Cerebral Cortex, 2018; 1–15

doi: 10.1093/cercor/bhy076
Original Article

ORIGINAL ARTICLE

Cross-Task Contributions of Frontobasal Ganglia Circuitry in Response Inhibition and Conflict-Induced Slowing

Sara Jahfari[1,2], K. Richard Ridderinkhof[2,3], Anne G.E. Collins[4], Tomas Knapen[1,5], Lourens J. Waldorp[2] and Michael J. Frank[6]

[1]Spinoza Centre for Neuroimaging, 1105 BK Amsterdam, The Netherlands, [2]Amsterdam Brain & Cognition (ABC), University of Amsterdam, 1018 WB Amsterdam, The Netherlands, [3]Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands, [4]Department of Psychology, University of California, Berkeley, 94720 CA, USA, [5]Department of Cognitive Psychology, Vrije Universiteit Amsterdam, 1081 BT Amsterdam, The Netherlands and [6]Department of Cognitive, Linguistic and Psychological Sciences, and Brown Institute for Brain Sciences, Brown University, Providence, 02912 Rhode Island, USA

Address correspondence to Sara Jahfari, Spinoza Centre for Neuroimaging, Royal Netherlands Academy of Arts and Sciences (KNAW) Meibergdreef 75 1105 BK Amsterdam, The Netherlands. Email: sara.jahfari@gmail.com; Michael J. Frank, Department of Cognitive, Linguistic and Psychological Sciences and Brown Institute for Brain Sciences, Brown University, 190 Thayler St., Providence RI 02912-1821, USA. Email: michael_frank@brown.edu

Abstract

Why are we so slow in choosing the lesser of 2 evils? We considered whether such slowing relates to uncertainty about the value of these options, which arises from the tendency to avoid them during learning, and whether such slowing relates to frontosubthalamic inhibitory control mechanisms. In total, 49 participants performed a reinforcement-learning task and a stop-signal task while fMRI was recorded. A reinforcement-learning model was used to quantify learning strategies. Individual differences in lose–lose slowing related to information uncertainty due to sampling, and independently, to less efficient response inhibition in the stop-signal task. Neuroimaging analysis revealed an analogous dissociation: subthalamic nucleus (STN) BOLD activity related to variability in stopping latencies, whereas weaker frontosubthalamic connectivity related to slowing and information sampling. Across tasks, fast inhibitors increased STN activity for successfully canceled responses in the stop task, but decreased activity for lose–lose choices. These data support the notion that fronto-STN communication implements a rapid but transient brake on response execution, and that slowing due to decision uncertainty could result from an inefficient release of this "hold your horses" mechanism.

Key words: basal ganglia systems, Bayesian hierarchical modeling, fMRI effective and functional connectivity, reinforcement learning, response inhibition

Optimal foraging entails learning to select among decision alternatives, based on their (hidden) probabilistic values. Individuals differ in their exploration/exploitation balance, and hence the degree to which they sample options with lower valued outcomes, during reinforcement learning. Such interindividual variability may help understand the mechanisms involved in choosing the lesser of 2 evils. Value-based decision-making often requires choice between options that have similar learned values but may never have been presented together (e.g., a novel choice between miso soup and corn chowder). These kinds of choices can elicit conflict arising

from either the novel pairing of 2 previously desired outcomes (win–win) or undesired outcomes (lose–lose). Despite identical value differences, the novel pairing of 2 lose–lose options is consistently associated with prolonged decision times when compared with win–win conflict (Frank, Samanta, et al. 2007; Cavanagh et al. 2011; Jocham et al. 2011; Ratcliff and Frank 2012; Cavanagh, Wiecki, et al. 2014). While the relative speeding for high valued options is attributed to effects of reward expectation (and dopamine levels) on reaction time (RT), the literature has generally not considered the impact of differential uncertainty about choice values. Consider a common reinforcement-learning task in which an agent learns to choose among pairs of options with different reinforcement probabilities (e.g., 80% vs. 20%, 70% vs. 30%, and 60% vs. 40%) (Frank et al. 2005). While one can optimize rewards in this task by exploiting/maximizing (i.e., always choosing the more rewarded option), this strategy would prevent the agent from exploration and hence from acquiring a precise representation about the value of the lesser options (Gittins 1979; Cohen et al. 2007). Critically, this exploitation strategy would also then make it more difficult to later choose between a 40% and 20% option (a high-conflict lose–lose choice), due to less sampling and greater uncertainty about their true values.

What are the neural mechanisms that can leverage such uncertainty to adjust decision times? Prior studies indicate that when presented with decision conflict, increased activity in the STN acts to delay response execution by inhibiting action altogether (Aron and Poldrack 2006; Aron et al. 2007; Jahfari et al. 2011, 2012) or by raising the decision threshold, that is, the level of evidence required to make a choice (Frank 2006; Frank, Samanta, et al. 2007; Cavanagh et al. 2011; Ratcliff and Frank 2012; Zaghloul et al. 2012; Green et al. 2013; Wiecki and Frank 2013; Frank et al. 2015; Zavala et al. 2015; Herz et al. 2016). Intuitively, a common mechanism for response inhibition and threshold adjustment seemingly implies that faster or more efficient inhibition would relate to more conflict-induced slowing. However, in the case of lose–lose conflict such a fixed increase in decision threshold mechanism is maladaptive when the learned information for the optimal choice is sparse (i.e., it could engender decision paralysis). Instead, simulation studies suggested that the STN "hold your horses" mechanism is dynamic, with a fast initial STN surge that is followed by a steep decline of activation ("releasing the horses"), facilitating choice even when the evidence is sparse (Ratcliff and Frank 2012; Wiecki and Frank 2013). This dynamic could even suggest an efficient initial STN surge, and hence rapid response inhibition, might actually lead to less uncertainty-induced slowing.

We aimed to specify these relationships with the examination of 2 tasks. Functional magnetic resonance imaging (fMRI) data was recorded while participants performed a reinforcement-learning task followed by a test-phase containing novel win–win, lose–lose, and win–lose pairs without feedback (Fig. 1*a*). Here, the degree of exploration/exploitation (and hence subsequent uncertainty in learned values of lose options) was assessed during learning by stochasticity in choices and quantified with a reinforcement-learning model that reliably predicted trial-to-trial choices (Fig. 2). With far less samples to refine beliefs about the precise probabilities, exploiters should only have a rough value estimate for loss stimuli. Although the reward probabilities of each stimulus are complementary (e.g., A is 80% and B is 20%) and hence in principle knowledge of this structure could facilitate inference about the value of B by selecting A, this inference is indirect and participants were never explicitly told about the complementary relationship within each learning option. Previous studies with this task have also shown that learning about A and B is independent and one can excel at choose-A and not at avoid-B and vice-versa. Choose-A performance is related to brain responses to positive feedback and striatal dopaminergic signaling, whereas avoid-B performance is related to neural response to negative feedback and oppositely impacted by dopaminergic manipulations (Frank et al. 2004; Frank, Moustafa, et al. 2007; Kravitz et al. 2012; Collins and Frank 2014; Frank 2016). Importantly, we also administered a stop-signal task to assess the efficiency of response inhibition in the absence of learning (Fig. 1*b*), and to relate this to the behavioral and neural markers of conflict-based slowing.

We assess how choice strategies and the efficiency of response inhibition each relate to slowing in 1) reaction times, 2) the BOLD response of the STN region, and 3) the strength of effective connectivity in the frontosubthalamic pathway by using a model-driven effective connectivity approach termed ancestral graphs (AG) (Waldorp et al. 2011). This last explorative analysis followed prior studies suggesting that the communication from PFC into the STN region, the so-called hyperdirect pathway (Nambu et al. 2002), is enhanced under response conflict (Aron et al. 2007; Isoda and Hikosaka 2007, 2008; Frank et al. 2015) to motivate a brake (Jahfari et al. 2012, 2015; Aron et al. 2016), or decision threshold adjustments on striatal reward-based choice in order to prevent impulsive or premature responses (Frank 2006; Cavanagh et al. 2011; Wiecki and Frank 2013; Herz et al. 2017).

Materials and Methods

Participants

A total of 49 young adults (25 males; mean age = 22 years; range: 19–29 years) participated in this study. Four participants were excluded from all analyses due to movement (2), incomplete sessions (1), or misunderstanding of task instructions (1). One participant did not complete the stop-task, and for one we were unable to obtain reliable SSRT estimates (stopping latency) therefore they were only included for the reinforcement learning task (RL-task) RL-task analysis. All participants had normal or corrected-to-normal vision and provided written consent before the scanning session, in accordance with the declaration of Helsinki. The ethics committee of the University of Amsterdam approved the experiment, and all procedures were in accordance with relevant laws and institutional guidelines.

Tasks and Procedure

As shown in Figure 1, participants performed a reinforcement-learning task (Frank et al. 2004) and a stop-signal task in the MRI scanner. All stimuli were presented on a black-projection screen that was viewed via a mirror-system attached to the MRI head coil. For each experiment faces with natural expressions were used as stimuli and selected from the Radboud Face Database (Langner et al. 2010). Faces had neither hair nor glasses and were trimmed to remove all external features (neck, hairline). To control for carryover effects, faces used for the stop-task were not used in the RL-task, and participants were told explicitly that the faces used in the stop-task are unrelated to, and different from, the ones used in the RL-task.

Reinforcement Learning Task

The RL-task consisted of 2 phases; an initial reinforcement learning phase and a subsequent test-phase. During the

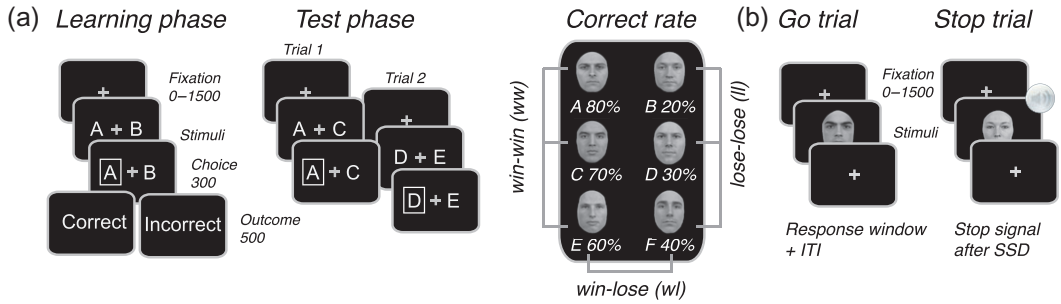

Figure 1. Experimental design. (*a*) Reinforcement-learning task. During learning, 2 faces were presented at each trial, and participants learned to select the most optimal face stimulus (A, C, E) solely through probabilistic feedback (probability of correct is displayed beneath each stimulus). The learning-phase only contained 3 face pairs (AB, CD, ED) for which feedback was given. In the test-phase, faces were arranged into 15 combinations. Trials were further identical to the learning-phase with the exception of feedback. (*b*) Stop-signal task. Each trial started with the presentation of a fixation-cross followed by a male or female face stimulus, indicating a left or right response. During stop trials, a tone was played at a variable delay (SSD) after the presentation of the go stimulus. The tone instructed participants to suppress the indicated response.

learning phase, 3 different male or female face pairs (AB, CD, EF) were presented in random order and participants learned to choose 1 of the 2 faces (Fig. 1*a*). Probabilistic feedback followed each choice to indicate "correct" (happy smiley) or "incorrect" (sad smiley) (Jocham et al. 2011). Choosing face-A lead to "correct" on 80% of the trials, whereas face-B leads to "incorrect". Other ratios for "correct" were 70:30 (CD) and 60:40 (EF). Each trial had a fixed duration of 4000 ms, and started with a jitter interval of 0, 500, 1000, or 1500 ms to obtain an interpolated temporal resolution of 500 ms. During this interval, a white fixation cross was presented and participants were asked to maintain fixation. Two faces were then presented left and right of the fixation-cross and remained on screen up to response, or trial end (4000 ms). If a response was given on time, a white box surrounding the chosen face was shown (300 ms) and followed (interval 0–450 ms) by feedback (500 ms). Omissions were followed by the text "miss" (2000 ms). The test-phase contained the 3 face-pairs from the learning phase, and 12 novel combinations, in which participants had to select which item they thought had been more rewarding during learning. High-conflict win–win trials were defined as choices that involved 2 previously rewarding stimuli (i.e., AC, AE, CE), whereas high-conflict lose–lose trials were defined as choices that involved 2 previously losing stimuli (BD, BF, DF). Low-conflict win–lose stimuli served as controls for selection among novel pairs but which invoked little conflict (AD, AF, CB, etc.). Test-phase trials (4000 ms) were identical to the learning phase but no feedback was provided. In addition to the jitter used at the beginning of each trial, null trials (4000 ms) were randomly interspersed during the learning (60 trials) and test (72 trials) phase. Across the whole task, each face was presented equally often on the left or right side, and choices were indicated with the right-hand index (left) or middle (right) finger.

Before the MRI session, participants performed a complete learning phase to familiarize with the task (300 trials with different faces). In the MRI scanner, participants performed 2 learning blocks of 150 trials each (300 trials total; equal numbers of AB, CD, and EF), and 3 test phase blocks of 120 trials each (360 total; 24 presentations of each pair).

Stop-signal task

Each trial started with a white fixation cross followed by a male or female face stimulus. Participants were asked to identify the gender of the face presented with a left (index finger right hand) or right response (middle finger of the right hand). During stop trials a tone was presented after a variable interval. The tone instructed participants to suppress the indicated gender response (Fig. 1*b*). Trials started with a random jitter interval of 0–1500 ms (steps of 500 ms), during which a white fixation cross was presented in the center of the screen. A face stimulus was then presented for a period of 500 ms. On 30% of the trials, the go stimulus was followed by a high tone (stop signal). The stop signal delay (SSD) between the go stimulus and the stop signal was initially set at 250 ms and adjusted according to standard staircase methods to ensure convergence to *P*(inhibit) = 0.5 (for the full description of the staircase method used please see the task code on https://github.com/sarajahfari/Control_Conflict.git). Instructions emphasized that participants should do their best to respond as quickly as possible while also doing their best to stop when an auditory stop signal occurred. Each trial had a fixed duration of 4000 ms, and trials were further separated by an occasional null trial with only a fixation (4000 ms; 15 trials). Outside the scanner, participants performed a brief practice block of 30 trials to familiarize with the task. In the MRI scanner, participants subsequently performed a total of 150 trials (100 go trials, 50 stop trials).

Reinforcement-Learning Model

We quantitatively characterized participants' learning curves using a variant of the Q learning RL algorithm (Watkins and Dayan 1992; Frank, Moustafa, et al. 2007; Daw 2011), using hierarchical Bayesian parameter estimation, allowing us to separately estimate learning rates from choice stochasticity/exploration. Based on previous work we defined separate learning rate parameters for positive (αgain) and negative (αloss) reward prediction errors (Frank, Moustafa, et al. 2007; Kahnt et al. 2009; Niv et al. 2012). Q-learning assumes participants represent reward expectations for each stimulus/action (A-to-F). After observing a particular reward outcome, the expected value (Q) for selecting a stimulus *i* (A-to-F) on the next trial is updated as follows:

$$Q_i(t+1) = Q_i(t) + \begin{cases} \alpha_{Gain}\,[r_i(t) - Q_i(t)], & \text{if } r = 1 \\ \alpha_{Loss}\,[r_i(t) - Q_i(t)], & \text{if } r = 0 \end{cases}$$

where $0 \le \alpha_{gain}$ or $\alpha_{loss} \le 1$ represent learning rates, *t* is trial number, and $r = 1$ (positive feedback) or $r = 0$ (negative feedback). The probability of selecting one response over the other (i.e., A over B) is computed as follows:

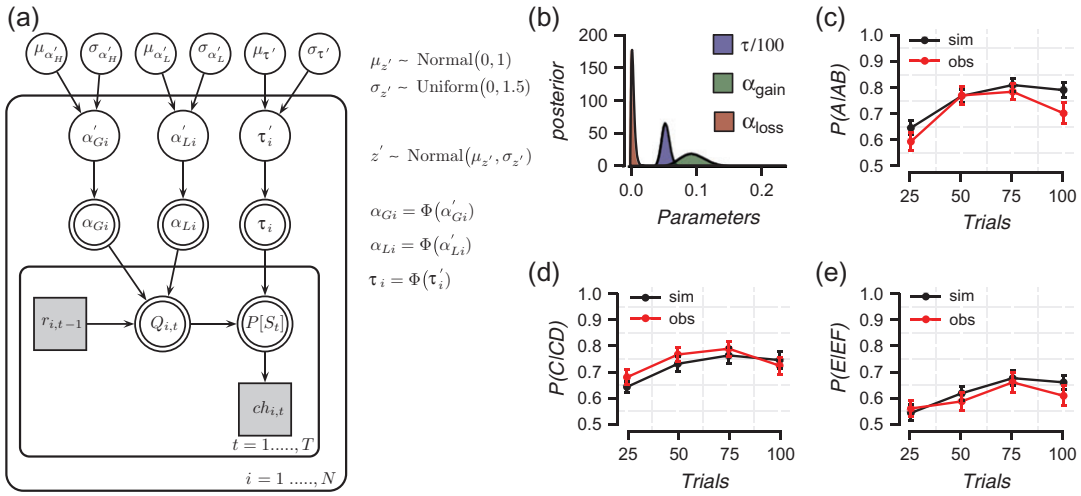

Figure 2. Q-learning model and performance. Graphical Q-learning model for hierarchical Bayesian parameter estimation (*a*). Φ () is the cumulative standard normal distribution function. The model consists of an outer subject (i = 1,…,N), and an inner trial plane (t = 1,…,T). Nodes represent variables of interest. Arrows are used to indicate dependencies between variables. Double borders indicate deterministic variables. Continues variables are denoted with circular nodes, and discrete with square nodes. Observed variables are shaded in grey. The right panel shows group-level posteriors for all Q-learning parameters (with τ/100) (*b*), and model performance where data is simulated with the estimated parameters and evaluated against the observed data for the AB (*c*), CD (*d*), or EF (*e*) pairs. Error bars represent standard error of the mean (SEM).

$$P_A(t) = \frac{\exp(\tau \times Q_t(A))}{\exp(\tau \times Q_t(B)) + \exp(\tau \times Q_t(A))}$$

with $0 \leq \tau \leq 100$ known as the inverse temperature governing the degree to which learned Q values are exploited. Higher estimates of τ indicate that decisions are mostly determined by the relative difference in value (exploitation), whereas lower estimates show a more stochastic choice pattern but which facilitates better learning of the underlying values of the lesser options.

The Q-learning algorithm was fit to the learning-phase trials using a Bayesian hierarchical estimation method where parameters for individual subjects are drawn from a group-level distribution. This hierarchical structure is preferred for parameter estimation as it allows for the simultaneous estimation of both group level parameters and individual parameters, and confers greater statistical strength for estimating and recovering parameters (Wetzels et al. 2010; Ahn et al. 2011; Lee 2011; Steingroever et al. 2013; Wiecki et al. 2013; Jahfari and Theeuwes 2017). Figure 2*a* shows a graphical representation of the model. The quantities $r_{i,\,t-1}$ (reward for participant i on trial t–1) and $ch_{i,t}$ (choice for participant *i* on trial *t*) are obtained directly from the data. The quantities αGi, αLi and τ*i* are deterministic, and are transformed during estimation by using their respective probit transformations Z′*i* (α′Gi, α′Li, τ′*i*). The probit transform is the inverse cumulative distribution function of the normal distribution. The parameters Z′*i* lie on the probit scale covering the entire real line. Parameters Z′*i* were drawn from group-level normal distributions with mean $\mu_{z'}$ and standard deviation $\delta_{z'}$. A normal prior was assigned to group-level means $\mu_{z'} \sim N(0,1)$, and a uniform prior to the group-level standard deviations $\delta_{z'} \sim U(1,1.5)$ (Wetzels et al. 2010; Steingroever et al. 2013). Model fits were implemented in Stan (Homan and Gelman 2014; Stan Development Team 2014). Multiple chains were generated to ensure convergence, and evaluated with the Rhat statistics (i.e., all Rhats were close to 1.0) (Gelman and Rubin 1992). The right panel of Figure 2 shows group-level posteriors on model parameters, and simulations from these parameters yield reasonable learning curves that match those observed empirically.

Behavioral Analysis

In the RL-task accuracy rates were based on choosing the most optimal stimulus (e.g., A over C, or A over B). Accuracy rates and median reaction times (RTs) for the RL-task test-phase were separated into high-conflict win–win (ww; AC, AE, CE), high-conflict lose–lose (ll; BD, BF, DF), and low-conflict win–lose (wl; AD, AF, CB, CF, EB, ED) pairs (Fig. 1*a*). Pairs that were presented during the learning phase (AB, CD, EF) were excluded from the win–lose condition, so that all conditions only contained novel pairs. Repeated measures ANOVA with Tukey's test were used to assess how conflict affects performance (RT and Accuracy). The stop-signal reaction time (SSRT) for the stop-task was estimated using the so-called "integration method" (Logan and Cowan 1984; Verbruggen and Logan 2009). This method takes the percentile of the go RT corresponding the individuals exact chance of responding given a stop signal, and subtracts the mean SSD from this value. Overall, the chance of responding given a signal was close to 0.5 (M = 0.47, SD = 0.06), the average SSD was 530.83 (SD = 284.71), and SSRT was 274.07 ms (SD = 105.67). Mean RTs of failed stops (M = 803.62, SD = 307.7) were faster than correct go trials (M = 845.14, SD = 330.20), and this difference was significant (t[42] = 2.30, P = 0.03) validating the independence assumption of the race-model. Robust regressions, and robust multiple regressions were used to focus on the relationship between conflict-induced slowing/errors and choice strategies (*τ*) or the efficiency to stop (SSRT), after reinforcement learning.

Magnetic Resonance Imaging Scanning Procedure

The fMRI data for the RL-task was acquired in a single scanning session with 2 learning and 3 test-phase runs on a 3-T scanner (Philips Achieva TX, Andover, MA) using a 32-channel head coil. Each scanning run contained 340 functional T2*-weighted echo-planar images for the learning phase, and 290 T2*-weighted echoplanar images for the test phase (TR = 2000 ms; TE = 27.63 ms; FA = 76.1°; 3 mm slice thickness; 0.3 mm slice spacing; FOV = 240 × 121.8 × 240; 80 × 80 matrix; 37 slices, ascending slice order). After a short break of 10 min with no scanning, data

collection was continued with a 3D T1 scan for registration purposes (repetition time [TR] = 8.5080 ms; echo time [TE] = 3.95 ms; flip angle [FA] = 8°; 1 mm slice thickness; 0 mm slice spacing; field of view [FOV] = 240 × 220 × 188), and the fMRI data collection for the stop-task (335 T2* weighted echoplanar images; TR = 2000 ms; TE = 27.63 ms; FA = 76.1°; 3 mm slice thickness; 0.3 mm slice spacing; FOV = 240 × 121.8 × 240; 80 × 80 matrix; 37 slices, ascending slice order).

Overall, participants first went into the scanner for approximately 45 min to perform the RL-task with 2 learning phases, and 3 test phase blocks. We subsequently introduced a short ±10 min break, where participants were taken out of the scanner and served with the traditional "stroopwafel" (Dutch cookie). Once participants went back into the scanner we first started with the recording of the structural T1 scan. During this period, participants had no task and were asked to relax. The stop-task then followed with an approximate duration of 15 min. Although we cannot rule out fatigue or attentional deterioration in the stop-task, our behavioral and BOLD observations were very similar to those observed in other fMRI studies using the stop-signal task (Aron et al. 2007; Jahfari et al. 2011).

Preprocessing

Preprocessing was performed using FEAT (FMRI Expert Analysis Tool) version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The first 6 volumes were discarded to allow for T1 equilibrium effects. Preprocessing steps included motion correction, high-pass filtering in the temporal domain (σ = 50), and prewhitening (Woolrich et al. 2001). All functional data sets were individually registered into 3D space using the participant's individual high-resolution anatomical images. The individual 3D representation was then used to normalize the functional data into Montreal Neurological Institute (MNI) space by linear and nonlinear scaling.

fMRI Analysis Procedure and ROI Selection

The analysis procedure of the fMRI data was 2-fold. First, an anatomically defined template of the STN region was used to explore how the STN BOLD response relates to SSRT (control) or τ (choice uncertainty based on past learning) in the test phase after learning, and in the stop-signal task. The STN template was derived from a recent study using ultrahigh 7 T scanning (Keuken et al. 2014), and selected for its use in previous fMRI studies focusing on reinforcement based conflicted choices (Frank et al. 2015), or response inhibition (Jahfari et al. 2012, 2015; De Hollander et al. 2017). Because our 3 T protocol with 3 mm isotropic voxels might lack the resolution to separate the STN from the Red Nucleus (RN), or Substantia Nigra (SN) with certainty (De Hollander et al. 2015) we refer to the STN as the STN region in the description and discussion of our findings.

Second, the model-driven AG connectivity method was used for selecting the optimal network in describing PFC and BG coactivation patterns during test-phase trials (for the analysis of the stop-task using AG please see Jahfari et al. 2011, 2012, 2015), and, to subsequently explore how the strength of PFC-BG connectivity in the optimal model relates to decision times, SSRT, or τ. The regions of interest (ROIs) used for the AG analysis included: 1) masks based on a whole brain cluster-corrected analysis of the learning-phase fMRI data to identify regions that covary with trial-by-trial signed reward prediction errors; and 2) a priori selected PFC and BG anatomical masks for regions typically associated with fronto-BG decision-making, or choice evaluations (Mink 1996; Frank 2006; Isoda and Hikosaka 2008; Nambu 2009; Shenhav et al. 2014). Please see below for a detailed description of each step.

Deconvolution Analysis of the STN

To more precisely examine the time course of activations in the STN region, we performed finite impulse response estimation (FIR) on the STN BOLD signals. After motion correction, temporal filtering and percent signal change conversion, data from the STN region were averaged across voxels, and upsampled from 0.5 to 3 Hz. This allows the FIR fitting procedure to capitalize on the random timings (relative to TR onset) of the stimulus presentations and decisions in the experiment. For this analysis, stimulus onset was chosen as t0 of the FIR time course. FIR time courses for all trial types were then estimated simultaneously using a least-squares fit, as implemented in the FIRDeconvolution package (Knapen and Gee 2016). Resulting single-participant response time-courses were then used to evaluate the contribution of SSRT and choice strategies for each timepoint separately, using multiple regression as implemented in the statsmodels package (Seabold and Perktold 2010). Here, alpha value for the contributions of SSRT and choice strategy was set to 0.0125 (i.e., a Bonferroni corrected value of 0.05 given the interval of interest between 0 and 8 s). Confidence intervals in Fig. 5 were estimated using bootstrap analysis across participants (n = 1000), where the shaded region represents the SEM across participants (i.e., bootstrapped 68% confidence interval).

Ancestral Graphs Method

To focus on frontobasal ganglia dynamics when participants make reinforcement guided-decisions after learning, the fMRI data recorded during test-phase was analyzed using AG (Waldorp et al. 2011). AG infer functional or effective connectivity by taking into account the distribution of BOLD activation per ROI, across trials, per subject, and so are not dependent on the low temporal resolution of the time series in fMRI. A graphical model reflects the joint distribution of several neuronal systems with the assumption that for each individual the set of active regions is the same. The joint distribution (graphical model) of 2 nodes is estimated from the replications of "condition specific trials" (e.g., win–win or lose–lose), and not from the time series. With this method, we can infer 3 types of connections: 1) effective connectivity (directed connection –>), 2) functional connectivity (undirected connection –), and 3) unobserved systems (bidirected connection <–>). Directed connections are regression parameters in the usual sense (denoted by β) and undirected connections are partial covariances (unscaled partial correlations; denoted by λ). The bidirected connections refer to the covariance of the residuals from the regressions (denoted by ω). These 3 types of connections can be identified by modeling the covariance matrix (denoted by Σ) as follows:

$$\sum = \beta^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & \Omega\end{pmatrix}(\beta^{-1})^T$$

where β contains the regression coefficients, Λ contains the partial covariances, and Ω contains the covariances between residuals. A random effects model is used to combine models across subjects to then compare different models over the whole group using Bayes information criterion (BIC). The graph with the lowest BIC value will be selected.

To infer directions from the ancestral graph, it is required that a change in direction implies a change in probability distribution. This is not always the case. For example, a chain from A to B to C is in terms of conditional independencies equivalent to a chain with the directions reversed, that is from C to B to A (for more details see Waldorp et al. 2011). Two equivalent models, such as those just mentioned, will result in the same BIC value, indicating that directionality cannot be inferred. The most important structure is when 2 arrowheads meet (a collider). This will always result in a change in BIC value. The causal interpretations of the connections from an ancestral graph that is the best model according to the BIC can be briefly described as follows:

- A –> B: A is a cause of B [effective connectivity]
- A – B: A is a cause of B and/or B is a cause of A [functional connectivity]
- A <–> B: there is a latent common cause of A and B [missing region]

For a more detailed description and cautions on causal interpretations see Zhang (2008).

The method of AG relies on conditional independencies implied by the topology of the network. Therefore, different models (e.g., different directions of connections) result in different fits to the data. The differences between models is characterized by BIC, which combines both accurate descriptive (for the data at hand) and predictive (for future data) value.

For the purpose of testing differences between connections, Waldorp et al. (2011) combined the estimation of AG with a random-effects model in which the parameters (connections) of each subject are from a normal distribution with unknown mean and variance. The main assumption of the random effects model is that all participants are from the same population, but that they can differ in connection strength. The model is compared at the group level to other models and is tested for fit at the individual level. The resulting ancestral graph is the best representation at the group level and at least an adequate representation at the individual level.

Once the model with the lowest BIC is selected, individual (subject) fits are obtained by using an adjusted goodness-of-fit test, indicating whether the model explains the data well enough. To assess relative fit between the selected model and saturated model, the AG method makes use of a modified version of the likelihood ratio (LR) test. For AG, the modified LR test is defined as the ratio of the model of interest (hypothesis) and the unrestricted (saturated) model. The test is corrected for being overly sensitive because the data can deviate from normality slightly (Yuan and Bentler 1997). The corrected test, has asymptotically a 2 distribution with $p(p + 1)/2 - q$ degrees of freedom, where p is the number of variables and q is the number of parameters. The test represents the relative difference in fit between the saturated model and the hypothesized model. Smaller values indicate a good relative fit to the observed data, compared with the full-saturated model; that is, smaller values mean that leaving out connections still corresponds well to the data. A significance level of 0.05 was used to reject models with a poor fit at the subject level.

The main differences between AG described in (Waldorp et al. 2011) and dynamic causal modeling (DCM) or structural equation modeling (SEM) are: 1) inference is based on trial-by-trial variation in the estimated BOLD signal and not on the time series as in DCM or SEM because of the low frequency sampling in fMRI, 2) both functional and effective connectivity can be represented in a single ancestral graph, 3) a common unobserved (latent) cause of a connection can be detected, 4) the definition of a circular system is only possible in undirected systems, and 5) the selected model is always compared against the full saturated model to evaluate relative fits, and to ensure that leaving out connections can still correspond well to the observed pattern of BOLD responses across trials.

ROI Definition for AG Connectivity

AG connectivity was evaluated using the test-phase trials of the RL-task. Because lose–lose options can prolong decisions, the first aim of our connectivity evaluation was to describe how PFC and BG decision-making regions collaborate to reach, and implement, a value-driven choice. We previously showed that a stop-network (please see Supplementary Fig. 1a), with projections from the right IFG and preSMA into the hyperdirect (STN –> GPi), and indirect (STR –> GPe –> GPi) pathways fits well to observed BOLD activity patterns when a choice is omitted, but is insufficient to describe across trial coactivation patterns when a manual choice is initiated (Jahfari et al. 2012, 2015). For similar observations with the BOLD pattern of win–win or lose–lose trials, using this right hemispheric stop-network, please see Supplementary Figure 1b. To optimize fits for the description of win–win or lose–lose decisions, all ROIs were selected based on their potential involvement in value-driven decision-making. The definition of ROIs used for AG connectivity in the test-phase relied on 1) the analysis of fMRI data in the learning-phase, to identify regions (voxels) within the striatum and vmPFC that correlate specifically with the signed reward prediction error (RPE), and 2) on previous work linking specific regions within the PFC and BG to value-driven choice.

First, the 2 learning blocks were used to identify voxels within the vmPFC and striatum that respond to ongoing reward prediction errors during reinforcement-guided decision-making. For this purpose, the onset of each outcome was modeled as a separate delta function and convolved with the hemodynamic response function. We used a parametric GLM design with orthogonalized regressors where positive or negative outcomes were parametrically modulated by demeaned trial-wise prediction errors derived from the Q-learning model. Individual contrast images were computed for positive and negative error related responses and taken to a second-level random effect analysis using one-sample t-test. For the whole-brain analysis Z (Gaussianzied T/F) statistic images were thresholded using clusters determined by $z > 2.3$ (contrast positive RPE correlation) and a cluster-corrected significance threshold of $P = 0.05$. Note, that this liberal threshold was only used for the definition of ROI masks that covary with RPE during learning; to be used only as masks for the evaluation of connectivity in the subsequent test-phase. During the test-phase feedback is no longer presented and internal representations of action values become vital to the selection process. Because reward prediction errors are thought to act as a teaching signal, ROI definition for the striatum (center of gravity [cog]: 1, 5, –4) and vmPFC (cog: –3, 52, –1) nodes made use of the positive correlation RPE contrast in the learning phase, with the exclusion of voxels in the ventricles.

Second, a priori anatomical masks were defined for the following regions: preSMA (cog: [–] 9, 25, 50), DLPFC (cog: [–] 37, 37, 27), STN (cog: [L] –9, –14, –7; [R] 10, –13, –7), globus pallidus interna (GPi) (cog: [L] –18, –8, –3; [R] 19, –7, –3), globus pallidus

externa (GPe) (cog: [L] −19, −5, 0; [R] 20, −3, 0), thalamus (cog: [L] −10, −19, 7; [R] 11, −18, 7), and primary motor cortex (M1) (cog: −18, −26, 61). All selected ROIs were bilateral. The DLPFC template was obtained from a recent study, linking especially the posterior part to action execution (Cieslik et al. 2013). The STN, GPe, and GPi templates were derived from a previous study using ultrahigh 7 T scanning (Keuken et al. 2014), thresholded to exclude the lowest 25% voxels, and then binarized. All other ROIs were created from cortical and subcortical structural atlases available in FSL.

Single-Trial Parameter Extraction for AG Connectivity

For each ROI (anatomical or RPE based) we subsequently obtained a single parameter estimate (averaged normalized β estimate across voxels in each ROI mask) for each trial of the recorded test-phase, per subject. The average number of parameters (based on trials) per ROI was 71.1 (SD = 1.7) for win–win, 71.4 (SD = 1.3) for lose–lose, and 213.7 (SD = 5.2) for win–lose. Misses were excluded from connectivity analysis. Connectivity analysis was conducted in R-Cran (version 3.0.2), including the packages ggm (version 1.995-3), graph (version 1.40.0), and RBGL (version 1.38.0). Please see Supplementary Figures 2 and 3 for the average (across trials) cluster corrected activity maps of each behavior, overlaid with the outline of masks used for connectivity. Note however that these activity maps display only the average response across trials, whereas our connectivity analysis focused on trial-by-trial coactivation patterns.

Model Definition for AG Connectivity

To examine how frontal and basal-ganglia nodes work together in selecting a response during the test phase, model fits were performed on the following trials: 1) win–win, 2) lose–lose, and 3) win–lose choices. A set of 7 potential choice models containing the direct (PFC–Striatum–GPi–Thalamus–M1), hyperdirect (PFC–STN region–GPi–Thalamus–M1), or indirect (PFC–Striatum–GPe–GPi–Thalamus–M1) PFC-BG pathways was tested to find the most optimal model in explaining the pattern of activation in the predefined regions. PFC consisted of vmPFC, DLPFC and preSMA, and each PFC region was defined to project into BG (see above for specification in the separate pathways). Because all PFC regions projected into BG, connections between PFC nodes could only be defined as undirected (functional connectivity). To optimize fits, all models were evaluated separately for left (right hand index finger) and right (right hand middle finger) responses (Table 1), and win–lose trials were first subdivided into 3 smaller chunks based on value-differences between pairs (small, 30; medium, 40; large, 50).

Connectivity towards GPe and GPi was differentially evaluated based on the theoretical description of the direct (Striatum –> GPi), indirect (Striatum –> GPe –> GPi), and hyperdirect (STN region –> GPi) pathways. The differential evaluation of 2 adjacent nodes (i.e., GPi and GPe) is not trivial and was justified by 2 critical observations. First, our model selection approach illustrated a better balance between variance and bias, with substantial decreases in BIC values (> 860 points in each condition; please see Table 1: model 4 vs. model 7), when projections towards the GPe were defined ("indirect" pathway) alongside the projections towards only GPi ("hyperdirect" and "direct" pathways). Second, for each participant, across trial partial correlations (pcor) were computed among all the regions included in the connectivity analysis. Across subjects, the strength of pcor observed between the adjacent GPe/GPi nodes (pcor M = 0.37, SD = 0.11), was highly similar to relationships found between more distant regions such as for example the DLPFC and preSMA (pcor M = 0.36, SD = 0.19; t[44] = 0.30, P = 0.77), or the preSMA and vmPFC (pcor M = 0.38, SD = 0.15; t[44]= −0.39, P = 0.70).

Because connection strengths did not differ for win–lose divisions, parameter estimates of the winning model were averaged for the win–win, lose–lose, and win–lose condition to align with the behavioral analysis. To compare the contribution of each model with the BIC criterion all 9 regions were always entered into the model, but the defined relationship (or connections) among regions varied across models.

Results

Uncertainty and Conflict-Induced Slowing

In our RL-task participants learned to select among choices with different probabilities of reinforcement (i.e., AB 80:20, CD 70:30, and EF 60:40). A subsequent test-phase, where feedback was omitted, required participants to select the optimal option among novel pairs involving low (win–lose) or high (win–win and lose–lose) decision conflict. During the test-phase, as expected, accuracy in choosing the most optimal stimulus was reduced for both high-conflict lose–lose and win–win pairs compared with low-conflict win–lose pairs (F[2,88] = 18.8, P < 0.0001; Fig. 3*a*). Slowed RTs were observed for only the lose–lose pairs (F[2,88] = 21.4, P < 0.0001; Fig. 3*b*).

To understand how the experience of conflict is influenced by the uncertainty associated with learned values that arises from information sampling, we quantified such sampling via the softmax τ parameter estimated from the reinforcement learning model. Higher estimates of τ index a greater tendency to exploit higher valued stimuli and as such predicted higher accuracies (r_{AB} = 0.79, r_{CD} = 0.74, r_{EF} = 0.47; all P's < 0.01; Fig. 3*c*) with steeper learning curves (Fig. 3*d–f*) in the learning phase. As

Table 1 BIC values for model fits across test-phase conditions

		Win–win		Lose-lose		Win–lose l		Win–lose m		Win–lose s	
	Model specification	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
1	PFC + BG direct	18 045	18 222	18 333	18 453	12 928	13 321	13 281	13 306	13 192	13 226
2	PFC + BG indirect	17 491	17 507	17 731	17 753	12 535	12 900	12 862	12 905	12 743	12 748
3	PFC + BG hyperdirect	18 709	18 644	18 987	19 036	13 367	13 678	13 674	13 653	13 636	13 612
4	PFC + BG direct&hyperdirect	17 831	17 906	18 133	18 320	12 768	13 119	13 056	13 082	13 031	12 986
5	PFC + BG direct&indirect	17 427	17 453	17 681	17 705	12 480	12 832	12 801	12 855	12 667	12 673
6	PFC + BG indirect&hyperdirect	16 660	16 614	16 902	16 918	11 962	12 296	12 235	12 279	12 104	12 116
7	PFC + BG direct&indirect&hyperdirect	16 588	16 564	16 848	16 869	11 906	12 215	12 176	12 218	12 018	12 029

Lower BIC values indicate a better balance between the variance and bias of the estimated model connections.

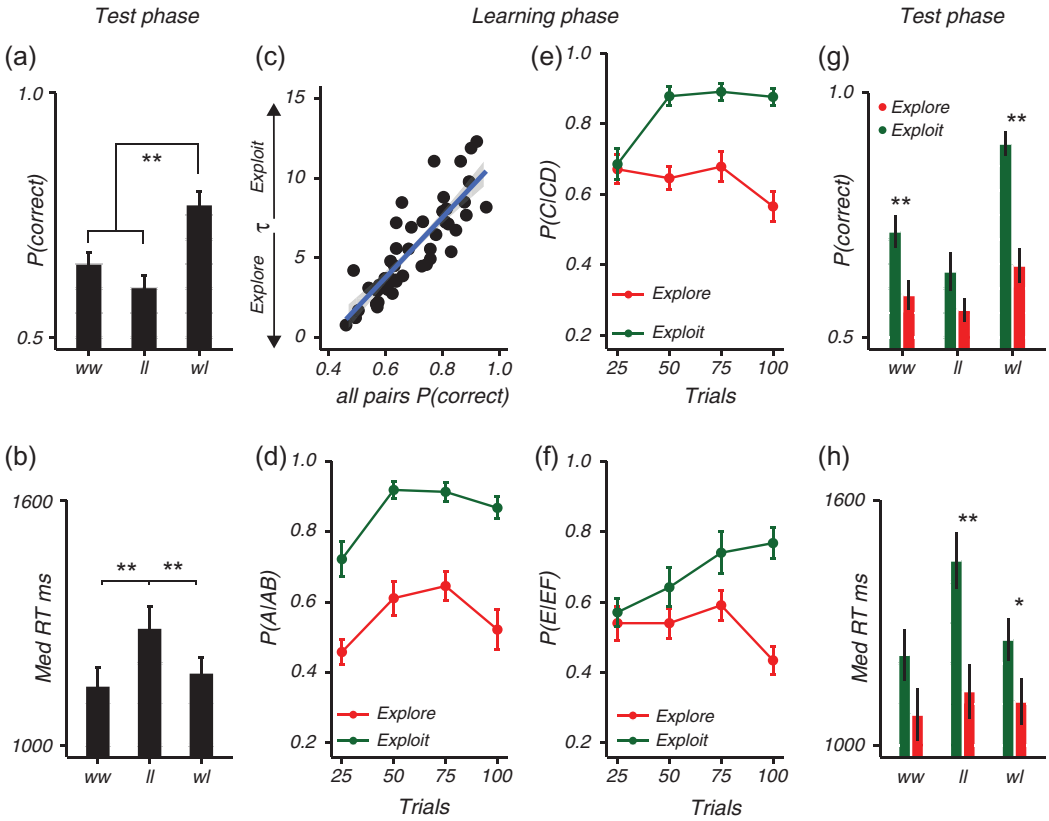

Figure 3. The exploit/explore trade-off in future decisions. Percentage of correct responses (*a*) and median reaction times (*b*) in the test-phase. (*c*) Participants with an exploitative choice strategy learned well by mostly choosing the optimal options during the learning phase (*d–f*) and were more accurate (*g*) but slowed (*h*) in the test-phase; especially for the mostly neglected lose–lose pairs. The groups in plots *d–h* were created with a median split on β, and plotted to illustrate learning differences over time. Error bars represent SEM. $^{**}P < 0.01$, $^{*}P < 0.05$

noted above, however, we posited that such exploitation would increase the uncertainty about the values of under-sampled loss stimuli in future test-phase choices. A repeated measures ANOVA with the between subject variable strategy (Exploit/Explore; defined as the continuous variable τ) and within subject factor Conflict (win–win, lose–lose, win–lose) revealed that exploitation during learning was related to improved accuracy ($F[1,43] = 72.1$, $P < 0.0001$; please see Fig. 3*g* for a visualization based on a median split on τ), but also prolonged reaction times in the test phase ($F[1,43] = 9.0$, $P < 0.01$; Fig. 3*h*). Critically, these effects were qualified by an interaction between Strategy and Conflict (accuracy: $F[2,86] = 3.4$, $P = 0.04$; RT: $F[2,86] = 6.9$, $P < 0.01$), revealing especially large costs for lose–lose decisions in exploiters. In particular, compared with explorers, exploiters exhibited the most prominent RT cost for lose–lose ($t[42] = 3.5$, $P = 0.001$) and less so for win–lose ($t[42] = 2.1$, $P = 0.04$), and win–win ($t[42] = 1.7$, $P = 0.09$). Similarly, although they performed more accurately overall, exploiters showed significant gains in accuracy only for choices involving a win stimulus (win–win $t[42] = 3.2$, $P < 0.01$; win–lose $t[42] = 6.3$, $P < 0.0001$), and not for lose–lose choices ($t[42] = 1.8$, $P = 0.08$).

Hence, while it is unsurprising that overall, participants performing more accurately during training also do so at test, these exploitative participants were characterized by relatively selective RT costs for the lose–lose choices in the test phase. These costs are expected given that they had not sampled these stimuli as much and hence should exhibit larger uncertainty when choosing among them. We next considered whether such RT costs were mitigated by response inhibition, separately from choice strategy.

Control and Conflict-Induced Slowing

While prolonged decision-times (RTs) in the test-phase were related to exploitative choice strategies, we also were interested to assess the role of response inhibition independently of learning and uncertainty. Previous work has attributed conflict-induced slowing to the same STN mechanism associated with outright response inhibition (Frank 2006; Aron et al. 2007) via either dynamic modulation of decision thresholds and/or an initial delay that precedes the decision-process (Ratcliff and Frank 2012). Therefore, we additionally examined an independent measure of inhibitory control efficiency in the stop-signal task termed the stop-signal reaction time (SSRT, Fig. 4*a*).

We hypothesized that if conflict-induced slowing is simply associated with more overall response inhibition (or a fixed increase in decision threshold), then subjects engaging this mechanism would exhibit more inhibition and slower conflict-induced RTs. If, on the other hand, conflict-induced slowing involves a transient threshold increase that then collapses, then efficient response inhibition should relate to less conflict-induced slowing. Moreover, for exploiters, this release of a transient brake should particularly censor the tail of the RT distribution, which would otherwise have more density due to uncertainty in the evidence. Please note that $N = 43$, in contrast to the analysis above with $N = 45$, as there were 2 more fall-outs in the stop task.

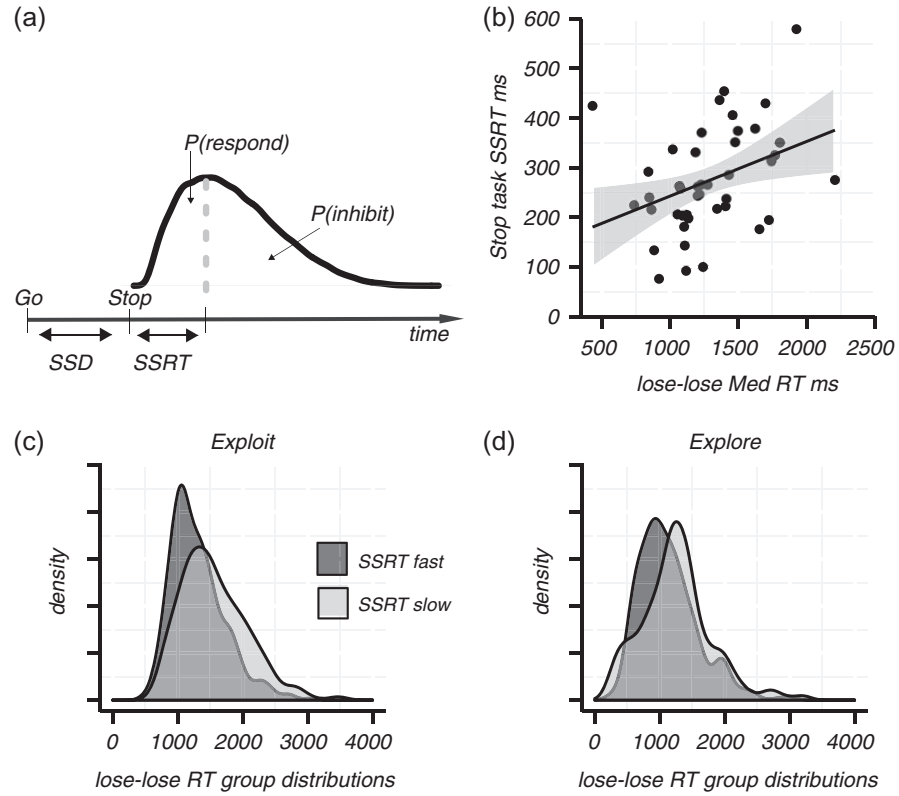

Figure 4. The efficiency to implement control predicts lose–lose slowing. (*a*) Graphic representation of the race model estimation for SSRT. A distribution of go trial RTs is shown beneath the curve. SSRT represents the average time needed to suppress a planned response. The efficiency to stop (SSRT) predicted response times during lose–lose trials (*b*). Exploitative participants who were more efficient in inhibition showed a steeper decline in the tail of the lose–lose reaction time distribution (*c*), this was not seen for explorative participants (*d*). Median splits were used to create the Exploit/Explore or fast/slow SSRT groups.

Indeed, overall, faster SSRTs (more efficient inhibition) were related to faster lose–lose response times ($t[41] = 3.36$, $P = 0.002$, robust regression; Fig. 4*b*). (No such relationship was seen for win–win RT; $t[41] = 1.57$, $P = 0.13$.) SSRT was unrelated to exploration vs exploitation in choices during learning ($t[41] = 1.82$, $P = 0.08$), suggesting that the 2 factors might contribute independent variance to the lose–lose decision times. Indeed, a robust multiple regression showed a significant contribution of both τ ($b_{\tau} = 42.02$, $t[40] = 2.992$, $P = 0.005$), and SSRT ($b_{ssrt} = 1.181$, $t[40] = 2.989$, $P = 0.005$) to lose–lose response times. Furthermore, while the inhibition effect was observable in both exploiters and explorers—consistent with an independent effect of SSRT on implementing and releasing the brake—its impact on the tail of the distributions was observed only in exploiters (Fig. 4*c*,*d*). This result is consistent with the notion that exploiters have more uncertainty about action outcomes, and hence without an efficient brake they exhibit longer tails. SSRTs were not related to lose–lose accuracy performance ($P = 0.34$).

These results explain lose–lose RT as a function of both choice strategies (previous sampling of information and hence uncertainty) and active but transient inhibitory control. Highly slowed participants were exploitative during learning, and inefficient in the implementation of a fast brake.

The Efficacy of Control in the STN During Full Stops and Conflict

At the neural level, the STN is well known for its role in global stopping (fast full brake) and the modulation of decision requirements. To evaluate how our behavioral observations relate to this literature, the time-course of activity within the STN region was estimated for both the stop-signal task and the test-phase of the RL-task. Multiple regressions were then used to evaluate how the STN region activity in each task relates to one's efficacy to inhibit a planned response (SSRT), or choice strategy τ.

In the stop-signal task ($N = 43$), the estimated STN region activity (Fig. 5*a*) was strongest for failed stop trials, corroborating a recent 7 T study focusing on the STN in this task (De Hollander et al. 2017), and possibly reflects a reactive engagement to correct for the failure to stop (we return to this result in the discussion). Notably, efficient inhibition, as indexed by SSRT, was marginally correlated with the estimated STN region response only when participants succeeded to suppress a planned response on time (i.e., successful stop trials); such that higher early BOLD responses in the STN region were related to faster or more efficient inhibition times (Fig. 5*b*,*c*). As expected, no relationship was observed between the STN region BOLD response and τ, when participants were engaged in the stop-task.

These 3 T observations focusing on the STN region in the stop-signal task replicate the 7 T findings reported by of De Hollander et al. (2017), by increasing both sample size and voxel size. The gains in temporal signal to noise ratio (tSNR) and power resulted in the estimation of robust task-related STN BOLD responses with a peak around 4–6 s, and condition specific replications of the 7 T report. Our large sample size additionally enabled us to examine how variability in BOLD relates to individual differences in stopping efficiencies. We extend current beliefs by showing how the efficacy to inhibit responses is only related to the STN regions BOLD response when the attempt is successful.

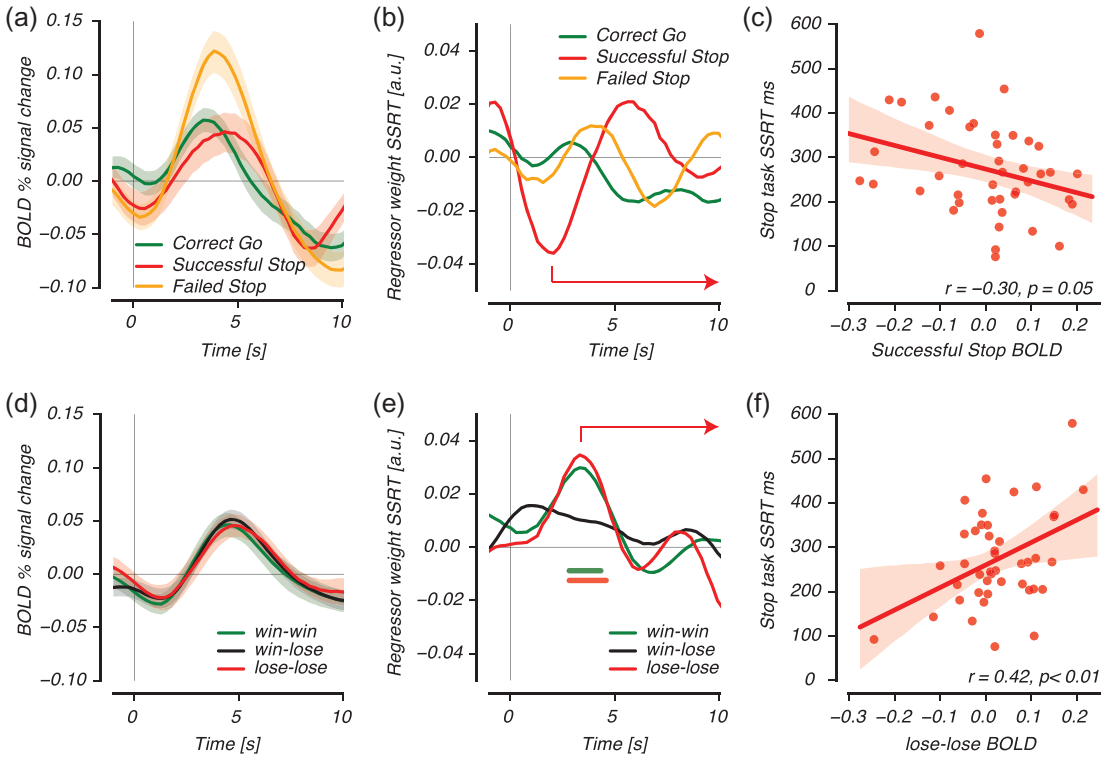

Figure 5. The efficiency to stop predicts the STN BOLD response in conflict and full response cancellation. The FIR-estimated STN BOLD signal time-course for trials in the stop-signal task (top panel) and the test-phase after learning (bottom panel), with the estimated regression coefficients SSRT shown for each trial type (mid panel). SSRT differentially related to activity patterns of the STN when inhibition was successful in the stop task (marginal effect with $P = 0.05$), or with the experience of conflict in test-phase trials. The horizontal lines show the interval in which SSRT contributed significantly to the multiple regression, for the conflicted lose–lose (red) and win–win (green) trials. The right panel highlights the differential relationship across task, with drawn correlation plots for successful stop trials (*c*), and the slowed lose–lose trials after learning (*f*).

We then turned to the test-phase of the RL-task to understand how STN region activity relates to the observed behavioral relationships between slowing and control ($N = 43$). Overall, the estimated STN region response was very similar across all trials of the test phase (Fig. 5*d*). The multiple regression, however, showed that STN response to high-conflict (win–win and lose–lose), but not low-conflict win–lose trials, was directly related to SSRT (Fig. 5*e*). STN region activity was unrelated to τ, pointing towards distinct effects of inhibitory control and choice uncertainty on response slowing. The positive relationship between SSRT and lose–lose STN activity (Fig. 5*f*) corresponds to the behavioral observation that more slowing was tied to longer SSRTs, consistent with the notion that it results from the inefficient implementation of a transient STN region brake.

Conflict-Induced Slowing in Corticobasal Ganglia Pathways

Finally, we used the model-driven AG approach (please see Materials and Methods for a detailed explanation) to analyze the information flow between PFC and BG during test-phase trials, and to explore how this interplay relates to the significant lose–lose slowing.

The evaluation of effective connectivity restricted the interplay between PFC and BG with the use of 3 pathways generally described in animal and human studies (Nambu et al. 2002; Nambu 2009; Jahanshahi and Rothwell 2017) (Fig. 6*a*). Here, most projections terminate in the striatum (STR), from where 2 (out of 3) pathways depart. A direct-pathway projects into thalamus via the globus pallidus interna (GPi) to facilitate action selection, while an indirect-pathway via the globus pallidus externa (GPe) can allow the integration of additional information by adaptively slowing the motor output. A third, hyperdirect-pathway directly projects from PFC into subthalamic nucleus (STN) and inhibits the thalamus output to primary motor cortex (M1) by exciting the GPi. These described PFC-BG pathways each play a specific (and therefore testable) role in the selection, regulation, or suppression of choices, and were therefore selected for the evaluation of 7 potential connectivity networks to describe information flows, or connectivity, between the PFC and BG during test-phase decisions (please see Materials and Methods for the definition of all 7 models).

Concurrent with the literature, random effects analysis across the whole group indicated that a connectivity network comprising the direct, indirect and the hyperdirect pathway best describes the pattern of activity during all choice trials (Table 1). Assessment of relative fit then ensured that the model is a good representation of the observed activation patterns in all 45 participants. Figure 6*b* shows the graphical outline of this model with functional connectivity (undirected relationship) between all PFC nodes (i.e., DLPFC, preSMA, vmPFC) and effective connectivity from each PFC region into the striatum (STR) and STN region. Within BG, effective connectivity was defined from STN region into GPi, Striatum into GPe, GPe into GPi, GPi into thalamus, and finally Thalamus into primary motor cortex (M1) to select a response. To better understand the top-down dynamics of this network we next focused on connection strengths (regression values derived from AG) from PFC into STN region or STR in 2 steps.

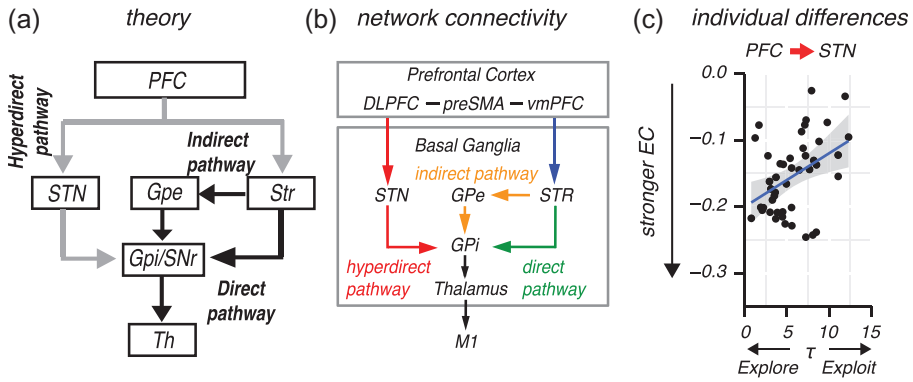

Figure 6. Schematic of the theoretical frontobasal ganglia pathways and effective connectivity results. (*a*) Theoretical frontobasal ganglia model with the direct, indirect and hyperdirect pathways. Gray arrows represent excitatory connections; black arrows represent inhibitory connections. (*b*) Graphical representation of the most representative effective connectivity network for all test-phase trials. Directed arrows represent effective connectivity (EC); undirected lines represent functional connectivity. For lose–lose trials, weaker PFC-into-STN connections related to a more exploitative choice strategy in the past and more uncertainty about the lose–lose options (*c*).

First, the strength of top-down connections was investigated using a repeated measure ANOVA with the factors Conflict (win–win, lose–lose, win–lose) and Connection (PFC –> STN region, PFC –> STR). There was a main effect of connection: PFC effective connectivity towards the STR (mean = −0.17, SD = 0.007) was stronger compared with that toward STN region (mean = −0.15, SD = 0.007; $F[1,44] = 10.38$, $P = 0.002$). There were no additional main effects or interactions modulated by conflict.

Second, we explored how past choice strategies, SSRTs, and test-phase decision times (RTs) each relate to the strength of connectivity (estimated regression strengths) from PFC into either the STN region or STR. The relationships are evaluated with $N = 45$ for RT and τ, or with $N = 43$ for SSRT, and reported with Bonferoni corrected P-values for the 3 behaviors evaluated using a critical alpha of 0.0167 (i.e., a Bonferroni corrected value of 0.05 given 3 tests). For lose–lose trials, PFC-into-STN connectivity correlated significantly with past choice strategies τ ($r = 0.41$, P_bonferoni=0.015; Fig. 6*c*); such that participants with the most exploitative strategies, and hence most uncertain about values of lose–lose options, exhibited the weakest PFC-into-STN region communication. No significant relationships were observed in the evaluation of PFC-into-STN region connectivity against SSRTs ($r = 0.21$, P_bonferoni=0.53), or lose–lose RTs ($r = 0.30$, P_bonferoni=0.14). Moreover, no relationships were observed in the evaluation of PFC-into-STN region connectivity during win–win or win–lose trials, or for PFC-into-STR connectivity (P's > 0.05).

These evaluations suggest that the communication between the PFC and STN region (the so called hyperdirect pathway) is disrupted during lose–lose trials because of choice, or information, uncertainty. The lack of a relationship between PFC-into-STN connectivity and stop-task performance is consistent with our previous work focusing on the stop-signal task (Jahfari et al. 2011, 2015), and possibly relates to the fast signal conduction within the hyperdirect pathway.

Discussion

A large body of work focuses on the neural mechanisms of reinforcement learning and value-based decision making, and how animals and humans can optimize learning and choice performance in stochastic environments. However, here we provide evidence for a tradeoff: subjects that appear to perform better during learning are less able to quickly avoid the worst of low value options in a later generalization test. Because exploitative subjects during learning did not sample the less valuable options, they obtained less information about their precise probabilities. The concomitant increased choice uncertainty for later decisions and was marked by altered communication strengths from the PFC into the STN region and slower response times. We additionally focused on the mechanisms of this lose–lose slowing and the related neural response in the STN region to show how both relate to the estimated stop signal reaction times (SSRTs), or the ability to rapidly and transiently implement inhibitory control.

The frontosubthalamic connections are thought to support conflict-induced slowing, and allow the integration of additional information by slowing or fully suppressing the motor output (Frank 2006; Cavanagh et al. 2011; Ratcliff and Frank 2012; Wiecki and Frank 2013; Herz et al. 2016). With the use of a model-driven connectivity approach, we found that the dynamic coupling between PFC and the STN region was weakest for the most uncertain and slowed participants. This relationship was further clarified by 3 additional findings. First, we observed that the magnitude of lose–lose slowing is best explained by considering not only past choice strategies (and hence uncertainty), but also, independently, SSRT. The SSRT is an estimate for one's efficiency to implement control (Logan and Cowan 1984; Verbruggen and Logan 2008) and extensively related to the STN region, which provides a fast and transient brake on all responses (Aron and Poldrack 2006; van den Wildenberg et al. 2006; Frank, Samanta, et al. 2007; Li et al. 2008; Schmidt et al. 2013; Obeso et al. 2014; Jahanshahi et al. 2015; Aron et al. 2007, 2016; Benis et al. 2016; Mallet et al. 2016; Fife et al. 2017). Accordingly, those subjects who were least efficient at rapid response inhibition (long SSRTs) exhibited more lose–lose slowing and a stronger STN surge during high-conflict trials. Moreover, while no direct relationship was observed between choice strategies and SSRT, the fast (and transient) implementation of control was especially helpful in the prevention of overly slow lose–lose choices, especially for uncertain exploitative learners.

In the last decade 2 parallel lines of literature have focused on the specific role of the STN in the modulation of evidence requirements/decision threshold adjustments (Bogacz 2007; Cavanagh et al. 2011; Ratcliff and Frank 2012; Herz et al. 2016),

or full response suppression (Aron and Poldrack 2006; Swann et al. 2011; Obeso et al. 2014; Jahanshahi et al. 2015). Response conflict has been consistently associated with the adaptation of evidence requirements (Verbruggen and Logan 2009; Jahfari et al. 2012; Wiecki and Frank 2013; White et al. 2014), including win–win and lose–lose conflict after reinforcement learning (Simen et al. 2006; Cavanagh et al. 2011; Cavanagh, Masters, et al. 2014; Cavanagh, Wiecki, et al. 2014). In this study, we evaluated the efficacy of control against the STN BOLD response during the experience of conflict, after learning, and in a separate task during full response suppression. We observed that the efficacy to implement a fast and full brake on all responses (SSRT) is differentially related to the STN region in each process.

In the stop-signal task, the activity pattern of the STN region was only related to stopping times when inhibition was successful. Here, the rise of the STN BOLD was highest for fast inhibitors. We note that this effect was only marginal in the 43 participants evaluated but consistent with the literature describing the STN in the stop task with rodents, or humans (Aron and Poldrack 2006; Aron et al. 2007; Schmidt et al. 2013; Schmidt and Berke 2017). The strongest BOLD response in the STN region was observed for trials where participants failed to inhibit a response on time (Li et al. 2008; De Hollander et al. 2017). Here, participants fail to inhibit the growth of activity for the go decision on time, and as a result might compensate by activating the STN without restraint or any regulation (Salinas and Stanford 2013; Greenhouse et al. 2015). This compensation effort could increase estimates of the slow BOLD response, but as observed, should have no causal contribution to the stop process for which the average inhibition time is estimated with SSRT.

In contrast to successful stop trials, slower inhibition times were associated with increased STN BOLD responses in the evaluation of value-based decisions. Critically, however, this relationship was specific to the high-conflict win–win and lose–lose trials, and not observed for easy win–lose decisions. Supporting the contrast between stopping and conflict, recent recordings from the STN have shown power increases in the STN to differ in the frequency range for conflict (2–8 Hz range) (Cavanagh et al. 2011; Zavala et al. 2014), or response inhibition (13–30 Hz) (Swann et al. 2009; Bastin et al. 2014; Aron et al. 2016; Wessel et al. 2016). However, beta-band adaptations (15–35 Hz) also occur at the resolve of conflict (Brittain et al. 2012). Our results suggest that a fast but transient STN brake, as posited by models showing a collapse in STN activity, might be helpful during conflicted-choices.

The efficiency to suppress all responses correlated with the STN region BOLD response during lose–lose and win–win conflict. Behaviorally, however, responses were only slowed and related to uncertainty, or SSRTs during lose–lose trials. Possibly, with the presentation of 2 negative options, the lack of information, negative value, and conflict all conspire to delay the selection process, or decision (Ratcliff and Frank 2012; Cavanagh and Frank 2013). In contrast, when conflict is the result of 2 positive options (win–win) there is more information, and the STN activates to counterbalance only the most impulsive choices with the increase of evidence requirements (Frank et al. 2005; Frank, Samanta, et al. 2007; Cavanagh et al. 2011). The lack of response slowing for win–win choices can largely be attributed to the impact of predicted reward on the decision process itself. Indeed, when the normal counterbalancing function of the STN is disrupted win–win choices become even faster than the easy win–lose (Frank, Samanta, et al. 2007; Ratcliff and Frank 2012).

The strength of fronto-STN connectivity or the magnitude of the STN BOLD both did not differ when compared between low-conflict (win–lose), or high-conflict (win–win, lose–lose) trials. Nevertheless, we observed selective relationships between uncertainty and fronto-STN connectivity during lose–lose decisions, or a relationship between control and the STN BOLD only at times of high conflict. These data imply frontosubthalamic involvements and activity to be condition specific—despite any differences in magnitude. In the literature, condition specific relationships between PFC-theta and RT are found for learned high-conflict choices—an effect that is reversed by deep brain stimulation of the STN—whereas no difference is found in overall theta power across conditions (Cavanagh et al. 2011). Moreover, the oscillatory activity of STN is related to behavior in opposing directions for low or high-conflict conditions (Herz et al. 2016). Our results complement these observations with the analysis of BOLD to show how high-conflict responses can be specifically tied to disrupted fronto-STN region dynamics, or inefficient control mechanisms.

Finally, previous time-sensitive reports have shown that the coherence between medial PFC and STN is increased in early periods of high-conflict (Zavala et al. 2014; Frank et al. 2015), with slower and more accurate responses when increases in the STN follow adaptations in medial PFC (Isoda and Hikosaka 2008). At first glance, our connectivity results contradict these findings. The critical difference here is that we evaluate connectivity, or coactivation patterns, with the use of trial-by-trial estimates of the slow BOLD response in PFC and BG nodes. In the STN region, these BOLD estimates include both the rise (implementation), and fall (release) (Ratcliff and Frank 2012) of the brake that is implemented to allow more time in conflicted decisions. With an identical rise, the trial-by-trial estimates of the STN region BOLD should be lower with faster releases of the brake. We found stronger negative PFC-into-STN region connections for the most certain participants, who responded faster, and chose the lesser options more often during learning. This pattern may suggest that when activity levels across PFC are raised sufficiently by information, the STN brake is released to allow choice. Consistently, PFC-into-STN region connectivity was disrupted, and tied to slowing, for the most uncertain participants who mostly avoided the lesser options during learning. Future work should refine this interpretation with high temporal resolution approaches to evaluate connectivity in both early and late phases of conflict-based decisions (Cohen 2011).

To summarize, these results describe the profound lose–lose slowing as a function of past learning choices, and individual differences in active but transient response suppression through the STN region (i.e., “hold and release your horses”). Moreover, they provide novel insights into the frontosubthalamic (“hyperdirect”) pathway involvement during the regulation of value-driven conflict.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online

Authors’ Contributions

S.J., K.R.R. and M.J.F. designed the study. S.J. collected the data. S.J., A.C., T.K., and L.W. contributed novel methods. S.J. and T.K. analyzed the data. S.J. and M.J.F. wrote the first draft of the article.

Funding

This project was funded by an ABC Talent Grant from the University of Amsterdam to S.J. and a National Science

Foundation (Grant #1460604) to M.J.F. The code and processed files supporting the findings can be downloaded from: https://github.com/sarajahfari/Control_Conflict.git. The raw data is available from the corresponding author in BIDS format upon reasonable request (sara.jahfari@gmail.com). To get an intuition for fitting the AG method using the stop-signal data reported in the supplementary see: https://github.com/sarajahfari/AG_example.git.

Notes

Conflict of Interest: None declared.

References

Ahn W-Y, Krawitz A, Kim W, Busemeyer JR, Brown JW. 2011. A model-based fMRI analysis with hierarchical Bayesian parameter estimation. J Neurosci Psychol Econ. 4:95–110.

Aron AR, Behrens TE, Smith SM, Frank MJ, Poldrack RA. 2007. Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. J Neurosci. 27:3743–3752.

Aron A, Herz D, Brown P, Forstmann B, Zaghloul K. 2016. Fronto-subthalamic circuits for control of action and cognition. J Neurosci. 36:11485–11495.

Aron AR, Poldrack RA. 2006. Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. J Neurosci. 26:2424–2433.

Bastin J, Polosan M, Benis D, Goetz L, Bhattacharjee M, Piallat B, Krainik A, Bougerol T, Chabardès S, David O. 2014. Inhibitory control and error monitoring by human subthalamic neurons. Transl Psychiatry. 4:e439.

Benis D, David O, Piallat B, Kibleur A, Goetz L, Bhattacharjee M, Fraix V, Seigneuret E, Krack P, Chabardès S, et al. 2016. Response inhibition rapidly increases single-neuron responses in the subthalamic nucleus of patients with Parkinson's disease. Cortex. 84:111–123.

Bogacz R. 2007. Optimal decision-making theories: linking neurobiology with behaviour. Trends Cogn Sci. 11:118–125.

Brittain J-S, Watkins KE, Joundi RA, Ray NJ, Holland P, Green AL, Aziz TZ, Jenkinson N. 2012. A role for the subthalamic nucleus in response inhibition during conflict. J Neurosci. 32:13396–13401.

Cavanagh JF, Frank MJ. 2013. Stop! Stay tuned for more information. Exp Neurol. 247:289–291.

Cavanagh JF, Masters SE, Bath K, Frank MJ. 2014. Conflict acts as an implicit cost in reinforcement learning. Nat Commun. 5:5394.

Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, Frank MJ. 2011. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat Neurosci. 14:1462–1467.

Cavanagh JF, Wiecki TV, Kochar A, Frank MJ. 2014. Eye tracking and pupillometry are indicators of dissociable latent decision processes. J Exp Psychol Gen. 143:1476–1488.

Cieslik EC, Zilles K, Caspers S, Roski C, Kellermann TS, Jakobs O, Langner R, Laird AR, Fox PT, Eickhoff SB. 2013. Is There "One" DLPFC in cognitive action control? Evidence for heterogeneity from co-activation-based parcellation. Cereb Cortex. 23:2677–2689.

Cohen MX. 2011. It's about time. Front Hum Neurosci. 5:2.

Cohen JD, McClure SM, Yu AJ. 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philos Trans R Soc Lond B Biol Sci. 362:933–942.

Collins AGE, Frank MJ. 2014. Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol Rev. 121:337–366.

Daw ND. 2011. Trial-by-trial data analysis using computational models. In: Delgado MR, Phelps EA, Robbins TW, editors. Decision making, affect, and learning: Attention and performance XXIII. Oxford: Oxford University Press. pp. 3–38.

De Hollander G, Keuken MC, Forstmann BU. 2015. The subcortical cocktail problem; mixed signals from the subthalamic nucleus and substantia nigra. PLoS One. 10:e0120572.

De Hollander G, Keuken MC, van der Zwaag W, Forstmann BU, Trampel R. 2017. Comparing functional MRI protocols for small, iron-rich basal ganglia nuclei such as the subthalamic nucleus at 7 and 3 T. Hum Brain Mapp. 38:3226–3248.

Fife KH, Gutierrez-Reed NA, Zell V, Bailly J, Lewis CM, Aron AR, Hnasko TS. 2017. Causal role for the subthalamic nucleus in interrupting behavior. eLife. 6. doi:10.7554/eLife.27689.

Frank MJ. 2006. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw. 19:1120–1136.

Frank MJ. 2016. Computational cognitive neuroscience approaches to deconstructing mental function and dysfunction. Comput Psychiatry New Perspect Ment Illn. 20:101–120.

Frank MJ, Gagne C, Nyhus E, Masters S, Wiecki TV, Cavanagh JF, Badre D. 2015. fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. J Neurosci. 35:485–494.

Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE. 2007. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci USA. 104:16311–16316.

Frank MJ, Samanta J, Moustafa AA, Sherman SJ. 2007. Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. Science. 318:1309–1312.

Frank MJ, Seeberger LC, O'Reilly RC. 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science. 306:1940–1943.

Frank MJ, Woroch BS, Curran T. 2005. Error-related negativity predicts reinforcement learning and conflict biases. Neuron. 47:495–501.

Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. Stat Sci. 7:457–472.

Gittins J. 1979. Bandit processes and dynamic allocation indices. J R Stat Soc Ser B. 41:148–177.

Green N, Bogacz R, Huebl J, Beyer AK, Kühn AA, Heekeren HR. 2013. Reduction of influence of task difficulty on perceptual decision making by STN deep brain stimulation. Curr Biol. 23:1681–1684.

Greenhouse I, Sias A, Labruna L, Ivry RB. 2015. Nonspecific inhibition of the motor system during response preparation. J Neurosci. 35:10675–10684.

Herz DM, Tan H, Brittain J-S, Fischer P, Cheeran B, Green AL, FitzGerald J, Aziz TZ, Ashkan K, Little S, et al. 2017. Distinct mechanisms mediate speed-accuracy adjustments in cortico-subthalamic networks. eLife. 6:357–381.

Herz DMM, Zavala BAA, Bogacz R, Brown P. 2016. Neural correlates of decision thresholds in the human subthalamic nucleus. Curr Biol. 26:916–920.

Homan MD, Gelman A. 2014. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J Mach Learn Res. 15:1593–1623.

Isoda M, Hikosaka O. 2007. Switching from automatic to controlled action by monkey medial frontal cortex. Nat Neurosci. 10: 240–248.

Isoda M, Hikosaka O. 2008. Role for subthalamic nucleus neurons in switching from automatic to controlled eye movement. J Neurosci. 28:7209–7218.

Jahanshahi M, Obeso I, Rothwell JC, Obeso JA. 2015. A fronto-striato-subthalamic-pallidal network for goal-directed and habitual inhibition. Nat Rev Neurosci. 16:719–732.

Jahanshahi M, Rothwell JC. 2017. Inhibitory dysfunction contributes to some of the motor and non-motor symptoms of movement disorders and psychiatric disorders. Phil Trans R Soc B. 372. doi:10.1098/rstb.2016.0198.

Jahfari S, Theeuwes J. 2017. Sensitivity to value-driven attention is predicted by how we learn from value. Psychon Bull Rev. 24:408–415.

Jahfari S, Verbruggen F, Frank MJ, Waldorp L, Colzato L, Ridderinkhof KR, Forstmann BU. 2012. How preparation changes the need for top-down control of the basal ganglia when inhibiting premature actions. J Neurosci. 32:10870–10878.

Jahfari S, Waldorp L, Ridderinkhof KR, Scholte HS. 2015. Visual information shapes the dynamics of corticobasal ganglia pathways during response selection and inhibition. J Cogn Neurosci. 27:1344–1359.

Jahfari S, Waldorp L, van den Wildenberg WP, Scholte HS, Ridderinkhof KR, Forstmann BU. 2011. Effective connectivity reveals important roles for both the hyperdirect (fronto-subthalamic) and the indirect (fronto-striatal-pallidal) fronto-basal ganglia pathways during response inhibition. J Neurosci. 31: 6891–6899.

Jocham G, Klein T a., Ullsperger M. 2011. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. J Neurosci. 31:1606–1613.

Kahnt T, Park SQ, Cohen MX, Beck A, Heinz A, Wrase J. 2009. Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. J Cogn Neurosci. 21:1332–1345.

Keuken MC, Bazin P-L, Crown L, Hootsmans J, Laufer A, Müller-Axt C, Sier R, van der Putten EJ, Schäfer A, Turner R, et al. 2014. Quantifying inter-individual anatomical variability in the subcortex using 7 T structural MRI. NeuroImage. 94: 40–46.

Knapen T, Gee JW De. 2016. FIRDeconvolution.

Kravitz AV, Tye LD, Kreitzer AC. 2012. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. Nat Neurosci. 15:816–818.

Langner O, Dotsch R, Bijlstra G, Wigboldus DH, Hawk ST, Knippenberg A. 2010. Presentation and validation of the Radboud Faces Database. Cogn Emot. 24:1377–1388.

Lee MD. 2011. How cognitive modeling can benefit from hierarchical Bayesian models. J Math Psychol. 55:1–7.

Li CR, Yan P, Sinha R, Lee TW. 2008. Subcortical processes of motor response inhibition during a stop signal task. NeuroImage. 41:1352–1363.

Logan GD, Cowan WB. 1984. On the ability to inhibit thought and action: a theory of an act of control. Psychol Rev. 91: 295–327.

Mallet N, Schmidt R, Leventhal D, Chen F, Amer N, Boraud T, Berke JD. 2016. Arkypallidal cells send a stop signal to striatum. Neuron. 89:308–316.

Mink JW. 1996. The basal ganglia: focused selection and inhibition of competing motor programs. Prog Neurobiol. 50: 381–425.

Nambu A. 2009. Functions of direct, indirect and hyperdirect pathways. Brain Nerve. 61:360–372.

Nambu A, Tokuno H, Takada M. 2002. Functional significance of the cortico-subthalamo-pallidal "hyperdirect" pathway. Neurosci Res. 43:111–117.

Niv Y, Edlund JA, Dayan P, O'Doherty JP. 2012. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. J Neurosci. 32:551–562.

Obeso I, Wilkinson L, Casabona E, Speekenbrink M, Luisa Bringas M, Alvarez M, Alvarez L, Pavón N, Rodríguez-Oroz MC, Macías R, et al. 2014. The subthalamic nucleus and inhibitory control: impact of subthalamotomy in Parkinson's disease. Brain. 137: 1470–1480.

Ratcliff R, Frank MJ. 2012. Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. Neural Comput. 24: 1186–1229.

Salinas E, Stanford TR. 2013. The countermanding task revisited: fast stimulus detection is a key determinant of psychophysical performance. J Neurosci. 33:5668–5685.

Schmidt R, Berke JD. 2017. A Pause-then-Cancel model of stopping: evidence from basal ganglia neurophysiology. Philos Trans R Soc Lond B Biol Sci. 372. doi:10.1098/rstb.2016.0202.

Schmidt R, Leventhal DK, Mallet N, Chen F, Berke JD. 2013. Canceling actions involves a race between basal ganglia pathways. Nat Neurosci. 16:1118–1124.

Seabold S, Perktold J 2010. Statsmodels: econometric and statistical modeling with Python. In: Proceedings of the 9th Python in Science Conference. p. 57–61.

Shenhav A, Straccia MA, Cohen JD, Botvinick MM. 2014. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nat Neurosci. 17: 1249–1254.

Simen P, Cohen JD, Holmes P. 2006. Rapid decision threshold modulation by reward rate in a neural network. Neural Netw. 19:1013–1026.

Stan Development Team. 2014. RStan: the R interface to Stan, Version 2.5.0.

Steingroever H, Wetzels R, Wagenmakers E-J. 2013. Validating the PVL-Delta model for the Iowa gambling task. Front Psychol. 4:898.

Swann NC, Poizner H, Houser M, Gould S, Greenhouse I, Cai W, Strunk J, George J, Aron AR. 2011. Deep brain stimulation of the subthalamic nucleus alters the cortical profile of response inhibition in the beta frequency band: a scalp EEG study in Parkinson's disease. J Neurosci. 31:5721–5729.

Swann NC, Tandon N, Canolty R, Ellmore TM, Mcevoy LK, Dreyer S, Disano M, Aron AR. 2009. Intracranial EEG reveals a time- and frequency-specific role for the right inferior frontal gyrus and primary motor cortex in stopping initiated responses. J Neurosci. 29:12675–12685.

van den Wildenberg WP, van Boxtel GJ, Van Der Molen MW, Bosch DA, Speelman JD, Brunia CH. 2006. Stimulation of the subthalamic region facilitates the selection and inhibition of motor responses in Parkinson's disease. J Cogn Neurosci. 18: 626–636.

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

Verbruggen F, Logan GD. 2009. Models of response inhibition in the stop-signal and stop-change paradigms. Neurosci Biobehav Rev. 33:647–661.

Waldorp L, Christoffels I, van de Ven V. 2011. Effective connectivity of fMRI data using ancestral graph theory: dealing with missing regions. NeuroImage. 54:2695–2705.

Watkins CJCH, Dayan P. 1992. Q-learning. Mach Learn. 8: 279–292.

Wessel JR, Ghahremani A, Udupa K, Saha U, Kalia SK, Hodaie M, Lozano AM, Aron AR, Chen R. 2016. Stop-related subthalamic beta activity indexes global motor suppression in Parkinson's disease. Mov Disord. 31:1846–1853.

Wetzels R, Vandekerckhove J, Tuerlinckx F, Wagenmakers E-J. 2010. Bayesian parameter estimation in the Expectancy Valence model of the Iowa gambling task. J Math Psychol. 54:14–27.

White CN, Congdon E, Mumford JA, Karlsgodt KH, Sabb FW, Freimer NB, London ED, Cannon TD, Bilder RM, Poldrack RA. 2014. Decomposing decision components in the stop-signal task: a model-based approach to individual differences in inhibitory control. J Cogn Neurosci. 26:1601–1614.

Wiecki TV, Frank MJ. 2013. A computational model of inhibitory control in frontal cortex and basal ganglia. Psychol Rev. 120:329–355.

Wiecki TV, Sofer I, Frank MJ. 2013. HDDM: hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. Front Neuroinform. 7:14.

Woolrich MW, Ripley BD, Brady M, Smith SM. 2001. Temporal autocorrelation in univariate linear modeling of fMRI data. NeuroImage. 14:1370–1386.

Yuan KH, Bentler P. 1997. Mean and covariance structure analysis: theoretical and practical improvements. J Am Stat Assoc. 92:767–774.

Zaghloul KA, Weidemann CT, Lega BC, Jaggi JL, Baltuch GH, Kahana MJ. 2012. Neuronal activity in the human subthalamic nucleus encodes decision conflict during action selection. J Neurosci. 32:2453–2460.

Zavala BA, Tan H, Little S, Ashkan K, Hariz M, Foltynie T, Zrinzo L, Zaghloul KA, Brown P. 2014. Midline frontal cortex low-frequency activity drives subthalamic nucleus oscillations during conflict. J Neurosci. 34:7322–7333.

Zavala B, Zaghloul K, Brown P. 2015. The subthalamic nucleus, oscillations, and conflict. Mov Disord. 30:328–338.

Zhang J. 2008. On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. Artif Intell. 172:1873–1896.