The right inferior frontal gyrus as pivotal node and effective regulator of the basal ganglia-thalamocortical response inhibition circuit

Qian Zhuang[a,b#], Lei Qiao[c#], Lei Xu[a,d], Shuxia Yao[a], Shuaiyu Chen[b], Xiaoxiao Zheng[a,e], Jialin Li[a], Meina Fu[a], Keshuang Li[a,f], Deniz Vatansever[g], Stefania Ferraro [a], Keith M. Kendrick[a,g*], Benjamin Becker[a,*]

[a] The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Laboratory for Neuroinformation, Center for Information in Medicine, University of Electronic Science and Technology of China, Chengdu, China

[b] Center for Cognition and Brain Disorders, The Affiliated Hospital of Hangzhou Normal University, Hangzhou, Zhejiang Province, China

[c] School of Psychology, Shenzhen University, Shenzhen, China

[d] Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, China

[e] Brain Cognition and Brain Disease Institute (BCBDI), Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

[f] School of Psychology and Cognitive Science, East China Normal University, Shanghai, China

[g] Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China

[#]Qian Zhuang and Lei Qiao are joint first author

*Corresponding authors: Benjamin Becker (ben_becker@gmx.de) and Keith M. Kendrick (k.kendrick.uestc@gmail.com)

No. 2006, Xiyuan Ave., West Hi-Tech Zone, Chengdu, Sichuan 611731, China.

Phone: +86-28-61830811; Fax: +86-28-61830811

Abstract

The involvement of specific basal ganglia-thalamocortical circuits in response inhibition has been extensively mapped in the last few decades. However, the pivotal brain nodes and directed casual regulation within this inhibitory circuit in humans remains controversial. Here, we capitalize on recent progress in robust and biologically plausible directed causal modelling (DCM-PEB) and a large fMRI response inhibition dataset (n=218) to determine key nodes, their causal regulation and modulation via biological variables (sex) and inhibitory performance in the inhibitory control circuit encompassing the right inferior frontal gyrus (rIFG), caudate nucleus (rCau), globus pallidum (rGP) and thalamus (rThal).The entire neural circuit exhibited high intrinsic connectivity and an increasing rIFG inflow and its causal regulation over the rCau and rThal during response inhibition. In addition, sex and behavioral performance influenced the architecture of the regulatory circuits such that women displayed increased rThal self-inhibition and decreased rThal to GP modulation, while better inhibitory performance was associated with stronger rThal to rIFG communication. Furthermore, control analyses did not reveal a similar key communication in a left lateralized model. Together these findings indicate a pivotal role of the rIFG as input and causal regulator of subcortical response inhibition nodes.

Keywords: response inhibition, basal ganglia, thalamus, inferior frontal gyrus, effective connectivity, DCM, cognitive control, sex difference

Introduction

Animal models and human neuroimaging studies convergently demonstrated that inhibitory control critically relies on highly specific basal ganglia-thalamocortical circuits (Alexander et al., 1986, 1991; Alexander and Crutcher, 1990; Aron et al., 2007; Jahfari et al., 2019; Morein-Zamir and Robbins, 2015; Pfeifer et al., 2022; Schall and Godlove, 2012; Stuphorn, 2015; Verbruggen and Logan, 2009; Wei and Wang, 2016). Dysregulations in this circuit have been implicated in disorders characterized by inhibitory control deficits, including addiction (Klugah-Brown et al., 2020; Morein-Zamir and Robbins, 2015; Zhou et al., 2018), attention deficit/hyperactivity (ADHD, Morein-Zamir et al., 2014; Sonuga-Barke, 2005), schizophrenia (Camchong et al., 2006; Feng et al., 2018; Mamah et al., 2007) and Parkinson Disorder (DeLong and Wichmann, 2015; Obeso et al., 2000). The key nodes within this response inhibition circuitry have been extensively mapped with convergent evidence suggesting critical contributions from the pre-supplementary motor area (pre-SMA) and lateral prefrontal cortex (lPFC), particularly inferior frontal gyrus (IFG, Aron et al., 2003; Dambacher et al., 2014; Hampshire et al., 2010; Maizey et al., 2020; Schaum et al., 2021; Verbruggen and Logan, 2008; Zhang et al., 2017) as well as striatal regions, in particular the caudate and putamen (Eagle et al., 2011; Ghahremani et al., 2012; Hampton et al., 2017; Kelly et al., 2004; Ott and Nieder, 2019; Robertson et al., 2015; Robbins, 2007).

Anatomical and neurochemical studies further suggest that response inhibitory control within this circuitry is modulated by dopaminergic and noradrenergic signaling (Bari et al., 2011; Ghahremani et al., 2012; Li et al., 2020; Pfeifer et al., 2022; Rae et al., 2016; Robertson

et al., 2015). Dopamine receptor availability in the fronto-striatal circuits is significantly related to inhibition-related neural responses (Ghahremani et al., 2012; Pfeifer et al., 2022) and dopamine receptor availability in the lPFC modulates motor control via downstream regulatory projections to the striatum (Ott and Nieder, 2019; Vijayraghavan et al., 2016). Enhanced norepinephrine signaling facilitates response inhibition via modulation of the IFG and its connections with the striatum (Chamberlain et al., 2009; Rae et al., 2016), while the dorsal striatum represents an important locus of dopaminergic control of response inhibition (Ghahremani et al., 2012; Robertson et al., 2015) and the IFG plays an important role in top-down control of the basal ganglia regions (Buschman and Miller, 2014; Hampshire et al., 2010; Jahfari et al., 2012; Kim, 2014; Puiu et al., 2020; Renteria et al., 2018; Schaum et al., 2020; Tops and Boksem, 2011). In the basal ganglia-thalamocortical model of response inhibition (Alexander et al., 1986, 1991; Alexander and Crutcher, 1990) the thalamus relays information between the basal ganglia and cortex (Collins et al., 2018; Haber and Mcfarland, 2001; Haber and Calzavara, 2009; McFarland and Haber, 2002) - thus facilitating response inhibition and performance monitoring (Bosch-Bouju et al., 2013; Huang et al., 2018; Saalmann and Kastner, 2015; Tanaka and Kunimatsu, 2011) - via dense reciprocal connections with the basal ganglia and PFC (Guillery, 1995; Phillips et al., 2021; Xiao et al., 2009; Tanaka and Kunimatsu, 2011).

Convergent evidence from human lesion studies and neuroimaging meta-analyses demonstrates a right-lateralized inhibitory control network encompassing the right IFG (rIFG), right caudate nucleus (rCau), right globus pallidum (rGP) and right thalamus (rThal)

(Aron et al., 2003; Chevrier et al., 2007; Garavan et al., 1999; Hung et al., 2018; Jahfari et al., 2011; Thompson et al., 2021). However, while extensive research has highlighted the critical role of these regions within a right-lateralized inhibitory control circuitry, the causal information flow and critical contribution of single nodes within this network have not been determined.

We therefore capitalized on a novel dynamic causal modelling (DCM) approach based on a priori specification of biologically and anatomically plausible models which allows estimation of directed causal influences between nodes and their modulation by changing task demands (Friston et al., 2003; Stephan et al., 2010) in the largest sample to-date (n=218). DCM further allows comparison of modulatory effective connectivity strength across different experimental conditions using Bayesian contrasts (Dijkstra et al., 2017) and in combination with the recently developed Parametrical Empirical Bayes (PEB) hierarchical framework (DCM-PEB method) allows modeling of both commonalities and differences in effective connectivity between subjects e.g. to determine the neurobiological basis of sex and behavioral performance variations (Friston et al., 2016; Zeidman et al., 2019a; Zeidman et al., 2019b).

To determine the causal information flow and critical nodes in the basal ganglia-thalamocortical circuits and whether these are modulated by biological factors (i.e. sex) and show functional relevance in terms of associations with performance we capitalized on DCM-PEB in combination with functional magnetic resonance imaging (fMRI) data collected in a large sample of healthy individuals (n=218) during a well-established response inhibition

paradigm (emotional Go/NoGo task, see also Zhuang et al., 2021). To unravel the key nodes and causal influences within the inhibitory control network, we firstly estimated the effective connectivity between and within key regions involved in response inhibitory control within the rIFG-rCau-rGP-rThal functional circuit (right lateralized model) and secondly estimated sex differences and behavioral performance effects on connectivity parameters. To validate the hemispheric asymmetry of the inhibitory control network, an identical model of nodes was tested in the left hemisphere (left lateralized model).

Given convergent evidence on a pivotal role of the right IFG in mediating top-down cortical-subcortical control during response inhibition (Aron et al., 2003; Dambacher et al., 2014; Hampshire et al., 2010; Maizey et al., 2020), we predicted a greater modulatory effect on rIFG and its directed connectivity to both rCau and rThal in the NoGo compared to Go condition. Additionally, based on previous findings we expected a modulation of the key pathways by biological (i.e. sex, Li et al., 2006; Ribeiro et al., 2021; Sjoberg and Cole, 2018) and performance variations (Chang et al., 2020; Jahfari et al., 2011; Wei and Wang, 2016; Xu et al., 2016) with better response inhibition being associated with stronger causal regulation in the inhibition circuitry. Finally, we hypothesized a different causal structure for the left and right models given the hemispheric asymmetry in the inhibitory network (Aron et al., 2003; Chevrier et al., 2007; Hung et al., 2018; Jahfari et al., 2011; Thompson et al., 2021).

Results

Behavioral Results

The two-way repeated-measures ANOVA on accuracy found a significant main effect of inhibition ($F(1,115)=21.73$, $p<0.001$, $\eta_p^2=0.16$), with a higher accuracy for Go compared to No Go trials (Go trials: mean± SEM=98.47% ±0.31, No Go trials: mean ±SEM=70.34% ±1.44, Cohen's d=2.48). No sex-differences were found for accuracy or reaction times (ps>0.18).

BOLD Activation (GLM) Analysis

Examination of domain general inhibition (contrast: NoGo>Go) revealed a widespread fronto-parietal cortical and thalamo-striatal subcortical network including the IFG, striatal, pallidal and thalamic regions (**Figure 1 and Table 1**) during response inhibition. Group-level peaks in the rIFG, rCau, rGP and rThal were selected as centers of the ROIs for model testing (**Figure 2a**). No significant sex difference were observed in BOLD activation.

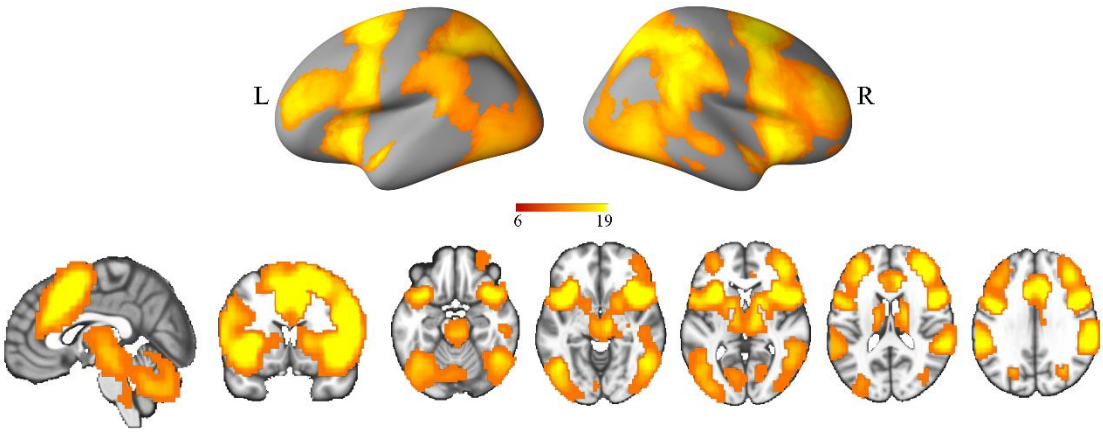

Figure 1. Brain activation maps for general response inhibition on whole brain level (contrast: NoGo > Go; p < 0.05 FWE, peak level). FWE, family-wise error; L, left; R, right.

Table 1. Activation and peak values for key regions included in the right model

Regions	Cluster K	Coordinates			t value
		X	Y	Z	
rIFG	611	51	12	18	21.40
rCau	144	15	-3	15	13.61
rGP	63	21	3	9	12.43
rThal	340	15	-6	12	14.30

Note: These clusters survived from the overlay between image masks of corresponding regions defined by Human Brainnetome Atlas ([84]) and group level brain activation maps (peak level, $p_{FWE} < 0.05$) and thus served as regions of interest combined with the individual peak location search on the individual level. FWE, family-wise error; Cau, caudate nucleus; GP, global pallidum; IFG, inferior frontal gyrus; r, right; Thal, thalamus.

Causal Connectivity (DCM) Analysis

For the A matrix, the diagonal cells represent self-connection which are unitless log scaling parameters and were multiplied with the default value of -0.5Hz (Zeidman et al., 2019a). Positive values indicate increased self-inhibition due to task condition and decreased responsivity to the inputs from the other regions of the network, while negative values indicate decreased self-inhibition and increased responsivity to the inputs from other nodes of the network (Zeidman et al., 2019a). Our findings revealed negative self-inhibition values for the rIFG, rCau and rThal but a positive value for the rGP (**Figure 2b, 2f**), indicating that the GP increased self-connection while the other nodes increased interaction with other nodes in the network.

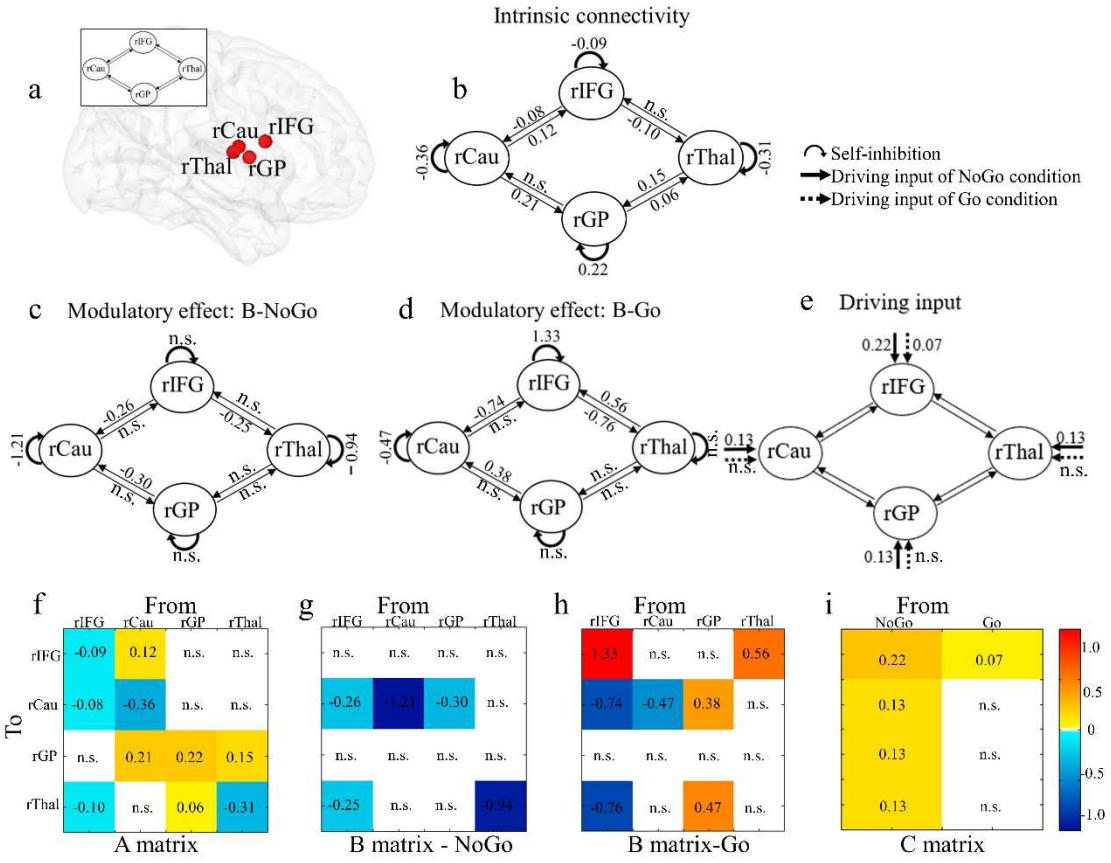

Figure 2. Location of regions included in the right model and group-level connectivity parameters. (a) Location of regions included in the right model. The A matrix: intrinsic connectivity across all experimental conditions (b, f). The B matrix: modulatory effect on effective connectivity between regions and self-inhibitions from NoGo (c, g) and Go condition (d, h). The C matrix: Driving inputs in ROIs in the NoGo and Go condition (e, i). Values in matrices reflect the connectivity parameters. Parameters with stronger evidence (posterior probability > 95%) are presented and subthreshold parameters are marked with "n.s.".

For the off-diagonal cells in the A matrix, the values (in Hz) reflect the rate of change in the activity of the target region caused by the source region per second. Positive values reflect

177 excitatory effects while negative values indicate inhibitory effects. In the forward direction
178 (e.g. rIFG-rThal-rGP-rCau-rIFG), we found a significant negative connectivity from rIFG to
179 rThal and positive connectivity from rThal to rGP as well as rCau to rIFG. In the backward
180 direction (e.g. rIFG-rCau-rGP-rThal-rIFG), rIFG exhibited a negative inhibitory influence
181 onto rCau, alongside an excitatory connection from rCau to rGP and rGP to rThal (**Figure 2b,**
182 **2f**). Although the connectivity from rThal to rIFG was not significant, a weak evidence
183 (posterior probability=57%) for this connection was observed with a more lenient threshold.
184 Values in the B matrix represent the rate of change, in Hz, in the connectivity from
185 source area to target area induced by the experimental conditions (Zeidman et al., 2019a).
186 During inhibitory control (NoGo condition) the rIFG exerted a negative influence onto the
187 rCau and rThal whereas the rGP exerted a negative influence on the rCau (**Figure 2c, 2g**). In
188 addition, we found negative self-inhibition values in both rCau and rThal respectively. During
189 the Go condition a negative influence of the rIFG on both rCau and rThal was observed
190 (**Figure 2d, 2h**), while the positive influence was observed from the rGP to rCau and from
191 rThal to rIFG. Moreover, we found a positive self-inhibition value in rIFG and a negative
192 value in rCau. A Bayesian contrast (NoGo > Go) allowed us to compare the connectivity
193 strength modulation during the different experimental conditions and revealed a very strong
194 evidence (posterior probability >99%) that the causal influence of the rIFG to both, the rCau
195 and rThal was stronger during inhibitory control (NoGo vs Go condition). This reflects that
196 response inhibition critically requires a causal top-down cortical-subcortical regulation via the
197 right IFG. We additionally found a very strong evidence (posterior probability >99%) for a

considerably stronger inhibitory connectivity from rGP to rCau in the NoGo compared to Go condition.

The C matrix represents the rate of change in neural response of one brain region due to the driving input from an experimental condition (Zeidman et al., 2019a). During inhibitory control (NoGo) all regions (rIFG, rCau, rGP and rThal) exhibited excitatory driving input while during the Go condition only the rIFG exhibited excitatory input (**Figure 2e, 2i**). Bayesian contrasts directly comparing the conditions (NoGo > Go) demonstrated an increasing driving input specifically in the rIFG during engagement of cognitive control (NoGo > Go condition) with a 100% posterior probability.

Sex Differences in Connectivity Parameters

Examining sex effects on intrinsic connectivity showed a negative influence from rThal to rGP in female compared to male subjects across all experimental conditions (**Figure 3a**). For the modulatory effects on connectivity, we found a greater self-inhibition in rThal in female than male subjects in the NoGo condition (**Figure 3b**). This suggests that for female subjects, rThal exhibits reduced sensitivity to inputs from the other regions of the selected network during response inhibition.

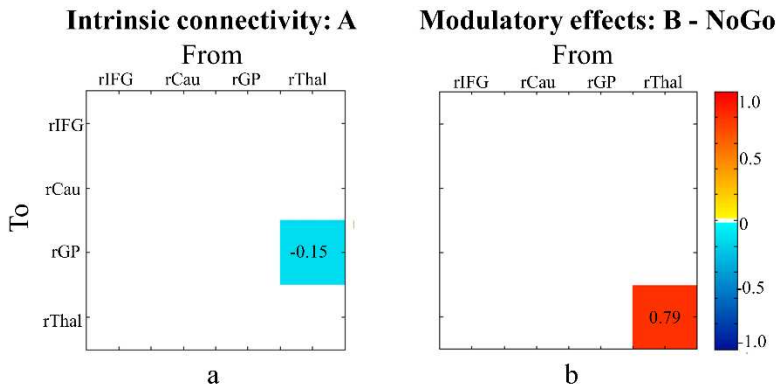

Figure 3. Sex effect on connectivity parameters in terms of A matrix and B

matrix. (a) For intrinsic connectivity in A matrix, female subjects showed a more negative influence from rThal to rGP compared to male subjects. (b) In the NoGo condition, there is a greater self-inhibition in rThal in female than male subjects in terms of B matrix. Parameters with stronger evidence (posterior probability > 95%) are presented.

Brain Behavior Associations: Inhibitory Behavioral Performance and Connectivity Parameters

Examining associations between inhibitory performance on the behavioral level (NoGo performance) and connectivity parameters revealed very strong evidence (posterior probability > 99%) that NoGo accuracy was positively associated with the directed connectivity from rThal to rIFG.

DCM Analyses in the Left Hemisphere

To further validate the hemispheric asymmetry of the inhibitory control network, an identical model for the left hemisphere including lIFG, lCau, lGP and lThal was tested. In contrast to the right model, no directed influences from IFG to subcortical regions were observed in terms of matrix A in the left model, and the hemispheric models different in terms of inhibition induced connectivity changes and differed in terms of the driving inputs. The different causal structure in the left and right model indicated a hemispheric asymmetry in the inhibition network (details see **Supplementary Materials Figure S3**). Additional Bayesian analyses confirmed the lack of a robust cortical-subcortical pathway in the left hemisphere (**Supplementary Material**).

Discussion

We capitalized on a combination of recent progress in biologically plausible causal hierarchical modelling (DCM-PEB) and a comparably large fMRI response inhibition dataset to determine causal information flow and key nodes within the extensively described basal ganglia-thalamocortical response inhibition circuits (Alexander et al., 1986, 1991; Alexander and Crutcher, 1990; Aron et al., 2007; Jahfari et al., 2019; Morein-Zamir and Robbins, 2015; Pfeifer et al., 2022; Schall and Godlove, 2012; Stuphorn, 2015; Verbruggen and Logan, 2009; Wei and Wang, 2016). Our neurocomputational model successfully validated a right-lateralized inhibitory control causal circuit and the best model showed significant intrinsic connectivity within this functional loop and captured an increasing causal influence of the cortical rIFG node on both the rCau and rThal as well as from the rGP to the rCau during inhibition. Direct comparison between different experimental conditions (e.g. NoGo and Go) revealed enhanced input into rIFG in terms of matrix C and increased connectivity from rIFG to rCau and rThal in the NoGo compared to the Go condition in terms of matrix B, suggesting a higher engagement of causal top-down cortical-to-subcortical control via the rIFG during inhibitory control. Although no sex differences were observed in inhibitory performance or BOLD activation, females exhibited decreased intrinsic connectivity from rThal to rGP and increased self-inhibition in rThal during the NoGo condition as compared to males. This indicates that similar behavioral performance in response inhibition might be mediated by different brain processes in men and women, particularly in thalamic loops. Moreover, a

higher NoGo response accuracy was associated with stronger causal information flow from the rThal to rIFG in the NoGo condition, suggesting a particular behavioral inhibitory relevance of this pathway. Finally, our findings showed different left and right model structures, suggesting a hemispheric asymmetry in the inhibitory control network and confirming a critical role of the rIFG in implementing response inhibition. Together these findings identified a pivotal role of the rIFG and its effective connectivity with the rCau/rThal within the basal ganglia-thalamocortical circuit during response inhibition.

Causal modelling successfully determined a right lateralized inhibitory control causal circuit encompassing the rIFG, rCau, rGP and rThal (Aron et al., 2003; Chevrier et al., 2007; Hung et al., 2018; Jahfari et al., 2011; Thompson et al., 2021). In terms of the A matrix, a significant rIFG-rCau-rGP-rThal loop was observed with rIFG exhibiting a negative influence onto rThal, alongside a positive information flow from from rThal to rGP and rCau to rIFG in the forward direction. In the backward direction, we found significant negative connectivity from rIFG to rCau and positive connectivity from rCau to rGP as well as rGP to rThal. A more lenient threshold additionally revealed rThal to rIFG connections (posterior probability = 57%). Importantly, accounting for behavioral task context revealed a significant positive modulatory effect on rIFG in both NoGo and Go condition in terms of matrix C which was considerably stronger during response inhibition. The direct driving inputs into the rIFG are in line with its role in top-down target detection and attentional control in the context of response inhibition (Hampshire et al., 2010; Krämer et al., 2013) and indicate that the rIFG represents the key regulator of other nodes. In line with this hypothesis the best model in

terms of matrix B revealed strong evidence for causal effective connectivity from the rIFG to both rCau and rThal during response inhibition (posterior probability >95%). This inhibitory pathway is consistent with previous reports on negative coupling between the rIFG and striatal regions during behavior control (Behan et al., 2015; Diekhof and Gruber, 2010). Notably, direct comparison using Bayesian contrast revealed a very strong evidence (posterior probability >99%) for increased modulatory connectivity from rIFG to rCau and rThal in the NoGo condition compared to the Go condition, suggesting the rIFG driven engagement of cortical-to-subcortical top-down control during response inhibition. Previous animal models and human neuroimaging meta-analyses have consistently identified the rIFG, as a key region implicated in dopaminergic and noradrenergic modulated inhibitory regulation (Bari et al., 2011; Hauber, 2010; Ott and Nieder, 2019; Pfeifer et al., 2022; Terra et al., 2020; Vijayraghavan et al., 2016; Zhukovsky et al., 2021) in particular during motor control and inhibition (Aron et al., 2003; Chamberlain and Sahakian, 2007; Puiu et al., 2020; Xu et al., 2016), while both, fronto-striatal and fronto-thalamic projections have been extensively involved in response inhibition (Ahissar and Oram, 2015; Bosch-Bouju et al., 2013; Marzinzik et al., 2008; Phillips et al., 2021; Schmitt et al., 2017; Sommer, 2003; Tanaka and Kunimatsu, 2011).

In addition to the cortical-subcortical pathways significant excitatory connectivity was observed from the rGP to rCau during the Go condition and switched to inhibitory connectivity when response inhibition was required during the NoGo condition. Direct comparison confirmed a considerably stronger inhibitory influence of the rGP on the rCau

during response inhibition (posterior probability >99%), suggesting that communication between basal ganglia nodes is crucial for context-appropriate behavioral response control. The involvement of this pathway is in line with extensive neurophysiological evidence showing that GABA inhibitory projections from the external segment of the GP to the striatum play an essential role in cancelling a planned response when it is inappropriate (Mallet et al., 2016; Wei and Wang, 2016) (but see also subthalamic nucleus to substantia nigra pars reticulata pathways in Hikosaka et al., 2006; Mallet et al., 2016).

With respect to sex differences we observed that females exhibited decreased intrinsic connectivity from rThal to rGP and increased modulation by NoGo condition on self-inhibition in rThal compared to male subjects in the absence of performance differences. While previous findings on sex differences in response inhibition remained inconsistent (Chung et al., 2020; Gaillard et al., 2020; Gaillard et al., 2021; Li et al., 2006; Ribeiro et al., 2021; Sjoberg and Cole, 2018) the present findings suggest that our model was sensitive to biological variables and that separable information processes may underly response inhibition in men and women (see also Chung et al., 2020; Li et al., 2006). The functional relevance of the identified pathways was further underscored by a significant association between response inhibition performance and the causal influence from the rThal to rIFG in the NoGo condition demonstrating that this pathway involved in motor inhibition critically mediates behavioral success during inhibition (Wei and Wang, 2016).

Finally, our modelling tests confirmed a hemispheric asymmetry and support the critical role of right IFG circuit in response inhibition (Hung et al., 2018; Jahfari et al., 2011; Maizey

et al., 2020). The different causal structures suggest a strong cortical-subcortical intrinsic connectivity and rIFG control on the right side. The left model revealed a different causal structure and null hypothesis tests showed moderate evidence for the difference between NoGo and Go condition's modulatory effects on effective connectivity from lIFG to lCau and to rThal (e.g. lIFG to lCau: Bayes factor = 5.47; lIFG to lThal: Bayes factor = 8.20).

Response inhibition impairments have been observed in several disorders and identification of the rIFG as critical input and top-down regulator for response inhibition opens new targets for regional or connectivity-based neuromodulation such as real-time neurofeedback which has been established for these regions (Li et al., 2019; Weiss et al., 2022; Zhao et al., 2019). For instance, rIFG and response inhibition deficits have been determined in ADHD (Clark et al., 2007; Morein-Zamir et al., 2014) and targeting the rIFG in ADHD may be a promising treatment.

There are several limitations in the current study. First, in line with our main aim we did not account for emotional valence in the DCM model which may affect response inhibition(Schimmack and Derryberry, 2005). Second, we focused on specific nodes that were based on established basal ganglia-thalamocortical circuits proposed by Alexander (Alexander et al., 1986, 1991; Alexander and Crutcher, 1990) (see also neuroimaging meta-analysis (Hung et al., 2018). Other regions such as the STN (Aron et al., 2016; Aron and Poldrack, 2006; Chen et al., 2020) could be integrated in future studies.

In conclusion, our findings demonstrated a critical role of the rIFG as well as top-down cortical-subcortical control from the rIFG to rCau and rThal in response inhibition. The nodes

and pathways of the model were sensitive to biological and performance variations. The nodes and pathways may represent promising targets to improve response inhibition in mental disorders.

Materials and Methods

Participants

N=250 healthy right-handed participants were enrolled in the current study and underwent a validated Go/NoGo fMRI paradigm. The data has been previously used to examine undirected functional connectivity within domain-general and emotion-specific inhibitory brain systems (Zhuang et al., 2021) and was part of larger neuroimaging project examining pain empathy (Li et al., 2018; Zhou et al., 2020), emotional face memory (Liu et al., 2022) and mirror neuron processing (Xu et al., 2022). After quality assessment n=218 subjects were included (104 males, **Supplementary Materials**). The study was approved by the local ethics committee and in accordance with the latest version of the Declaration of Helsinki.

Response Inhibition Paradigm

A validated mixed event-related block design linguistic emotional Go/NoGo fMRI paradigm was employed (Goldstein et al., 2007; Protopopescu et al., 2005, details see Zhuang et al., 2021). Participants were required to make responses as accurately and quickly as possible based on orthographical cues, i.e. words were presented in normal or italic font. For normal font words subjects were instructed to perform a button-press (Go trials), while inhibiting their response to words presented in italic font (NoGo trials). Positive, negative and neutral

words were included, however, given that the present study aimed to examine the causal influence within the general inhibition network and to increase statistical power in this respect the different emotional contexts were not further accounted for in the DCM analysis. Stimuli were presented in 2 runs and each run included 12 blocks (6 blocks: Go; 6 blocks: NoGo). Each Go block encompassed 18 normal font words (100% Go trials) while each NoGo block encompassed 12 normal font words (66.7% Go trials) and 6 italicized font words (33.3% NoGo trials). Further details in (Zhuang et al., 2021) and **Supplementary Materials**.

Behavioral Data Analysis

In our previous study we demonstrated that subjects exhibited more errors during inhibitory control (i.e., NoGo>Go) as well as faster responses in positive Go contexts and lower accuracy in positive NoGo contexts (Zhuang et al., 2021). Given that sex-differences were examined in the DCM model the present analyses additionally examined sex-differences on accuracy and reaction times (**Supplementary Materials)**. Given age-related effects on inhibition (Rey-Mermet et al., 2018; Rubia et al., 2007) age was included as covariate.

MRI Data Acquisition and Preprocessing

MRI data were collected on a 3T MRI system using standard sequences and were initially preprocessed using validated protocols in SPM 12 (details see **Supplementary Materials**)

GLM Analysis

An event-related general linear model (GLM) was established in SPM12. To examine domain general inhibitory control (irrespective of emotional context) the overarching inhibitory control contrast was modelled (e.g. all NoGo>all Go trials) and convolved with the canonical

hemodynamic response function (HRF). Six head motion parameters were included in the design matrix to control movement-related artifacts and a high-pass filter (1/128Hz) was applied to remove low frequency components. The contrast of interest (contrast: NoGo>Go) was created and subjected to one-sample t-test at the second level. In line with previous studies (Aron et al., 2003; Chevrier et al., 2007; Hung et al., 2018; Jahfari et al., 2011; Thompson et al., 2021), group-level (contrast: NoGo>Go) peaks in the IFG, Cau, GP and Thal within the identified general inhibition network were then used to define individual-specific regions of interest (ROIs) for the DCM analysis. Additionally, a two-sample t-test was conducted (contrast: NoGo>Go) to examine sex-dependent effects on the response inhibition network. Analyses were corrected for multiple comparisons using a conservative peak-level threshold on the whole brain level ($p<0.05$ family-wise error, FWE).

Dynamic Causal Modeling and Node Definition

A DCM analysis was employed to determine directed causal influences according to the circuitry model proposed by Alexander et al. (Alexander et al., 1986, 1991; Alexander and Crutcher, 1990). The DCM approach allows to construct a realistic neuronal model of interacting regions and to predict the underlying neuronal activity from the measured hemodynamic response (Friston et al., 2003; Stephan et al., 2007). To this end directed causal influences between the key regions including IFG, Cau, GP and Thal in the basal ganglia-thalamocortical loop and their modulation via experimental manipulations (engagement of motor inhibitory control) were examined. In line with previous neuroimaging studies and meta-analyses demonstrating a right-lateralized inhibition model (right model) encompassing

405 the rIFG, rCau, rGP, rThal (Aron et al., 2003; Chevrier et al., 2007; Hung et al., 2018; Jahfari
406 et al., 2011; Thompson et al., 2021) our main hypothesis testing focused on the right
407 lateralized network. To further validate the hemispheric asymmetry of the inhibitory control
408 network an identical model was tested for the left hemisphere including the lIFG, lCau, lGP,
409 and lThal. In line with previous studies, we combined atlas-based masks (Human
410 Brainnetome Atlas, Fan et al., 2016) with group-level and individual level activity maps to
411 generate the corresponding nodes (Fernández-Espejo et al., 2015; Holmes et al., 2021; Qiao et
412 al., 2020; Van Overwalle et al., 2020).

413 **Model Specification and Estimation**

414 A two-step DCM analysis was performed using the DCM-parametric empirical Bayes (PEB)
415 approach (Zeidman et al., 2019a; Zeidman et al., 2019b). On the first-level, time-series from
416 four ROIs (rIFG, rCau, rGP, rThal) were extracted. A full DCM model was specified for each
417 subject and all connectivity parameters in both forward (e.g. rIFG-rThal-rGP-rCau-rIFG)
418 and backward (e.g. rIFG-rCau-rGP-rThal-rIFG) directions were estimated. We estimated
419 three key DCM parameters: (1) the A matrix reflecting all connections including forward and
420 backward connectivity between ROIs and self-inhibitions in each ROI, (2) the B matrix
421 representing modulatory effects of Go and NoGo condition on all connections, (3) the C
422 matrix representing the driving inputs into ROIs from Go and NoGo conditions separately.
423 Given that all inputs in the model were mean-centered, intrinsic connectivity in the A matrix
424 indicates mean effective connectivity independent of all experimental conditions. The model
425 was estimated using Variational Laplace (Friston et al., 2007). Further details **Supplementary**

Material. At the second (group) level, we constructed a PEB model over the first-level estimated parameters. In accordance with previous studies (Bencivenga et al., 2021; Rupprechter et al., 2020), we evaluated the explained variance by the model on the individual level - higher values reflect better model inversion (Zeidman et al., 2019a) – and then we only included subjects with >10% of explained variance in the PEB model. A total of 118 subjects (56 males, age: mean ± SEM = 21.57 ± 0.21 years) were included for further analyses. The differences on behavioral performance were examined between the excluded and included subjects and no significant differences were found (all ps⩾0.23, for details see **Supplementary Material**), suggesting no evidence of biased selection.

The primary aim of the present study was to establish a causal neurobiological model for response inhibition and to determine the interaction between key players in this circuitry. To evaluate the model three PEB analyses were carried out separately for A, B and C matrices. Separate analyses examined sex and performance variations (details see **Supplementary Material**).

Next, to identify the model that best represented our data, Bayesian Model Reduction (BMR) was performed to compare the free energy of the full model with numerous reduced models for which specific parameters were “switched off” (Friston et al., 2016). An automatic greedy search procedure (iterative procedure) was employed to facilitate an efficient comparison of thousands of models. In this procedure parameters which do not contribute to free energy were pruned away. Next, Bayesian Model Average (BMA), performing a weighted average of the parameters of each model, was calculated over the 256 models

obtained from the final iteration (Friston et al., 2016).

Finally, to compare the effective connection strength, especially the cortical-subcortical connectivity and driving inputs into each region from different experimental conditions (NoGo and Go condition), Bayesian contrasts (Dijkstra et al., 2017) were computed over parameters from the B and C matrices. Group-level estimated parameters were thresholded at posterior probability > 95% (indicating strong evidence, Kass and Raftery, 1995) based on free energy.

Acknowledgements

This work was supported by the by the National Key Research and Development Program of China (grant number: 2018YFA0701400 - BB), National Natural Science Foundation of China (grant numbers 31530032 – KMK, 91632117 - BB), Key Technological Projects of Guangdong Province (grant number 2018B030335001 – KMK).

Author contributions

Qian Zhuang: Formal analysis; Investigation; Writing-original draft. **Lei Qiao**: Formal analysis; Validation. **Lei Xu**: Project administration. **Shuxia Yao**: Conceptualization; Supervision**. Shuaiyu Chen:** Formal analysis**. Xiaoxiao Zheng, Jialin Li, Meina Fu and Keshuang Li**: Project administration. **Deniz Vatansever and Stefania Ferraro:** Validation. **Keith M. Kendrick and Benjamin Becker:** Conceptualization; Funding acquisition; Project administration; Resources; Supervision; Validation; Writing.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with their research, authorship or the publication of this article.

References

Ahissar, E., and Oram, T. (2015). Thalamic relay or cortico-thalamic processing? Old question, new answers. Cerebral Cortex *25*, 845-848. 10.1093/cercor/bht296.

Alexander, G.E., and Crutcher, M.D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends in neurosciences *13*, 266-271. 10.1016/0166-2236(90)90107-l.

Alexander, G.E., Crutcher, M.D., and DeLong, M.R. (1991). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor,"prefrontal" and "limbic" functions. Progress in brain research *85*, 119-146. 10.1016/S0079-6123(08)62678-3

Alexander, G.E., DeLong, M.R., and Strick, P.L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual review of neuroscience *9*, 357-381. 10.1146/annurev.ne.09.030186.002041.

Aron, A.R., Durston, S., Eagle, D.M., Logan, G.D., Stinear, C.M., and Stuphorn, V. (2007). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. Journal of Neuroscience *27*, 11860-11864. 10.1523/JNEUROSCI.3644-07.2007. 10.1523/JNEUROSCI.3644-07.2007.

Aron, A.R., Fletcher, P.C., Bullmore, E.T., Sahakian, B.J., and Robbins, T.W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nature neuroscience *6*, 115-116. 10.1038/nn1003.

536 Aron, A.R., Herz, D.M., Brown, P., Forstmann, B.U., and Zaghloul, K. (2016).
537 Frontosubthalamic circuits for control of action and cognition. Journal of Neuroscience *36*,
538 11489-11495. 10.1523/JNEUROSCI.2348-16.2016.
539 Aron, A.R., and Poldrack, R.A. (2006). Cortical and subcortical contributions to stop signal
540 response inhibition: role of the subthalamic nucleus. Journal of Neuroscience *26*, 2424-2433.
541 10.1523/JNEUROSCI.4682-05.2006.
542 Bari, A., Mar, A.C., Theobald, D.E., Elands, S.A., Oganya, K.C., Eagle, D.M., and Robbins,
543 T.W. (2011). Prefrontal and monoaminergic contributions to stop-signal task performance in
544 rats. Journal of Neuroscience *31*, 9254-9263. 10.1523/JNEUROSCI.1543-11.2011.
545 Behan, B., Stone, A., and Garavan, H. (2015). Right prefrontal and ventral striatum
546 interactions underlying impulsive choice and impulsive responding. Human brain mapping
547 *36*, 187-198. 10.1002/hbm.22621.
548 Bencivenga, F., Sulpizio, V., Tullo, M.G., and Galati, G. (2021). Assessing the effective
549 connectivity of premotor areas during real vs imagined grasping: a DCM-PEB approach.
550 NeuroImage *230*, 117806. 10.1016/j.neuroimage.2021.117806.
551 Bosch-Bouju, C., Hyland, B.I., and Parr-Brownlie, L.C. (2013). Motor thalamus integration
552 of cortical, cerebellar and basal ganglia information: implications for normal and
553 parkinsonian conditions. Frontiers in computational neuroscience *7*, 163.
554 10.3389/fncom.2013.00163.

555 Buschman, T.J., and Miller, E.K. (2014). Goal-direction and top-down control. Philosophical
556 Transactions of the Royal Society B: Biological Sciences *369*, 20130471.
557 10.1098/rstb.2013.0471.
558 Camchong, J., Dyckman, K.A., Chapman, C.E., Yanasak, N.E., and McDowell, J.E. (2006).
559 Basal ganglia-thalamocortical circuitry disruptions in schizophrenia during delayed response
560 tasks. Biological psychiatry *60*, 235-241. 10.1016/j.biopsych.2005.11.014.
561 Chamberlain, S.R., Hampshire, A., Müller, U., Rubia, K., Del Campo, N., Craig, K.,
562 Regenthal, R., Suckling, J., Roiser, J.P., and Grant, J.E. (2009). Atomoxetine modulates right
563 inferior frontal activation during inhibitory control: a pharmacological functional magnetic
564 resonance imaging study. Biological psychiatry *65*, 550-555.
565 10.1016/j.biopsych.2008.10.014.
566 Chamberlain, S.R., and Sahakian, B.J. (2007). The neuropsychiatry of impulsivity. Current
567 opinion in psychiatry *20*, 255-261. 10.1097/YCO.0b013e3280ba4989.
568 Chang, J., Hu, J., Li, C.-S.R., and Yu, R. (2020). Neural correlates of enhanced response
569 inhibition in the aftermath of stress. Neuroimage *204*, 116212.
570 10.1016/j.neuroimage.2019.116212.
571 Chen, W., de Hemptinne, C., Miller, A.M., Leibbrand, M., Little, S.J., Lim, D.A., Larson,
572 P.S., and Starr, P.A. (2020). Prefrontal-subthalamic hyperdirect pathway modulates
573 movement inhibition in humans. Neuron *106*, 579-588. e573. 10.1016/j.neuron.2020.02.012.

574 Chevrier, A.D., Noseworthy, M.D., and Schachar, R. (2007). Dissociation of response
575 inhibition and performance monitoring in the stop signal task using event-related fMRI.
576 Human brain mapping *28*, 1347-1358. 10.1002/hbm.20355.
577 Chung, Y.S., Calhoun, V., and Stevens, M.C. (2020). Adolescent sex differences in cortico-
578 subcortical functional connectivity during response inhibition. Cognitive, Affective, &
579 Behavioral Neuroscience *20*, 1-18. 10.3758/s13415-019-00718-y.
580 Clark, L., Blackwell, A.D., Aron, A.R., Turner, D.C., Dowson, J., Robbins, T.W., and
581 Sahakian, B.J. (2007). Association between response inhibition and working memory in adult
582 ADHD: a link to right frontal cortex pathology? Biological psychiatry *61*, 1395-1401.
583 10.1016/j.biopsych.2006.07.020.
584 Collins, D.P., Anastasiades, P.G., Marlin, J.J., and Carter, A.G. (2018). Reciprocal circuits
585 linking the prefrontal cortex with dorsal and ventral thalamic nuclei. Neuron *98*, 366-379.
586 e364. 10.1016/j.neuron.2018.03.024.
587 Dambacher, F., Sack, A.T., Lobbestael, J., Arntz, A., Brugman, S., and Schuhmann, T.
588 (2014). A network approach to response inhibition: dissociating functional connectivity of
589 neural components involved in action restraint and action cancellation. European Journal of
590 Neuroscience *39*, 821-831. 10.1111/ejn.12425.
591 DeLong, M.R., and Wichmann, T. (2015). Basal ganglia circuits as targets for
592 neuromodulation in Parkinson disease. JAMA neurology *72*, 1354-1360.
593 10.1001/jamaneurol.2015.2397.

Diekhof, E.K., and Gruber, O. (2010). When desire collides with reason: functional interactions between anteroventral prefrontal cortex and nucleus accumbens underlie the human ability to resist impulsive desires. Journal of Neuroscience *30*, 1488-1493. 10.1523/JNEUROSCI.4690-09.2010.

Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M.A., and Friston, K. (2017). Distinct top-down and bottom-up brain connectivity during visual perception and imagery. Scientific reports *7*, 1-9. 10.1038/s41598-017-05888-8.

Eagle, D.M., Wong, J.C., Allan, M.E., Mar, A.C., Theobald, D.E., and Robbins, T.W. (2011). Contrasting roles for dopamine D1 and D2 receptor subtypes in the dorsomedial striatum but not the nucleus accumbens core during behavioral inhibition in the stop-signal task in rats. Journal of Neuroscience *31*, 7349-7356. 10.1523/JNEUROSCI.6182-10.2011.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., and Laird, A.R. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. Cerebral cortex *26*, 3508-3526. 10.1093/cercor/bhw157.

Feng, C., Becker, B., Huang, W., Wu, X., Eickhoff, S.B., and Chen, T. (2018). Neural substrates of the emotion-word and emotional counting Stroop tasks in healthy and clinical populations: A meta-analysis of functional brain imaging studies. NeuroImage *173*, 258-274. 10.1016/j.neuroimage.2018.02.023.

Fernández-Espejo, D., Rossit, S., and Owen, A.M. (2015). A thalamocortical mechanism for the absence of overt motor behavior in covertly aware patients. JAMA neurology *72*, 1442-1450. 10.1001/jamaneurol.2015.2614.

615 Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W. (2007). Variational
616 free energy and the Laplace approximation. Neuroimage *34*, 220-234.
617 10.1016/j.neuroimage.2006.08.035.
618 Friston, K.J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. Neuroimage
619 *19*, 1273-1302. 10.1016/s1053-8119(03)00202-7.
620 Friston, K.J., Litvak, V., Oswal, A., Razi, A., Stephan, K.E., Van Wijk, B.C., Ziegler, G., and
621 Zeidman, P. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies.
622 Neuroimage *128*, 413-431. 10.1016/j.neuroimage.2015.11.015.
623 Gaillard, A., Fehring, D.J., and Rossell, S.L. (2021). A systematic review and meta-analysis
624 of behavioural sex differences in executive control. European Journal of Neuroscience *53*,
625 519-542. 10.1111/ejn.14946.
626 Gaillard, A., Rossell, S.L., Carruthers, S.P., Sumner, P.J., Michie, P.T., Woods, W., Neill, E.,
627 Phillipou, A., Toh, W.L., and Hughes, M.E. (2020). Greater activation of the response
628 inhibition network in females compared to males during stop signal task performance.
629 Behavioural brain research *386*, 112586. 10.1016/j.bbr.2020.112586.
630 10.1073/pnas.96.14.8301.
631 Garavan, H., Ross, T., and Stein, E. (1999). Right hemispheric dominance of inhibitory
632 control: an event-related functional MRI study. Proceedings of the National Academy of
633 Sciences *96*, 8301-8306. 10.1073/pnas.96.14.8301.
634 Ghahremani, D.G., Lee, B., Robertson, C.L., Tabibnia, G., Morgan, A.T., De Shetler, N.,
635 Brown, A.K., Monterosso, J.R., Aron, A.R., and Mandelkern, M.A. (2012). Striatal dopamine

636 D2/D3 receptors mediate response inhibition and related activity in frontostriatal neural
637 circuitry in humans. Journal of Neuroscience *32*, 7316-7324. 10.1523/JNEUROSCI.4284-
638 11.2012.
639 Goldstein, M., Brendel, G., Tuescher, O., Pan, H., Epstein, J., Beutel, M., Yang, Y., Thomas,
640 K., Levy, K., and Silverman, M. (2007). Neural substrates of the interaction of emotional
641 stimulus processing and motor inhibitory control: an emotional linguistic go/no-go fMRI
642 study. Neuroimage *36*, 1026-1040. 10.1016/j.neuroimage.2007.01.056.
643 Guillery, R. (1995). Anatomical evidence concerning the role of the thalamus in
644 corticocortical communication: a brief review. Journal of anatomy *187*, 583.
645 Haber, S., and Mcfarland, N.R. (2001). The place of the thalamus in frontal cortical-basal
646 ganglia circuits. The Neuroscientist *7*, 315-324. 10.1177/107385840100700408.
647 Haber, S.N., and Calzavara, R. (2009). The cortico-basal ganglia integrative network: the role
648 of the thalamus. Brain research bulletin *78*, 69-74. 10.1016/j.brainresbull.2008.09.013.
649 Hampshire, A., Chamberlain, S.R., Monti, M.M., Duncan, J., and Owen, A.M. (2010). The
650 role of the right inferior frontal gyrus: inhibition and attentional control. Neuroimage *50*,
651 1313-1319. 10.1016/j.neuroimage.2009.12.109.
652 Hampton, W.H., Alm, K.H., Venkatraman, V., Nugiel, T., and Olson, I.R. (2017). Dissociable
653 frontostriatal white matter connectivity underlies reward and motor impulsivity. Neuroimage
654 *150*, 336-343. 10.1016/j.neuroimage.2017.02.021.
655 Hauber, W. (2010). Dopamine release in the prefrontal cortex and striatum: temporal and
656 behavioural aspects. Pharmacopsychiatry *43*, S32-S41. 10.1055/s-0030-1248300.

Hikosaka, O., Nakamura, K., and Nakahara, H. (2006). Basal ganglia orient eyes to reward. Journal of neurophysiology *95*, 567-584. 10.1152/jn.00458.2005.

Holmes, E., Zeidman, P., Friston, K.J., and Griffiths, T.D. (2021). Difficulties with speech-in-noise perception related to fundamental grouping processes in auditory cortex. Cerebral Cortex *31*, 1582-1596. 10.1093/cercor/bhaa311.

Huang, A.S., Mitchell, J.A., Haber, S.N., Alia-Klein, N., and Goldstein, R.Z. (2018). The thalamus in drug addiction: from rodents to humans. Philosophical Transactions of the Royal Society B: Biological Sciences *373*, 20170028. 10.1098/rstb.2017.0028.

Hung, Y., Gaillard, S.L., Yarmak, P., and Arsalidou, M. (2018). Dissociations of cognitive inhibition, response inhibition, and emotional interference: Voxelwise ALE meta-analyses of fMRI studies. Human brain mapping *39*, 4065-4082. 10.1002/hbm.24232.

Jahfari, S., Ridderinkhof, K.R., Collins, A.G., Knapen, T., Waldorp, L.J., and Frank, M.J. (2019). Cross-task contributions of frontobasal ganglia circuitry in response inhibition and conflict-induced slowing. Cerebral Cortex *29*, 1969-1983. 10.1093/cercor/bhy076.

Jahfari, S., Verbruggen, F., Frank, M.J., Waldorp, L.J., Colzato, L., Ridderinkhof, K.R., and Forstmann, B.U. (2012). How preparation changes the need for top–down control of the basal ganglia when inhibiting premature actions. Journal of Neuroscience *32*, 10870-10878. 10.1523/JNEUROSCI.0902-12.2012.

Jahfari, S., Waldorp, L., van den Wildenberg, W.P., Scholte, H.S., Ridderinkhof, K.R., and Forstmann, B.U. (2011). Effective connectivity reveals important roles for both the hyperdirect (fronto-subthalamic) and the indirect (fronto-striatal-pallidal) fronto-basal ganglia

678 pathways during response inhibition. Journal of Neuroscience *31*, 6891-6899.

679 10.1523/JNEUROSCI.5253-10.2011.

680 Kass, R.E., and Raftery, A.E. (1995). Bayes factors. Journal of the american statistical

681 association *90*, 773-795.

682 Kelly, A.C., Hester, R., Murphy, K., Javitt, D.C., Foxe, J.J., and Garavan, H. (2004).

683 Prefrontal-subcortical dissociations underlying inhibitory control revealed by event-related

684 fMRI. European Journal of Neuroscience *19*, 3105-3112. 10.1111/j.0953-816X.2004.03429.x.

685 Kim, H. (2014). Involvement of the dorsal and ventral attention networks in oddball stimulus

686 processing: A meta-analysis. Human brain mapping *35*, 2265-2284. 10.1002/hbm.22326.

687 Klugah-Brown, B., Di, X., Zweerings, J., Mathiak, K., Becker, B., and Biswal, B. (2020).

688 Common and separable neural alterations in substance use disorders: A coordinate-based

689 meta-analyses of functional neuroimaging studies in humans. Human brain mapping *41*,

690 4459-4477. 10.1002/hbm.25085.

691 Krämer, U.M., Solbakk, A.-K., Funderud, I., Løvstad, M., Endestad, T., and Knight, R.T.

692 (2013). The role of the lateral prefrontal cortex in inhibitory motor control. Cortex *49*, 837-

693 849. 10.1016/j.cortex.2012.05.003.

694 Li, C.-s.R., Huang, C., Constable, R.T., and Sinha, R. (2006). Gender differences in the neural

695 correlates of response inhibition during a stop signal task. Neuroimage *32*, 1918-1929.

696 10.1016/j.neuroimage.2006.05.017.

697 Li, J., Xu, L., Zheng, X., Fu, M., Zhou, F., Xu, X., Ma, X., Li, K., Kendrick, K.M., and

698 Becker, B. (2018). Common and dissociable contributions of alexithymia and autism to

699 domain-specific interoceptive dysregulations–a dimensional neuroimaging approach.
700 bioRxiv, 432971. 10.1159/000495122.

701 Li, J., Yang, X., Zhou, F., Liu, C., Wei, Z., Xin, F., Daumann, B., Daumann, J., Kendrick,
702 K.M., and Becker, B. (2020). Modafinil enhances cognitive, but not emotional conflict
703 processing via enhanced inferior frontal gyrus activation and its communication with the
704 dorsomedial prefrontal cortex. Neuropsychopharmacology *45*, 1026-1033. 10.1038/s41386-
705 020-0625-z.

706 Li, K., Jiang, Y., Gong, Y., Zhao, W., Zhao, Z., Liu, X., Kendrick, K.M., Zhu, C., and
707 Becker, B. (2019). Functional near-infrared spectroscopy-informed neurofeedback: regional-
708 specific modulation of lateral orbitofrontal activation and cognitive flexibility.
709 Neurophotonics *6*, 025011. 10.1117/1.NPh.6.2.025011.

710 Liu, X., Zhou, X., Zeng, Y., Li, J., Zhao, W., Xu, L., Zheng, X., Fu, M., Yao, S., and
711 Cannistraci, C.V. (2022). Medial prefrontal and occipito-temporal activity at encoding
712 determines enhanced recognition of threatening faces after 1.5 years. Brain Structure and
713 Function, 1-18. 10.1007/s00429-022-02462-5.

714 Maizey, L., Evans, C.J., Muhlert, N., Verbruggen, F., Chambers, C.D., and Allen, C.P.
715 (2020). Cortical and subcortical functional specificity associated with response inhibition.
716 NeuroImage *220*, 117110. 10.1016/j.neuroimage.2020.117110.

717 Mallet, N., Schmidt, R., Leventhal, D., Chen, F., Amer, N., Boraud, T., and Berke, J.D.
718 (2016). Arkypallidal cells send a stop signal to striatum. Neuron *89*, 308-316.
719 10.1016/j.neuron.2015.12.017.

720 Mamah, D., Wang, L., Barch, D., de Erausquin, G.A., Gado, M., and Csernansky, J.G.
721 (2007). Structural analysis of the basal ganglia in schizophrenia. Schizophrenia research *89*,
722 59-71. 10.1016/j.schres.2006.08.031.
723 Marzinzik, F., Wahl, M., Schneider, G.-H., Kupsch, A., Curio, G., and Klostermann, F.
724 (2008). The human thalamus is crucially involved in executive control operations. Journal of
725 Cognitive Neuroscience *20*, 1903-1914. 10.1162/jocn.2008.20124.
726 McFarland, N.R., and Haber, S.N. (2002). Thalamic relay nuclei of the basal ganglia form
727 both reciprocal and nonreciprocal cortical connections, linking multiple frontal cortical areas.
728 Journal of Neuroscience *22*, 8117-8132. 10.1523/JNEUROSCI.22-18-08117.2002.
729 Morein-Zamir, S., and Robbins, T.W. (2015). Fronto-striatal circuits in response-inhibition:
730 Relevance to addiction. Brain research *1628*, 117-129. 10.1016/j.brainres.2014.09.012.
731 Morein-Zamir, S., Dodds, C., van Hartevelt, T.J., Schwarzkopf, W., Sahakian, B., Müller, U.,
732 and Robbins, T. (2014). Hypoactivation in right inferior frontal cortex is specifically
733 associated with motor response inhibition in adult ADHD. Human brain mapping *35*, 5141-
734 5152. 10.1002/hbm.22539.
735 Obeso, J.A., Rodriguez-Oroz, M.C., Rodriguez, M., Lanciego, J.L., Artieda, J., Gonzalo, N.,
736 and Olanow, C.W. (2000). Pathophysiology of the basal ganglia in Parkinson's disease.
737 Trends in neurosciences *23*, S8-S19. 10.1016/s1471-1931(00)00028-8.
738 Ott, T., and Nieder, A. (2019). Dopamine and cognitive control in prefrontal cortex. Trends in
739 cognitive sciences *23*, 213-234. 10.1016/j.tics.2018.12.006.

740 Pfeifer, P., Sebastian, A., Buchholz, H.G., Kaller, C.P., Gründer, G., Fehr, C.,
741 Schreckenberger, M., and Tüscher, O. (2022). Prefrontal and striatal dopamine D2/D3
742 receptors correlate with fMRI BOLD activation during stopping. Brain imaging and behavior
743 *16*, 186-198. 10.1007/s11682-021-00491-y.
744 Phillips, J.M., Kambi, N.A., Redinbaugh, M.J., Mohanta, S., and Saalmann, Y.B. (2021).
745 Disentangling the influences of multiple thalamic nuclei on prefrontal cortex and cognitive
746 control. Neuroscience & Biobehavioral Reviews *128*, 487-510.
747 10.1016/j.neubiorev.2021.06.042.
748 Protopopescu, X., Pan, H., Altemus, M., Tuescher, O., Polanecsky, M., McEwen, B.,
749 Silbersweig, D., and Stern, E. (2005). Orbitofrontal cortex activity related to emotional
750 processing changes across the menstrual cycle. Proceedings of the National Academy of
751 Sciences *102*, 16060-16065. 10.1073/pnas.0502818102.
752 Puiu, A.A., Wudarczyk, O., Kohls, G., Bzdok, D., Herpertz-Dahlmann, B., and Konrad, K.
753 (2020). Meta-analytic evidence for a joint neural mechanism underlying response inhibition
754 and state anger. Human brain mapping *41*, 3147-3160. 10.1002/hbm.25004.
755 Qiao, L., Xu, M., Luo, X., Zhang, L., Li, H., and Chen, A. (2020). Flexible adjustment of the
756 effective connectivity between the fronto-parietal and visual regions supports cognitive
757 flexibility. NeuroImage *220*, 117158. 10.1016/j.neuroimage.2020.117158.
758 Rae, C.L., Nombela, C., Rodríguez, P.V., Ye, Z., Hughes, L.E., Jones, P.S., Ham, T.,
759 Rittman, T., Coyle-Gilchrist, I., and Regenthal, R. (2016). Atomoxetine restores the response
760 inhibition network in Parkinson's disease. Brain *139*, 2235-2248. 10.1093/brain/aww138.

761 Renteria, R., Baltz, E.T., and Gremel, C.M. (2018). Chronic alcohol exposure disrupts top-
762 down control over basal ganglia action selection to produce habits. Nature communications *9*,
763 1-11. 10.1038/s41467-017-02615-9.
764 Rey-Mermet, A., Gade, M., and Oberauer, K. (2018). Should we stop thinking about
765 inhibition? Searching for individual and age differences in inhibition ability. Journal of
766 Experimental Psychology: Learning, Memory, and Cognition *44*, 501. 10.1037/xlm0000450.
767 Ribeiro, F., Cavaglia, R., and Rato, J.R. (2021). Sex differences in response inhibition in
768 young children. Cognitive Development *58*, 101047. 10.1016/J.COGDEV.2021.101047.
769 Robbins, T.W. (2007). Shifting and stopping: fronto-striatal substrates, neurochemical
770 modulation and clinical implications. Philosophical Transactions of the Royal Society B:
771 Biological Sciences *362*, 917-932. 10.1098/rstb.2007.2097.
772 Robertson, C.L., Ishibashi, K., Mandelkern, M.A., Brown, A.K., Ghahremani, D.G., Sabb, F.,
773 Bilder, R., Cannon, T., Borg, J., and London, E.D. (2015). Striatal D1-and D2-type dopamine
774 receptors are linked to motor response inhibition in human subjects. Journal of Neuroscience
775 *35*, 5990-5997. 10.1523/JNEUROSCI.4850-14.2015.
776 Rubia, K., Smith, A.B., Taylor, E., and Brammer, M. (2007). Linear age-correlated functional
777 development of right inferior fronto-striato-cerebellar networks during response inhibition
778 and anterior cingulate during error-related processes. Human brain mapping *28*, 1163-1177.
779 10.1002/hbm.20347.
780 Rupprechter, S., Romaniuk, L., Series, P., Hirose, Y., Hawkins, E., Sandu, A.-L., Waiter,
781 G.D., McNeil, C.J., Shen, X., and Harris, M.A. (2020). Blunted medial prefrontal cortico-

limbic reward-related effective connectivity and depression. Brain *143*, 1946-1956. 10.1093/brain/awaa106.

Saalmann, Y.B., and Kastner, S. (2015). The cognitive thalamus. Frontiers in systems neuroscience *9*, 39. 10.3389/fnsys.2015.00039.

Schall, J.D., and Godlove, D.C. (2012). Current advances and pressing problems in studies of stopping. Current opinion in neurobiology *22*, 1012-1021. 10.1016/j.conb.2012.06.002.

Schaum, M., Pinzuti, E., Sebastian, A., Lieb, K., Fries, P., Mobascher, A., Jung, P., Wibral, M., and Tüscher, O. (2020). Cortical network mechanisms of response inhibition. biorxiv. 10.1101/2020.02.09.940841.

Schaum, M., Pinzuti, E., Sebastian, A., Lieb, K., Fries, P., Mobascher, A., Jung, P., Wibral, M., and Tüscher, O. (2021). Right inferior frontal gyrus implements motor inhibitory control via beta-band oscillations in humans. ELife *10*, e61679. 10.7554/eLife.61679.

Schimmack, U., and Derryberry, D.E. (2005). Attentional interference effects of emotional pictures: threat, negativity, or arousal? Emotion *5*, 55. 10.1037/1528-3542.5.1.55.

Schmitt, L.I., Wimmer, R.D., Nakajima, M., Happ, M., Mofakham, S., and Halassa, M.M. (2017). Thalamic amplification of cortical connectivity sustains attentional control. Nature *545*, 219-223. 10.1038/nature22073.

Sjoberg, E.A., and Cole, G.G. (2018). Sex differences on the Go/No-Go test of inhibition. Archives of Sexual Behavior *47*, 537-542. 10.1007/s10508-017-1010-9.

Sommer, M.A. (2003). The role of the thalamus in motor control. Current opinion in neurobiology *13*, 663-670. 10.1016/j.conb.2003.10.014.

803 Sonuga-Barke, E.J. (2005). Causal models of attention-deficit/hyperactivity disorder: from
804 common simple deficits to multiple developmental pathways. Biological psychiatry *57*, 1231-
805 1238. 10.1016/j.biopsych.2004.09.008.
806 Stephan, K.E., Harrison, L.M., Kiebel, S.J., David, O., Penny, W.D., and Friston, K.J. (2007).
807 Dynamic causal models of neural system dynamics: current state and future extensions.
808 Journal of biosciences *32*, 129-144. 10.1007/s12038-007-0012-5.
809 Stephan, K.E., Penny, W.D., Moran, R.J., den Ouden, H.E., Daunizeau, J., and Friston, K.J.
810 (2010). Ten simple rules for dynamic causal modeling. Neuroimage *49*, 3099-3109.
811 10.1016/j.neuroimage.2009.11.015.
812 Stuphorn, V. (2015). Neural mechanisms of response inhibition. Current opinion in
813 behavioral sciences *1*, 64-71. 10.1016/j.cobeha.2014.10.009.
814 Tanaka, M., and Kunimatsu, J. (2011). Contribution of the central thalamus to the generation
815 of volitional saccades. European Journal of Neuroscience *33*, 2046-2057. 10.1111/j.1460-
816 9568.2011.07699.x.
817 Terra, H., Bruinsma, B., de Kloet, S.F., van der Roest, M., Pattij, T., and Mansvelder, H.D.
818 (2020). Prefrontal cortical projection neurons targeting dorsomedial striatum control
819 behavioral inhibition. Current Biology *30*, 4188-4200. e4185. 10.1016/j.cub.2020.08.031.
820 Thompson, A., Schel, M.A., and Steinbeis, N. (2021). Changes in BOLD variability are
821 linked to the development of variable response inhibition. NeuroImage *228*, 117691.
822 10.1016/j.neuroimage.2020.117691.

823 Tops, M., and Boksem, M.A. (2011). A potential role of the inferior frontal gyrus and anterior
824 insula in cognitive control, brain rhythms, and event-related potentials. Frontiers in
825 psychology *2*, 330. 10.3389/fpsyg.2011.00330.
826 Van Overwalle, F., Van de Steen, F., van Dun, K., and Heleven, E. (2020). Connectivity
827 between the cerebrum and cerebellum during social and non-social sequencing using dynamic
828 causal modelling. NeuroImage *206*, 116326. 10.1016/j.neuroimage.2019.116326.
829 Verbruggen, F., and Logan, G.D. (2008). Response inhibition in the stop-signal paradigm.
830 Trends in cognitive sciences *12*, 418-424. 10.1016/j.tics.2008.07.005.
831 Verbruggen, F., and Logan, G.D. (2009). Models of response inhibition in the stop-signal and
832 stop-change paradigms. Neuroscience & Biobehavioral Reviews *33*, 647-661.
833 10.1016/j.neubiorev.2008.08.014.
834 Vijayraghavan, S., Major, A.J., and Everling, S. (2016). Dopamine D1 and D2 receptors make
835 dissociable contributions to dorsolateral prefrontal cortical regulation of rule-guided
836 oculomotor behavior. Cell reports *16*, 805-816. 10.1016/j.celrep.2016.06.031.
837 Wei, W., and Wang, X.-J. (2016). Inhibitory control in the cortico-basal ganglia-
838 thalamocortical loop: complex regulation and interplay with memory and decision processes.
839 Neuron *92*, 1093-1105. 10.1016/j.neuron.2016.10.031.
840 Weiss, F., Zhang, J., Aslan, A., Kirsch, P., and Gerchen, M.F. (2022). Feasibility of training
841 the dorsolateral prefrontal-striatal network by real-time fMRI neurofeedback. Scientific
842 reports *12*, 1-10. 10.1038/s41598-022-05675-0.

Xiao, D., Zikopoulos, B., and Barbas, H. (2009). Laminar and modular organization of prefrontal projections to multiple thalamic nuclei. Neuroscience *161*, 1067-1081. 10.1016/j.neuroscience.2009.04.034.

Xu, B., Sandrini, M., Wang, W.T., Smith, J.F., Sarlls, J.E., Awosika, O., Butman, J.A., Horwitz, B., and Cohen, L.G. (2016). PreSMA stimulation changes task-free functional connectivity in the fronto-basal-ganglia that correlates with response inhibition efficiency. Human brain mapping *37*, 3236-3249. 10.1002/hbm.23236.

Xu, L., Zheng, X., Yao, S., Li, J., Fu, M., Li, K., Zhao, W., Li, H., Becker, B., and Kendrick, K.M. (2022). The mirror neuron system compensates for amygdala dysfunction-associated social deficits in individuals with higher autistic traits. NeuroImage *251*, 119010. 10.1016/j.neuroimage.2022.119010.

Zeidman, P., Jafarian, A., Corbin, N., Seghier, M.L., Razi, A., Price, C.J., and Friston, K.J. (2019a). A guide to group effective connectivity analysis, part 1: First level analysis with DCM for fMRI. Neuroimage *200*, 174-190. 10.1016/j.neuroimage.2019.06.031.

Zeidman, P., Jafarian, A., Seghier, M.L., Litvak, V., Cagnan, H., Price, C.J., and Friston, K.J. (2019b). A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. Neuroimage *200*, 12-25. 10.1016/j.neuroimage.2019.06.032.

Zhang, R., Geng, X., and Lee, T. (2017). Large-scale functional neural network correlates of response inhibition: an fMRI meta-analysis. Brain Structure and Function *222*, 3973-3990. 10.1007/s00429-017-1443-x.

863 Zhao, Z., Yao, S., Li, K., Sindermann, C., Zhou, F., Zhao, W., Li, J., Lührs, M., Goebel, R.,
864 Kendrick, K.M., and Becker, B. (2019). Real-time functional connectivity-informed
865 neurofeedback of amygdala-frontal pathways reduces anxiety. Psychotherapy and
866 psychosomatics *88*, 5-15. 10.1159/000496057.
867 Zhou, F., Li, J., Zhao, W., Xu, L., Zheng, X., Fu, M., Yao, S., Kendrick, K.M., Wager, T.D.,
868 and Becker, B. (2020). Empathic pain evoked by sensory and emotional-communicative cues
869 share common and process-specific neural representations. Elife *9*, e56929.
870 10.7554/eLife.56929.
871 Zhou, F., Zimmermann, K., Xin, F., Scheele, D., Dau, W., Banger, M., Weber, B.,
872 Hurlemann, R., Kendrick, K.M., and Becker, B. (2018). Shifted balance of dorsal versus
873 ventral striatal communication with frontal reward and regulatory regions in cannabis-
874 dependent males. Human brain mapping *39*, 5062-5073. 10.1002/hbm.24345.
875 Zhuang, Q., Xu, L., Zhou, F., Yao, S., Zheng, X., Zhou, X., Li, J., Xu, X., Fu, M., Li, K.,
876 Vatansever, D., Kendrick, K.M., and Becker, B. (2021). Segregating domain-general from
877 emotional context-specific inhibitory control systems-ventral striatum and orbitofrontal cortex
878 serve as emotion-cognition integration hubs. NeuroImage *238*, 118269.
879 10.1016/j.neuroimage.2021.118269.
880 Zhukovsky, P., Morein-Zamir, S., Ziauddeen, H., Fernandez-Egea, E., Meng, C., Regenthal,
881 R., Sahakian, B.J., Bullmore, E.T., Robbins, T.W., and Dalley, J.W. (2021). Prefrontal Cortex
882 Activation and Stopping Performance Underlie the Beneficial Effects of Atomoxetine on

883 Response Inhibition in Healthy Volunteers and Those With Cocaine Use Disorder. Biological

884 Psychiatry: Cognitive Neuroscience and Neuroimaging. 10.1016/j.bpsc.2021.08.010.