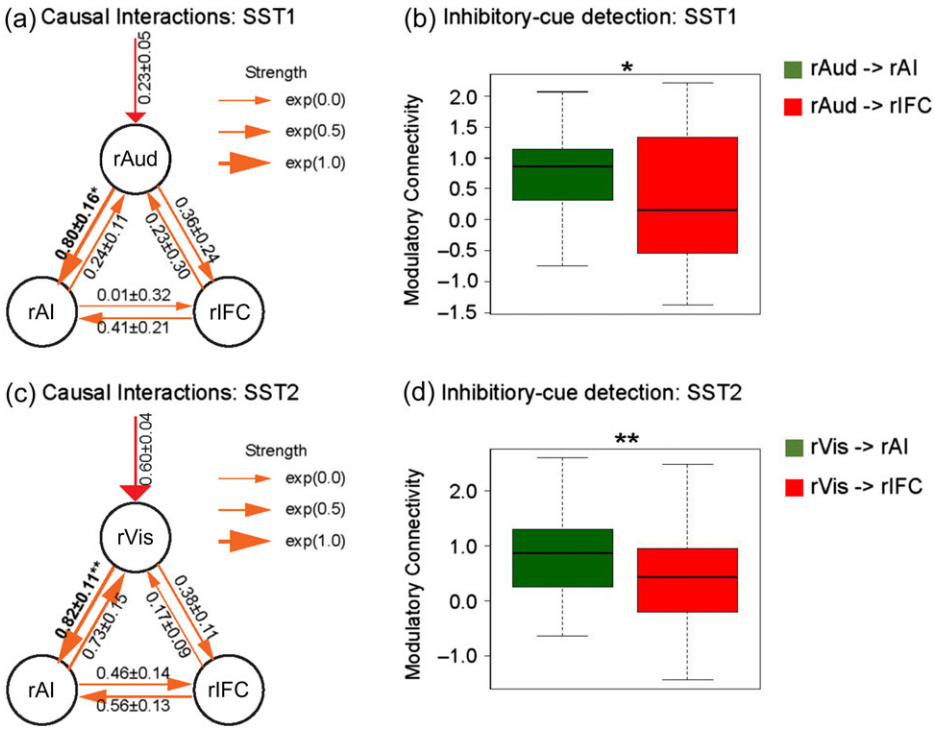

Figure 2. Dynamic causal interactions of the FOI during detection of inhibitory cues. (*a*, *b*) In SST1, the strength of causal interactions from the auditory cortex was significantly greater to the AI, compared with the IFC. (*c*, *d*) In SST2, the strength of causal interactions from the visual cortex was significantly greater to the AI, compared with the IFC. *$P < 0.05$; **$P < 0.01$. SST1, Stop-signal task 1; SST2, Stop-signal task 2; r, right. Line and arrow width is set as exponential power of estimated weight.

probabilistic anticipation of Stop signals. We predicted increased causal interactions from rIFC to rAI with high-probabilistic anticipation of Stop signals. Five thousand permutations were performed to generate a null distribution for the connectivity strength difference between 2 connections of interest from which the corresponding *P* value was obtained.

Results

Dynamic Interactions Between Sensory Cortex and FOI Control System Associated with Detection of Salient Inhibitory Cues

The first major goal of our study was to probe causal interactions between task-specific sensory cortex and rAI and rIFC subdivisions of right FOI associated with detection of inhibitory cues. We used DCM, a confirmatory method for examining causal interactions among a small number of task-relevant ROIs. In this approach, several plausible models are tested, and the model with the highest evidence is chosen as the one most likely to explain the data given the data and model assumptions.

We conducted DCM analyses in SST1 and SST2 with 2 distinct sensory inputs from rAud and rVis, respectively. At the group level, a BMS procedure with random effects (Stephan et al. 2009) was used to identify the "winning" model family with the highest evidence. We used these models to test the prediction that inhibitory cue detection is dominated by causal interactions between domain-specific sensory regions and rAI.

Causal Interactions Between Auditory Cortex and rAI/rIFC During Successful Stopping in SST1

We examined causal interactions between auditory cortex and rAI and rIFC on SuccStop trials in SST1. BMS revealed that the winning family was "rAud → rAI & rIFC" (Figure S3*a*). The posterior probabilities were 13%, 41%, and 46% for the "rAud → rAI," "rAud → rIFC," and "rAud → rAI & rIFC" families, respectively. The exceedance probabilities were 2%, 42%, and 56%, respectively. Aggregating the models within the winning family, we specifically tested the strength of modulatory connectivity of rAud → rAI versus that of rAud → rIFC (Fig. 2*a*) and found that the strength of rAud → rAI connectivity was significantly greater than that of rAud → rIFC ($P < 0.05$, Fig. 2*b* & Figure S5*a*).

Causal Interactions Between the Visual Cortex and rAI/rIFC During Successful Stopping in SST2

We examined causal interactions between the visual cortex and the rAI and rIFC on SuccStop trials in SST2. BMS revealed that the winning family was "rVis → rAI & rIFC" (Figure S3*b*). The posterior probabilities were 32%, 18%, and 50% for the "rVis → rAI," "rVis → rIFC," and "rVis → rAI & rIFC" families, respectively. The exceedance probabilities were 13.2%, 0.5%, and 86.3%, respectively. Aggregating the models within the winning family, we specifically tested the strength of modulatory connectivity of rVis → rAI versus rVis → rIFC (Fig. 2*c*) and found that the strength of rVis → rAI connectivity was significantly greater than that of rVis → rIFC ($P < 0.001$, Fig. 2*d* & Figure S5*b*).

Dynamic Interactions of FOI Control System Associated with Anticipation of Inhibitory Cues

The second goal of our study was to investigate how causal interactions between domain-specific sensory regions and subdivisions of right FOI during successful stopping are modulated by subjective expectations of stop signals. Trial-by-trial

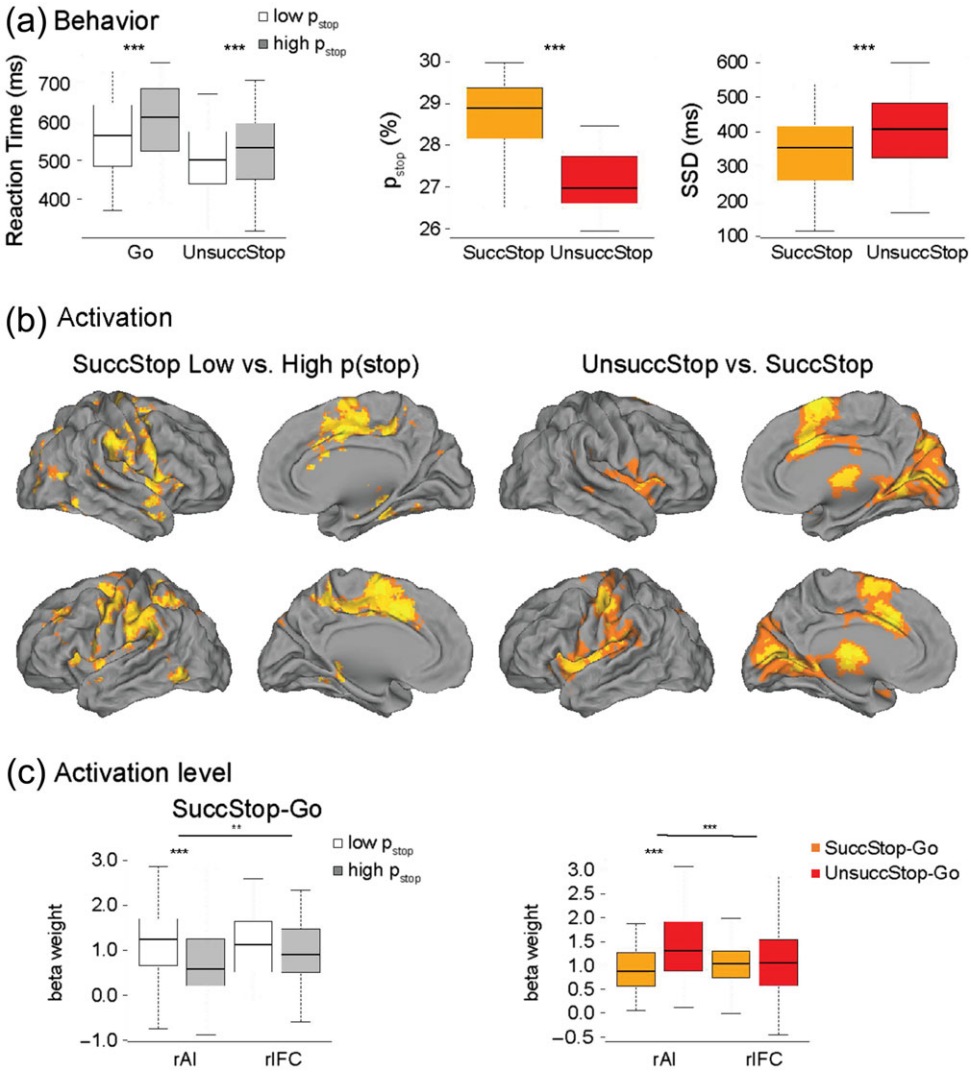

Figure 3. Bayesian prediction of behavior and brain activation associated with anticipation of inhibitory cues. (*a*) Participants showed faster RTs on Go and unsuccessful stop trials (UnsuccStop) with low, compared with high, probability of expecting an inhibitory cue (low-p_{stop} vs. high-p_{stop}; "left panel"). Participants had higher stop-signal expectation ("middle panels"), and had shorter SSD times on successful, compared with unsuccessful stop trials (SuccStop vs. UnsuccStop; "right panel"). (*b*) Bilateral activation of AI and IFC associated with successfully stopping on trials with low versus high probability of expecting an inhibitory cue ("left panel"), as well as unsuccessful, compared with successful stop trials (UnsuccStop vs. SuccStop, "right panel"). (*c*) Activation levels in unbiased AI and IFC ROIs from a previous study of right FOI parcellation (Cai et al., 2014). $^{**}P < 0.01$; $^{***}P < 0.001$.

probabilistic stop-signal expectation p_{stop} was estimated using BPM, and SuccStop trials were split by median p_{stop} to allow contrasting of "behavior, brain response, and dynamic causal interaction" between SuccStop-low-p_{stop} and SuccStop-high-p_{stop}. This analysis was conducted only in SST2 because of the small sample size and small number of SuccStop-low-p_{stop} trials in SST1. Our hypothesis is that anticipation of stop signals or proactive control involves top-down modulation of the salience detection system via interaction between rIFC and rAI.

Behavioral Effects of Inhibitory Cue Anticipation in SST2

As predicted, we found that reaction times (RTs) on go trials with high p_{stop} (mean ± standard error of mean: 599 ± 14 ms) were significantly longer than RTs on go trials with low p_{stop} (561 ± 14 ms) ($t = 9.37$, $P < 0.001$, Cohen's $d = 1.35$, Fig. 3*a* & Figure S6*a*, left panel). RTs on UnsuccStop trials with high p_{stop} (526 ± 14 ms) were also significantly longer than those on UnsuccStop trials with low p_{stop} (501 ± 13 ms) ($t = 4.35$, $P < 0.001$, Cohen's $d = 0.63$ Fig. 3*a* & Figure S6*a*, right panel). These results suggest that as the expectation of stop signals increases, participants slow down their responses.

Average p_{stop} was significantly higher on SuccStop (29 ± 0.1%) than on UnsuccStop trials (27 ± 0.1%) ($t = 11.65$, $P < 0.001$, Cohen's $d = 1.68$, Fig. 3*a* & Figure S6*b*). Note that average stop-signal delays (SSDs) were also significantly shorter on SuccStop (344 ± 15 ms) than on UnsuccStop trials (402 ± 15 ms) ($t = 33.21$, $P < 0.001$, Cohen's $d = 4.79$, Fig. 3*a* & Figure S6*c*). These findings suggest that successful stopping is influenced by how fast stop signals are detected, which could be facilitated by (1) higher expectation of stop signals so that participants are better prepared to detect stop signals and (2) shorter stop-signal delays such that earlier stop signals allowed more time to cancel a prepotent response.

Brain Responses to Inhibitory Cue Anticipation in SST2

Next, we compared brain responses to stop trials with high and low p_{stop}. We found that rAI, along with anterior cingulate cortex (ACC), showed significantly greater activation on SuccStop trials with low, compared with high, p_{stop} (activation height $P < 0.01$, cluster size $P < 0.01$, Fig. 3*b*). There was a much weaker activation difference on SuccStop trials with low, compared with high, p_{stop} for rIFC. A similar pattern of stronger activation in rAI and ACC and weaker activation in rIFC was found for UnsuccStop versus SuccStop trials (activation height $P < 0.01$, cluster size $P < 0.01$, Fig. 3*b*).

To further confirm the findings of dissociations in causal interactions associated with rAI and rIFC, we conducted additional ROI analyses. Because our goal is to test the anatomical specificity of findings from our previous work, we used the same rAI and rIFC ROIs from our previous parcellation study (Cai et al. 2014) (Figure S1*b*). We performed an ANOVA with factors of Region (rAI vs. rIFC) and p_{stop} (low vs. high). As in the previous analysis, beta weights contrasting activation between SuccStop and go trials were used. We found a significant Region x p_{stop} interaction ($F = 8.37$, $P < 0.01$) and a significant main effect of p_{stop} ($F = 10.37$, $P < 0.005$) (Fig. 3*c*). Post hoc analysis revealed that rAI showed significantly greater activation for low, compared with high, p_{stop} trials ($t = 4.11$, $P < 0.001$, Cohen's $d = 0.59$), while this difference was only marginally significant for rIFC ($t = 1.90$, $P = 0.06$, Cohen's $d = 0.27$). We then used an ANOVA to examine the interaction between Region (rAI vs. rIFC) and activation on SuccStop versus UnsuccStop trials. We found a significant Region × Trial type interaction ($F = 31.83$, $P < 0.001$) and significant main effects of Region ($F = 4.54$, $P < 0.05$) and Trial type ($F = 10.69$, $P < 0.01$) (Fig. 3*c*). Post hoc analysis revealed significantly greater activation on UnsuccStop, compared with SuccStop, trials in rAI ($t = 4.66$, $P < 0.001$, Cohen's $d = 0.67$) but not rIFC ($t = 0.64$, $P = 0.5$, Cohen's $d = 0.09$).

Differential Causal Interactions Between rAI and rIFC Associated with Inhibitory Cue Anticipation in SST2

Finally, we compared the strength of causal interactions between visual cortex and rAI and rIFC on SuccStop trials with low, compared with high, p_{stop} in the SST2 task (anticipation model). BMS revealed that the winning family was "rVis → rAI & rIFC" (Figure S4). The posterior probabilities were 19%, 14%, and 67%, and the exceedance probabilities were 0.02%, 0%, and 99.8% for the "rVis → rAI," "rVis → rIFC," and "rVis → rAI & rIFC" families, respectively. Aggregating the models within the winning family, we found a strong interaction between factors p_{stop} (high vs. low) and causal direction (rAI → rIFC vs. rIFC → rAI) ($P < 0.01$, Fig. 4*a*). Specifically, the strength of rIFC → rAI connectivity was significantly greater on SuccStop trials with high, compared with low, p_{stop} ($P < 0.01$, Fig. 4*b*).

Interestingly, the strength of extrinsic input into visual cortex was significantly greater on SuccStop trials with low, compared with high, p_{stop} ($P = 0.01$), which suggests that infrequent stop signals have a greater influence on the rVis-rAI-rIFC network when participants are less prepared to detect stop signals.

Discussion

We investigated the differential roles of rAI and rIFC subdivisions of right FOI in detection and anticipation of inhibitory cues. A novel aspect of our study is the analysis of dynamic causal interactions of rAI and rIFC with primary sensory regions and how these interactions change with subjective anticipation of inhibitory cues. Across 2 different experiments, we demonstrate that sensory regions have a stronger causal influence on rAI than on rIFC, suggesting that rAI plays a more important role in inhibitory cue detection. Crucially, we also found that the strength of causal influence from rIFC to rAI increased when participants had higher levels of anticipation of inhibitory cues. Our findings significantly advance knowledge of how a key FOI control system interacts with inputs from primary sensory regions and how these interactions are modulated by subjective changes in expectation of inhibitory cues.

Primary Sensory Cortex has Stronger Influences on rAI than on rIFC During Detection of Inhibitory Cues

While rAI and rIFC subdivisions of FOI have both been consistently implicated in detection of behaviorally salient and infrequent inhibitory cues, how these regions interact with sensory cortex is poorly understood. Here we show that primary sensory cortices have differential influences on rAI and rIFC during inhibitory cue detection. Specifically, the strength of causal influences to rAI from sensory regions is greater than that of influences to rIFC. Importantly, we replicated this finding in 2 independent studies in which stop signals were presented through different sensory channels. These results demonstrate that sensory inputs help dissociate the differential roles of the 2 subdivisions of FOI, with rAI playing a more dominant role in inhibitory cue detection during cognitive control.

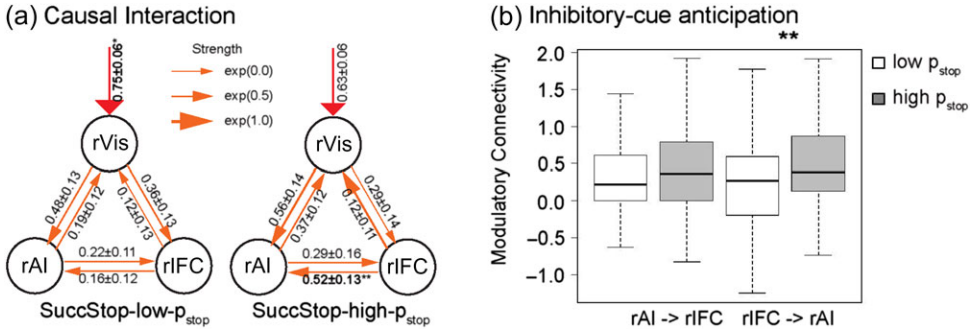

Figure 4. Dynamic causal interactions of the FOI associated with anticipation of inhibitory cues. (*a*) Causal interactions associated with successfully stopping on trials with low and high probability of anticipating inhibitory cues (SuccStop-low-p_{stop}, SuccStop-high-p_{stop}). (*b*) The strength of casual interactions from the IFC to AI was significantly greater during trials with high, compared with low, probability of anticipating inhibitory cues (low-p_{stop} vs high-p_{stop}). *$P < 0.05$; **$P < 0.01$. Line and arrow width is set as exponential power of estimated weight.

Our findings are relevant to a more general reconsideration of the role of rAI and rIFC in detection of behaviorally salient signals via interactions with sensory cortex. Previous studies of the classic "oddball" task have also consistently reported coactivation of rAI and rIFC during detection of infrequent targets (e.g. Linden et al. 1999). Like the stop-signal task, the oddball task requires inhibiting a frequent/dominant response followed by either withholding a response or initiating a different response (Crottaz-Herbette and Menon 2006). An influential model has proposed that the ventral attention system is specialized for stimulus-driven (or bottom-up) attention, with rIFC as the key prefrontal cortex node (Corbetta and Shulman 2002). The role of AI in stimulus-driven attention has, however, been less clear despite consistent reports of its coactivation with rIFC (Levy and Wagner 2011). Our findings are consistent with a network model, which posits that AI, a key node of the salience network, plays a more crucial role in detecting behaviorally important or salient signals and triggering access to the dorsal fronto-parietal attentional network (Menon and Uddin 2010). A previous study also revealed a strong causal influence from a multisensory processing region to rAI while detecting oddball signals (Chen et al. 2015). Our findings extend the current understanding of how rAI and rIFC interact with bottom-up sensory inputs.

rAI and rIFC Respond Differently to Low-Probability Inhibitory Cues

Inhibitory control can be modulated not only by task demands but also by trial-by-trial variations in inhibitory cue anticipation (Verbruggen and Logan 2009; Cai et al. 2011). The question of how fronto-operculo-insular inhibitory control regions dynamically adjust their interactions with sensory cortex in response to changing subjective inhibitory cue expectations has not received adequate attention. Our study was designed to test a key prediction related to previous findings of differential responses of rAI and rIFC during unsuccessful stop trials (Cai et al. 2014). It has been shown that participants make more errors on stop trials when the estimated subjective probability of stop signals is low (Ide et al. 2013). While rAI and rIFC show similar levels of activation on successful stop trials, rAI shows greater activation than rIFC on unsuccessful stop trials (Cai *et al.* 2014). Because stop signals occur late and unexpectedly during unsuccessful stop trials (Logan et al. 1984; Verbruggen and Logan 2008), greater activation is likely related to detection of these highly unexpected and salient signals (Ide et al. 2013). It is therefore likely that this pattern may extend more generally to trials with low inhibitory cue expectation. Indeed, we found that rAI shows greater activation than rIFC on successful stop trials with low relative to high inhibitory cue expectation. These findings demonstrate that rAI responses can be dissociated from those of rIFC and that rAI sensitivity to less expected events is important for successful inhibition. They also suggest a more crucial role for rAI in the bottom-up attention system during detection of surprising events, consistent with the hypothesized role of this region within the salience network (Menon and Uddin 2010; Cai et al. 2014; Uddin 2015).

Causal Influences from rIFC to rAI are Stronger in Response to High-Probability Inhibitory Cues

The next question we addressed was how dynamic interactions between rAI, rIFC, and sensory regions are altered by anticipation of inhibitory cues. We compared the strength of dynamic interactions on trials with low versus high stop-signal probability. We found that the strength of causal influences from rIFC to rAI was stronger on successful stop trials with high stop-signal probability. These results point to a "top-down" control signal from rIFC to rAI when participants believe that stop signals are more likely to occur, a finding consistent with the putative role of rIFC in proactive control (Aron 2011). Our findings are consistent with and extend results from previous studies, which have pointed to the involvement of rIFC in the preparatory phase of response inhibition (Chikazoe et al. 2009; Boehler et al. 2010; Jahfari et al. 2010; Swann et al. 2012). However, these previous studies have not examined the differential role of rAI or dynamic causal interactions between rAI and rIFC in proactive control. Our findings demonstrate that rIFC exerts stronger top-down modulation over rAI when stop-signal expectation is high, thereby facilitating detection and inhibition of responses to behaviorally salient inhibitory cues.

Dissociable Roles of rAI and rIFC in Inhibitory Control

Our study provides strong and replicable evidence that rAI and rIFC play different roles in anticipation and detection of behaviorally salient signals in inhibitory control. Analysis of context-specific dynamic causal interactions, together with comparison of activation levels, demonstrates that rAI is more involved in bottom-up salient inhibitory cue detection, especially when detecting less expected signals. In contrast, rIFC plays a more important role in top-down modulation of the rAI node of the salience network when expectation of inhibitory cue is high. This double dissociation has important implications for advancing our understanding of trial-by-trial variations in stimulus-driven detection and goal-relevant expectation during inhibitory control. We propose that when inhibitory cue expectation is low, rAI signals other brain regions in the cognitive control network to implement inhibitory control in a reactive control mode. In contrast, when anticipation of a need for inhibitory control is high, rIFC modulates rAI, thereby facilitating inhibitory cue detection and proactive control.

Our findings help address unresolved issues regarding the role of rIFC in inhibitory control versus attention capture. While most previous studies have treated rAI and rIFC subdivisions of FOI as one functionally homogeneous cluster, the present findings provide robust evidence for further dissociations between the functional roles of rAI and rIFC in inhibitory cue anticipation and detection. Thus, "inhibitory control" should not be viewed as a unitary construct but rather as a dynamical process of interactive top-down and bottom-up influences that vary on a trial-by-trial basis with changing expectations and surprise.

Limitations and Future Directions

Beyond the FOI, inhibitory control also involves several other cortical and subcortical regions (Levy and Wagner 2011; Swick et al. 2011; Cai et al. 2014). Extending analysis of dynamic causal interactions to a large number of nodes remains an important challenge for DCM because of the exponential increase in the number of models to be tested with a linear increase in the number of nodes. Future studies will need to extend this work to include pre-supplementary motor area, ACC, and basal ganglia to develop a more comprehensive computational model of the inhibitory control network. Finally, a limitation of DCM is that Bayesian evidence for the "winning" models may not be significantly higher than models with the second highest

evidence. However, the fact that the same winning model was chosen in multiple tests across 2 different data sets points to the robustness of our findings. As reproducibility is a major concern in all of neuroscience and particularly in fMRI studies (Button et al. 2013), it is more critical to demonstrate reproducibility and stability of causal interactions as we have.

Conclusions

Our findings provide new insights into dynamic causal mechanisms underlying inhibitory control in the human brain. We focused on rAI and rIFC, 2 key nodes of the salience and ventral attentional networks, and demonstrated their distinct roles in inhibitory control. We found that sensory inputs have a stronger causal influence on rAI during inhibitory cue detection. Importantly, we replicated this finding in 2 independent data sets and in 2 sensory domains. Furthermore, greater inhibitory cue anticipation is associated with stronger casual influences from rIFC to rAI. Our findings advance current knowledge of the functional architecture of neurocognitive systems involved in cognitive control and have the potential to inform investigations of cognitive control in disorders, such as ADHD, autism, and schizophrenia.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/

Funding

National Institutes of Health (grant number MH105625, NS086085, DA026990, DA023248, and AA021449); National Science Foundation (grant number BCS1309260).

Notes

We are grateful to OpenfMRI.org for data sharing, Dr Sheng Zhang for assistance with data acquired at the Yale University School of Medicine, and Rachel Rehert and Jonathan Nicholas for their careful proofreading of the manuscript. *Conflict of Interest*: None declared.

References

Aron AR. 2011. From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. Biol Psychiatry. 69:e55–68.

Aron AR, Fletcher PC, Bullmore ET, Sahakian BJ, Robbins TW. 2003. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nature Neuroscience. 6:115–116.

Boehler CN, Appelbaum LG, Krebs RM, Hopf JM, Woldorff MG. 2010. Pinning down response inhibition in the brain—conjunction analyses of the stop-signal task. Neuroimage. 52:1621–1632.

Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafo MR. 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 14:365–376.

Cai W, Oldenkamp CL, Aron AR. 2011. A proactive mechanism for selective suppression of response tendencies. J Neurosci. 31:5965–5969.

Cai W, Ryali S, Chen T, Li CS, Menon V. 2014. Dissociable roles of right inferior frontal cortex and anterior insula in inhibitory control: evidence from intrinsic and task-related functional parcellation, connectivity, and response profile analyses across multiple datasets. J Neurosci. 34:14652–14667.

Cai Y, Li S, Liu J, Li D, Feng Z, Wang Q, Chen C, Xue G. 2016. The role of the frontal and parietal cortex in proactive and reactive inhibitory control: a transcranial direct current stimulation study. J Cogn Neurosci. 28:177–186.

Casey BJ, Castellanos FX, Giedd JN, Marsh WL, Hamburger SD, Schubert AB, Vauss YC, Vaituzis AC, Dickstein DP, Sarfatti SE, et al. 1997. Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 36:374–383.

Chambers CD, Bellgrove MA, Stokes MG, Henderson TR, Garavan H, Robertson IH, Morris AP, Mattingley JB. 2006. Executive "brake failure" following deactivation of human frontal lobe. J Cogn Neurosci. 18:444–455.

Chen T, Michels L, Supekar K, Kochalka J, Ryali S, Menon V. 2015. Role of the anterior insular cortex in integrative causal signaling during multisensory auditory-visual attention. Eur J Neurosci. 41:264–274.

Chikazoe J, Jimura K, Hirose S, Yamashita K, Miyashita Y, Konishi S. 2009. Preparation to inhibit a response complements response inhibition during performance of a stop-signal task. J Neurosci. 29:15870–15877.

Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci. 3:201–215.

Crottaz-Herbette S, Menon V. 2006. Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. J Cognitive Neurosci. 18:766–780.

Friston KJ, Harrison L, Penny W. 2003. Dynamic causal modelling. Neuroimage. 19:1273–1302.

Hu S, Ide JS, Zhang S, Li CS. 2015. Anticipating conflict: Neural correlates of a Bayesian belief and its motor consequence. Neuroimage. 119:286–295.

Ide JS, Shenoy P, Yu AJ, Li CS. 2013. Bayesian prediction and evaluation in the anterior cingulate cortex. J Neurosci. 33:2039–2047.

Jahfari S, Stinear CM, Claffey M, Verbruggen F, Aron AR. 2010. Responding with restraint: what are the neurocognitive mechanisms? J Cogn Neurosci. 22:1479–1492.

Levy BJ, Wagner AD. 2011. Cognitive control and right ventrolateral prefrontal cortex: reflexive reorienting, motor inhibition, and action updating. Ann N Y Acad Sci. 1224:40–62.

Linden DE, Prvulovic D, Formisano E, Vollinger M, Zanella FE, Goebel R, Dierks T. 1999. The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. Cereb Cortex. 9:815–823.

Logan GD, Cowan WB, Davis KA. 1984. On the ability to inhibit simple and choice reaction time responses: a model and a method. J Exp Psychol Hum Percept Perform. 10:276–291.

Menon V. 2011. Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn Sci. 15:483–506.

Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct. 214:655–667.

Palaniyappan L, Simmonite M, White TP, Liddle EB, Liddle PF. 2013. Neural primacy of the salience processing system in schizophrenia. Neuron. 79:814–828.

Petersen SE, Posner MI. 2012. The attention system of the human brain: 20 years after. Annu Rev Neurosci. 35:73–89.

Shenoy P, Yu AJ. 2011. Rational decision-making in inhibitory control. Front Hum Neurosci. 5:48.

Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. 2009. Bayesian model selection for group studies. Neuroimage. 46:1004–1017.

Swann NC, Cai W, Conner CR, Pieters TA, Claffey MP, George JS, Aron AR, Tandon N. 2012. Roles for the pre-supplementary motor area and the right inferior frontal gyrus in stopping action: electrophysiological responses and functional and structural connectivity. Neuroimage. 59:2860–2870.

Swick D, Ashley V, Turken U. 2011. Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. Neuroimage. 56: 1655–1665.

Uddin LQ. 2015. Salience processing and insular cortical function and dysfunction. Nat Rev Neurosci. 16:55–61.

Uddin LQ, Supekar K, Lynch CJ, Khouzam A, Phillips J, Feinstein C, Ryali S, Menon V. 2013. Salience network-based classification and prediction of symptom severity in children with autism. JAMA Psychiatry. 70:869–879.

Verbruggen F, Logan GD. 2008. Response inhibition in the stop-signal paradigm. Trends Cogn Sci. 12:418–424.

Verbruggen F, Logan GD. 2009. Proactive adjustments of response strategies in the stop-signal paradigm. J Exp Psychol Hum Percept Perform. 35:835–854.

Wessel JR, Conner CR, Aron AR, Tandon N. 2013. Chronometric electrical stimulation of right inferior frontal cortex increases motor braking. J Neurosci. 33:19611–19619.

Xue G, Aron AR, Poldrack RA. 2008. Common neural substrates for inhibition of spoken and manual responses. Cereb Cortex. 18:1923–1932.

Yu AJ, Cohen J. 2008. Sequential effects: superstition or rational behavior? In: Koller D, Schuumans D, Bengio Y, Bottou L, editors. Advances in Neural Information Processing Systems. Vancouver: MIT. p. 1873–1880.

Zhang S, Li CS. 2012. Functional networks for cognitive control in a stop signal task: independent component analysis. Hum Brain Map. 33:89–104.