

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/87832>

Please be advised that this information was generated on 2021-03-03 and may be subject to change.

# The Human Basal Ganglia Modulate Frontal-Posterior Connectivity during Attention Shifting

Martine R. van Schouwenburg,<sup>1,2</sup> Hanneke E. M. den Ouden,<sup>1</sup> and Roshan Cools<sup>1,2</sup>

<sup>1</sup>Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, Centre for Cognitive Neuroimaging, 6500 HB, Nijmegen, The Netherlands, and <sup>2</sup>Radboud University Nijmegen Medical Centre, Department of Psychiatry, 6500 HB, Nijmegen, The Netherlands

Current models of flexible cognitive control emphasize the role of the prefrontal cortex. This region has been shown to control attention by biasing information processing in favor of task-relevant representations. However, the prefrontal cortex does not act in isolation. We used functional magnetic resonance imaging combined with nonlinear dynamic causal modeling to demonstrate that the basal ganglia play a role in modulating the top-down influence of the prefrontal cortex on visual processing in humans. Specifically, our results reveal that connectivity between the prefrontal cortex and stimulus-specific visual association areas depends on activity in the ventral striato-pallidum, elicited by salient events leading to shifts in attention. These data integrate disparate literatures on top-down control by the prefrontal cortex and selective gating by the basal ganglia and highlight the importance of the basal ganglia for high-level cognitive control.

## Introduction

The limited processing capacity of our brain requires us to select relevant information for further processing and filter out irrelevant information from our complex environment. According to the biased competition model, this selection is biased by salience (bottom-up processing) as well as behavioral relevance (top-down processing) (Bundesen, 1990; Desimone and Duncan, 1995). Active maintenance of goal-relevant representations in the prefrontal cortex (PFC) allows top-down biasing of attention by modulation of visual processing in posterior cortical regions (Miller and Cohen, 2001; Wallis et al., 2001; Gazzaley et al., 2007; Pessiglione et al., 2008). To facilitate flexibility of attention in response to changes in the environment, these goal-relevant representations need to be updated constantly (Rougier et al., 2005).

The PFC does not act in isolation but rather interacts with other regions, such as the basal ganglia (BG) to bias attentional flexibility. However, their respective contributions are unclear. The BG have long been implicated in the control of movement, and the anatomy of the BG is perfectly suited to selectively gate a desired motor plan to the motor cortex while simultaneously inhibiting competing motor plans (Mink, 1996). Computational modeling work has suggested that the role of the BG in selective gating is not limited to motor processes but extends to cognitive functions. For instance, it has been proposed that goal-relevant representations in PFC are updated only when the BG “open the gate” for cortical processing (Braver and Cohen, 2000; Frank et

al., 2001). This hypothesis is in line with empirical evidence from functional imaging and patient studies revealing a role for the BG in attention switching (Cools et al., 2004; Leber et al., 2008). For example, patients with focal lesions in the BG (Cools et al., 2006) as well as patients with Parkinson’s disease, characterized by BG dysfunction, exhibit attention-switching deficits (Cools et al., 2001a,b, 2003).

In this study, we aimed to elucidate the mechanism by which the BG control attention switching by integrating the hitherto segregated literatures on the role of the PFC in top-down biasing of attention and the role of the BG in selective gating, using functional magnetic resonance imaging (fMRI). In contrast to traditional attention-switching paradigms (e.g., reversal learning, task switching, and set shifting), we used an attention-switching paradigm in which subjects did not switch their attention based on an explicit, top-down cue. Rather, the need to shift attention was signaled by a bottom-up cue consisting of a change in stimuli. We hypothesize that attention switching under such salience-driven conditions is mediated by modulatory influences of the BG on interactions between the PFC and stimulus-specific visual regions in the posterior cortex. To test this hypothesis, we used dynamic causal modeling (DCM), a generic Bayesian framework for inferring effective connectivity from neuroimaging data (Friston et al., 2003). Specifically, we used a nonlinear extension to DCM (Stephan et al., 2008; den Ouden et al., 2010) that allowed us to investigate modulatory influences of activity in the BG on the effective connectivity between PFC and posterior visual regions.

## Materials and Methods

**Subjects.** Twenty healthy right-handed volunteers participated in this study, which was approved by the local ethics committee. Exclusion criteria were claustrophobia, neurological or cardiovascular diseases, psychiatric disorders, regular use of medication or marijuana, use of psychotropic drugs, heavy smoking, or metal parts in the body. All subjects gave written informed

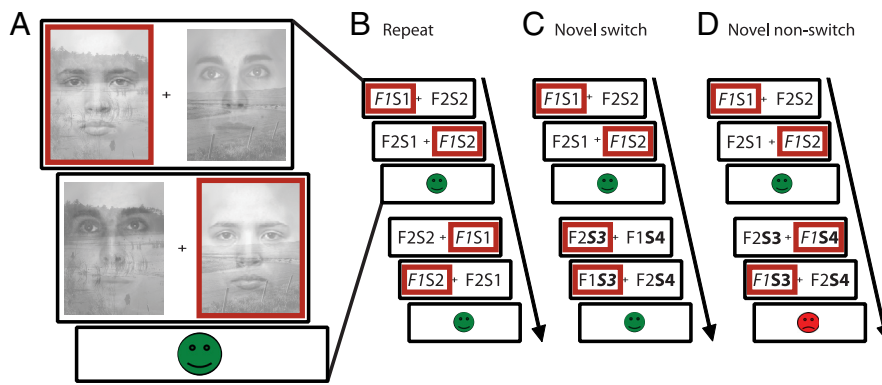
Received March 4, 2010; revised May 12, 2010; accepted June 11, 2010.

R.C. and M.R.v.S. were supported by a Vidi Grant from the Innovational Research Incentives Scheme of the Netherlands Organisation for Scientific Research. We thank Paul Gaalman for assistance with data collection.

Correspondence should be addressed to Martine R. van Schouwenburg, Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, Centre for Cognitive Neuroimaging, P.O. Box 9101, 6500 HB, Nijmegen, The Netherlands. E-mail: martine.vanschouwenburg@donders.ru.nl.

DOI:10.1523/JNEUROSCI.1111-10.2010

Copyright © 2010 the authors 0270-6474/10/309910-09\$15.00/0



**Figure 1.** The attention-switching paradigm used in this study required subjects to select one stimulus exemplar (left vs right) within one dimension (faces vs scenes) on every trial. **A**, Each trial consisted of two consecutive responses followed by feedback. Red boxes indicate a possible response sequence. **B–D** show two consecutive trials with responses defining the three different trial types. For clarification, the stimuli are displayed schematically (F1, face 1; S1, scene 1; F2, face 2; S2, scene 2). **B**, In this example, the subject is attending to F1 on the first trial (attended stimuli are displayed in italic). On the next trial, no novel stimuli are introduced and the subject keeps attending to F1. The second trial is thus defined as a repeat trial. **C**, On a novel switch trial, novel stimuli of the unattended dimension, in this case scenes, are introduced (S3 and S4). The subject detects this change and switches attention to one of two novel stimuli (here S3). **D**, Alternatively the subject can fail to detect the novel stimuli and keep responding to the previously relevant stimulus exemplar, in this case F1. The subject will then receive negative feedback and the second trial is defined as a novel nonswitch trial.

consent and were compensated for participation. Two subjects were excluded from additional analysis because of abnormal performance on the task (see below). Accordingly, data are reported from 18 subjects [seven male; age  $22.4 \pm 0.6$  years (mean  $\pm$  SEM)].

**Paradigm.** A novel attention-switching paradigm was used in which subjects switched attention when they detected a change in the stimulus exemplars of an unattended dimension of two-dimensional stimuli. Subjects were presented with a series of stimulus pairs, i.e., two images presented side by side, each consisting of an overlapping face exemplar and scene exemplar (Fig. 1A). At the beginning of each block, subjects were instructed to select one of the two dimensions (faces or scenes), focus on this dimension, and ignore the other dimension. Within the chosen dimension, subjects then selected one of the two exemplars by making a left (left index finger) or right (right index finger) response, depending on the location of the exemplar of their choice. This self-chosen exemplar was then set as the correct stimulus. Subjects were instructed to continue selecting the correct stimulus on subsequent trials. We used a design similar to that used by Hampshire and Owen (2006), in which stimulus pairs were presented twice within each trial. The combination of face and scene was reversed on the second presentation (F2S1 and F1S2) relative to the first (F1S1 and F2S2). This enabled us to identify the attended stimulus (Fig. 1). At the end of each trial, feedback was presented. Feedback was positive (a green “smiley” face) only if the subject selected the correct stimulus twice within the trial. If subjects selected the pattern that did not contain the correct exemplar or did not respond within a personalized cutoff time, then negative feedback (a red “sad” face) was presented. Thus, a trial consisted of two successive responses followed by feedback, and subjects were explicitly instructed to always respond to the same exemplar within each trial.

After a variable number of correct trials (that is, 2–5 positive feedback events, or 4–10 correct responses) stimuli of the ignored dimension were replaced with novel exemplars. Subjects were instructed to shift their attention to this other dimension and to choose one of the two novel exemplars, whenever they detected a change. On trials on which novel exemplars were introduced (novel trials), subjects either detected the change and switched to one of the novel exemplars [novel switch trials (Fig. 1C)] or they failed to detect the novel exemplars and kept responding to the previously correct exemplar [novel nonswitch trials (Fig. 1D)]. If they failed to detect the change, negative feedback was presented, usually leading subjects to switch on the subsequent trial. Trials on which no novel stimuli were introduced are defined as repeat trials (Fig. 1B).

In the main experiment, subjects were presented with, on average  $355 \pm 15$  trials (mean  $\pm$  SEM), of which 86 were novel trials. The trials were

distributed across four blocks, separated by 23 s breaks. The sequence of the presented faces and scenes was randomized across subjects. For details on the exact timing of the paradigm, we refer to supplemental materials (available at [www.jneurosci.org](http://www.jneurosci.org)).

The paradigm was programmed using Presentation software (Neurobehavioural Systems).

**Localizer.** After completion of the main experiment, subjects performed a one-back task using alternating blocks of face stimuli and scene stimuli to localize the stimulus-specific visual association cortices [i.e., fusiform face area (FFA) (Kanwisher et al., 1997) and parahippocampal place area (PPA) (Epstein and Kanwisher, 1998)]. Subjects were presented with 16 s blocks of 20 face stimuli, 20 scene stimuli (each presented for 300 ms, intertrial interval of 500 ms), and rest periods (seven blocks of each type) and were instructed to press buttons with their left and right index finger whenever they noticed an immediate (1-back) repeat of a stimulus. Acquisition and preprocessing of fMRI data was performed as for the main experiment, and the statistical analysis was conducted using the normalized and smoothed images. In the general linear model (GLM), we included three regressors of interest (scene blocks, face blocks, and rest blocks), and the six realignment parameters as regressors of no interest. The blocks were modeled at the onset of the first stimulus presentation, with a duration of 16 s and convolved with a canonical hemodynamic response. Our contrasts of interest were (1) faces versus scenes and (2) scenes versus faces.

**Behavioral analysis.** The switch likelihood was calculated as the percentage of switches on novel trials. The primary reaction time (RT) data analyses focused on three trial types of interest: novel switch trials, novel nonswitch trials, and repeat trials. Excluded from these primary RT analyses were the first trial of each block, all trials on which subjects received negative feedback (except for the novel nonswitch trials, which by definition resulted in negative feedback), and the trials following negative feedback. Median rather than mean RTs were reported to minimize the influence of outliers. Planned contrasts were assessed using paired sample *t* tests. The statistical threshold was set at  $p < 0.05$  (two-tailed). All results are reported as mean  $\pm$  SEM unless stated otherwise.

**fMRI data acquisition.** Whole-brain imaging was performed on a 3 T MR scanner (Magnetom Trio Tim; Siemens Medical Systems). Functional data were obtained using a gradient-echo echo-planar scanning sequence with blood oxygenation level-dependent (BOLD) contrast (30 axial-oblique slices; repetition time, 1990 ms; echo time, 30 ms; voxel size,  $3.5 \times 3.5 \times 3.0$  mm; interslice gap, 0.5 mm; field of view, 224 mm; flip angle,  $80^\circ$ ). Visual stimuli were projected on a screen and were viewed through a mirror attached to the head coil. In addition, a high-resolution T1-weighted magnetization-prepared rapid-acquisition gradient echo anatomical scan was obtained from each subject (192 sagittal slices; repetition time, 2300 ms; echo time, 3.03 ms; voxel size,  $1.0 \times 1.0 \times 1.0$  mm; field of view, 256 mm).

**fMRI analysis.** Univariate data analysis was performed using SPM5 software (Statistical Parametric Mapping; Wellcome Trust Centre for Cognitive Neuroimaging, London, UK). For the DCM analysis, SPM8 software was used. The first four functional scans of each dataset were discarded to avoid T1 equilibrium effects. Anatomical images were spatially coregistered to the mean of the functional images and normalized using a unified segmentation approach. Preprocessing procedures of functional images included within-subject realignment, spatial normalization using the same transformation matrix as estimated from the anatomical images, and spatial smoothing using a Gaussian kernel of 6 mm full-width at half-maximum. These preprocessed images were used for all analyses.

In a GLM (model A), we included three regressors of interest: novel switch trials, novel nonswitch trials, and repeat trials. In addition, we modeled trials following nonswitch trials, on which subjects switched their attention based on feedback (regressor 4), all error trials, missed trials and trials after an error or after a missed trial (regressor 5), and the six realignment parameters (regressors 6–11) as regressors of no interest. All paradigm-related regressors were modeled as delta functions at the onset of the first stimulus pair presentation within a trial and were convolved with a canonical hemodynamic response function including time derivatives. Time series were high-pass filtered (128 s).

We focused on the following four contrasts: (1) novel switch versus repeat, (2) novel switch versus novel nonswitch, (3) novel nonswitch versus repeat, and (4) novel (both switch and nonswitch) versus repeat. Contrasts from the first (subject-specific) level were used in a second-level random-effects analysis to test for consistent effects across subjects.

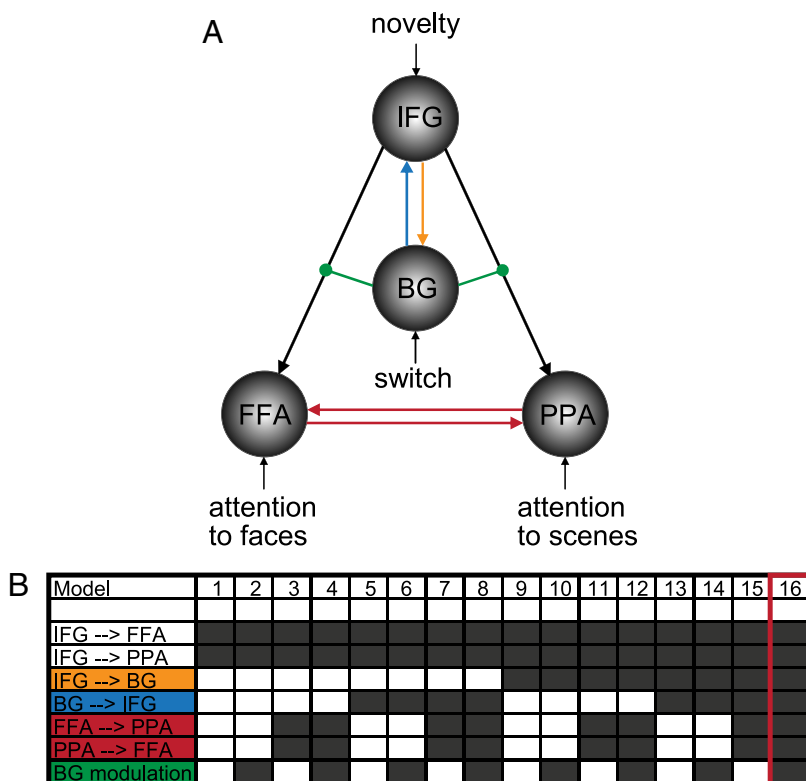
To investigate any stimulus-specific effects in the PPA and FFA, we specified a second GLM (model B) in which novel switch, novel nonswitch, and repeat trials were separated according to whether subjects were attending to faces or scenes. The following trial types were categorized as trials on which subjects attended to faces (vice versa for scenes): (1) novel switch trials on which subjects switched attention to a face, (2) novel nonswitch trials on which subjects failed to detect a novel scene, and (3) repeat trials on which subjects attended to a face. This additional separation of trial types led to a reduction in the number of trials per trial type.

For statistical analysis, we included only those subjects with at least 10 trials per trial type in each comparison (for additional details, see supplemental materials, available at [www.jneurosci.org](http://www.jneurosci.org)).

We report the results of a random-effects analysis, with inferences drawn at the cluster-level, familywise error corrected for multiple comparisons ( $p < 0.05$ ) over the volumes of interest (VOIs). The height threshold at the voxel level was set at  $p < 0.001$  uncorrected for multiple comparisons. Large activation clusters from the insula often blended into clusters in the BG and PFC as a result of smoothing and intersubject differences in anatomy. Therefore, we also report the second or third largest peak voxel if the maximum peak voxel in a VOI was at the border with the insula.

**Volumes of interest.** VOIs in the BG, the PFC, and the primary visual cortex (V1) were defined using the Automated Anatomical Labeling (AAL) interface (Tzourio-Mazoyer et al., 2002). The VOI of V1 was defined as the calcarine sulcus. The VOI of the BG included the caudate nucleus, the putamen, and the pallidum. VOI analyses of the PFC focused on the (right) inferior frontal gyrus pars opercularis (IFG), based on previous results indicating that the right IFG plays an important role in the deliberate and selective focusing of attention on currently relevant information (Gazzaley et al., 2004; Hampshire et al., 2007, 2009; Petrides and Pandya, 2009). For example, the right IFG, but not the middle frontal gyrus, was shown recently to rapidly tune to selectively respond to current targets, becoming less responsive to those same objects when the task demands change (Hampshire et al., 2009). It might be noted that the pattern of BOLD responses in the IFG reported in Figure 5B does not differ qualitatively from that in other regions of the PFC.

Because of large variation in the localization of the FFA and PPA, these were individually defined using an independent localizer task as described above. To define the FFA and PPA VOIs, we used a combination of functional and anatomical constraints. Within the anatomical masks of fusiform gyrus (FFA) and parahippocampal and lingual gyri (PPA)



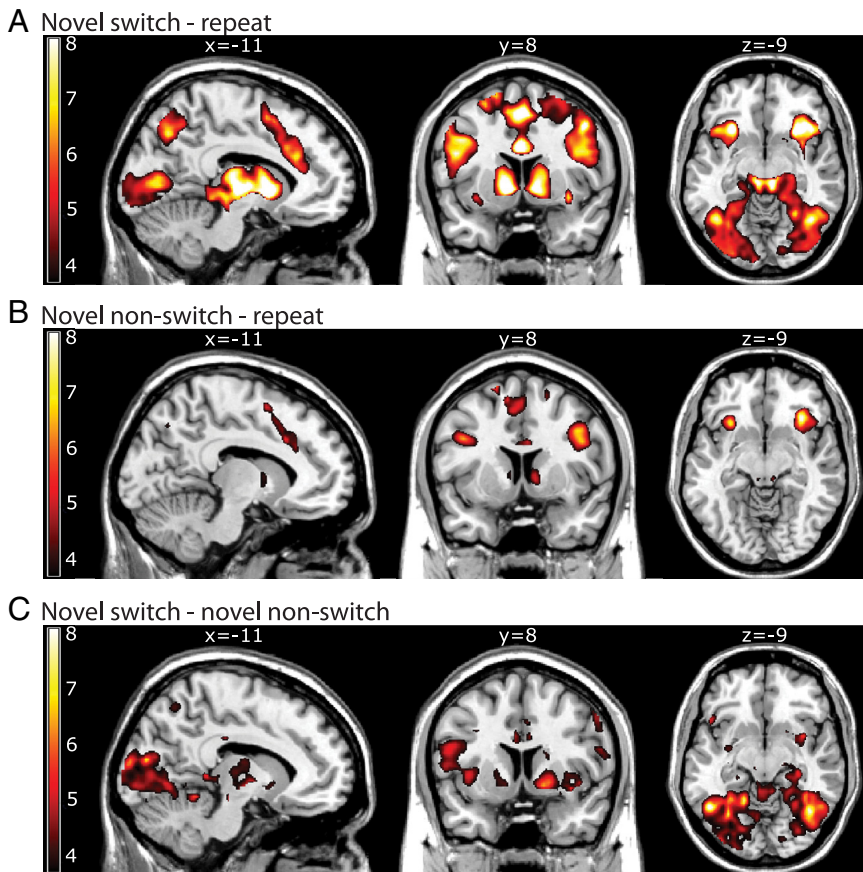
**Figure 2.** Dynamic causal modeling was used to investigate the modulation of connections between the IFG and FFA/PPA by switch-related activity in the BG. **A**, The basic architecture of the model included connections from the IFG to the FFA and PPA (black) and the following inputs: novelty to the IFG, switch to the BG, attention to faces to the FFA, and attention to scenes to the PPA. **B**, We tested 16 alternative models that could include connections from IFG to BG (orange), from BG to IFG (blue), reciprocal connection between FFA and PPA (red), and modulation of IFG to FFA/PPA connectivity by the BG (green). Dark gray boxes indicate that this connection was included in the model. The best model (16) included all connections.

(defined using the AAL interface), the voxel with the highest  $t$  value was determined in the faces versus scenes and scenes versus faces contrasts, respectively, for every subject separately. Voxels that (1) were within the anatomical masks, (2) were within a sphere (radius of 3 mm) around the peak voxel, and (3) exceeded a statistical threshold of  $p = 0.05$  (uncorrected) were included in the subject's FFA and PPA VOIs.

Inferences were drawn based on the whole-brain or VOI analysis, corrected for multiple comparisons at the cluster level. For illustration purposes, we also plotted the  $\beta$  weights for the different trial types for each VOI [extracted using MarsBaR (Brett et al., 2002)]. For the BG, the IFG, and V1,  $\beta$  weights were extracted from the peak voxel at the group level from the novel switch versus repeat contrast. To show the stimulus-specific effects,  $\beta$  weights were extracted from the supplementary GLM (model B, with separate regressors for attention to faces and attention to scenes) from the individually defined FFA and PPA VOIs and averaged over the whole VOI.

**Dynamic causal modeling.** DCM is a hypothesis-driven model of neural dynamics that uses a bilinear or nonlinear state equation to characterize an experimentally perturbed cognitive system (Friston et al., 2003). The original bilinear implementation allows one to estimate effective connectivity between areas as well as modulations of these connections by external parameters. Recently, a nonlinear extension was introduced that allows one to test modulation of effective connectivity between two areas by activity in a third area. We used this nonlinear DCM to test our hypothesis, based on the GLM results and previous findings that top-down influences from the PFC to posterior visual regions were modulated by activity in the BG. More specifically, we tested whether the increased activity in the BG that accompanied novel switch trials modulated connectivity between the IFG and the FFA/PPA.

For a given model, nonlinear DCM models the hidden neural dynamics of a system of interacting brain regions. Using a nonlinear state equation, neural state changes are governed by four sets of parameters: (1)



**Figure 3.** BOLD responses from a whole-brain analysis. Bars indicate  $t$  values, and figures are thresholded for a  $t$  value of 3.65, corresponding to a  $p$ -value of 0.001 uncorrected for multiple comparisons. **A**, Contrasting novel switch trials with repeat trials showed increased responses in the BG, anterior cingulate cortex, IFG, midbrain, parietal cortex, and posterior visual regions. **B**, When comparing novel nonswitch trials with repeat trials, the BG and frontoparietal regions also showed an increase in BOLD responses, but this effect was not observed in posterior visual regions. **C**, Contrasting novel switch trials with novel nonswitch trials showed increased responses in posterior visual regions and the BG.

direct input parameters that model how brain regions respond to external stimuli, known as the “driving inputs,” (2) fixed effective connectivity parameters that reflect the coupling between modeled regions in the absence of input, the “endogenous or intrinsic connections,” (3) changes of these connections induced by experimental conditions, or the “modulatory inputs,” and (4) modulation of intrinsic connections by the neural activity of one of the modeled regions. This model of neural dynamics is combined with a hemodynamic model that describes the transformation of neural activity into a BOLD response. More details about DCM can be found in previous studies (Friston et al., 2003; Penny et al., 2004b; Stephan et al., 2008, 2010).

The posterior probabilities of the parameters from the neural as well as the hemodynamic model are estimated from the measured BOLD data using a Bayesian inversion scheme that rests on an expectation-maximization algorithm (Friston et al., 2003). The posterior distributions of the estimated parameters can then be used to test hypotheses about connection strengths, context-dependent connectivity changes, or the effect of activity in one region on coupling strength between two other regions. In addition, several models can be compared (e.g., including or excluding a particular connection) to test which estimated model optimally describes the measured BOLD responses, using Bayesian model selection (BMS) (as described below).

**DCM specification.** Based on our GLM results, we constructed a nonlinear DCM including the right BG, the IFG, the PPA, and the FFA (Fig. 2). We compared several alternative models, all of which included connections from the IFG to the FFA and the PPA. In addition to this basic architecture, models could include (1) reciprocal connections between the FFA and the PPA to model mutual interaction between these regions, (2) a connection from the IFG to the BG and (3) a connection from the BG to the IFG to test

functional contributions of known recurrent loops between these regions (Alexander et al., 1986), and (4) modulation of the connections from the IFG to the FFA and the PPA by BG activity to test our hypothesis of interest. Connections from the BG to the FFA and PPA were not included based on the fact that our GLM results could not be accounted for by direct effects of BG activity on signal in the FFA and PPA. Varying these model features in a factorial manner resulted in a model space of 16 models (Fig. 2). Note that comparing DCMs with these connections is not equivalent to testing whether these connections do or do not exist anatomically but rather whether these connections play a functional role in the process modeled.

Attention to faces and attention to scenes were modeled as input to the FFA and PPA, respectively. In our paradigm, the need to switch attention between faces and scenes was signaled by novelty. Novelty responses were larger in the IFG than in the BG (see Fig. 5). Accordingly, we modeled novelty as input in the IFG and switching as input to the BG.

Following the notation in previous DCM publications (Friston et al., 2003; Stephan et al., 2008), the hidden neural dynamics of the areas  $x_{1-n}$  in the tested models are described by the following equation:

$$\frac{dx}{dt} = \left( A + \sum_{j=1}^n x_j D^{(j)} \right) x + Cu.$$

Here,  $x$  is the state vector, with each state variable representing the population activity in one region of the model, within total  $n$  regions ( $n = 4$  for FFA, PPA, BG, and IFG).  $t$  is continuous time, and thus  $dx/dt$  is the change in activity in areas  $x$  over time  $t$ . The  $A$  matrix represents the endogenous connection strengths between the modeled regions  $x$ , and  $u$  is the experimentally controlled inputs (attention to faces, attention to scenes, switching, novelty). As can be seen in Figure 2, these external inputs to the system only directly enter into the different areas, the weight of which is represented by the  $C$  matrix, i.e., there are no external modulatory inputs, hence the absence of the  $B$  matrix in this equation. Finally, the  $D^{(j)}$  matrices encode how connection strengths are modulated or gated by activity in area  $j$  (for details, see Stephan et al., 2008).

**Time series extraction.** Because the exact locations of activation maxima varied across participants, we used subject-specific anatomical and functional constraints for selection of regional time series (cf. Stephan et al., 2007a). For the BG, we determined the individual peak voxel that (1) exceeded a threshold of  $p < 0.05$  (uncorrected) in the novel switch versus novel nonswitch contrast, (2) was within the anatomical VOI of the BG, and (3) was within 12 mm of the group maximum in the novel switch versus novel nonswitch contrast. To summarize the regional time series, we computed the first eigenvector across all suprathreshold voxels ( $p < 0.05$  uncorrected) within 3 mm of this peak voxel. For the IFG, we determined the individual peak voxel that (1) exceeded a threshold of  $p < 0.05$  (uncorrected) in the novel versus repeat contrast and (2) was within 6 mm of the group maximum in the novel versus repeat contrast. We then again computed the first eigenvector across all suprathreshold voxels within 3 mm of this peak voxel. For the FFA and PPA, we computed the first eigenvector across all voxels in the individual VOIs.

We were able to extract time series for all four areas in 16 of 18 participants. We could not obtain a BG time series in two participants because of failure to meet the anatomical and functional criteria above. Given

that the complete models could not be specified, these participants were excluded from the DCM analysis.

**Bayesian model selection.** BMS provides a principled foundation for comparing competing models of different complexity (Penny et al., 2004a). We used the negative free energy approximation to the log model evidence (cf. Friston and Stephan, 2007; Stephan et al., 2007b) to compare models at the group level, using random-effects BMS (Stephan et al., 2009). This method is considerably more robust than either the conventional fixed-effects analysis using the group Bayes factor (Stephan et al., 2007b) or frequentist tests applied to model evidences, especially in the presence of outliers (Stephan et al., 2009). It uses variational Bayes to infer the posterior density of the models per se. One can then derive the exceedance probability  $\varphi_k$ , i.e., the probability that a particular model  $k$  is more likely than any other model considered, given the group data.

Note that the model evidence is defined with respect to one particular dataset and that it is therefore not possible to compare models with different numbers of nodes.

## Results

### Behavioral results

There was large individual variability in terms of the likelihood of switching when novel stimuli were introduced, ranging from 31 to 94% (mean  $\pm$  SEM,  $65 \pm 4\%$ ). Two subjects with a switch likelihood above 90% were excluded from additional analysis because of insufficient numbers of novel nonswitch trials.

Subjects were significantly slower on novel switch trials compared with novel nonswitch trials ( $t_{(17)} = 6.0$ ,  $p < 0.001$ ) and compared with repeat trials ( $t_{(17)} = 7.5$ ,  $p < 0.001$ ) (RTs: novel switch,  $1118 \pm 71$  ms; novel nonswitch,  $817 \pm 54$  ms; repeat,  $678 \pm 37$  ms). Conversely, there was no significant difference in RT between novel nonswitch trials and repeat trials ( $t_{(17)} = 1.2$ ,  $p < 0.3$ ). Thus, subjects' performance did not differ between trials in which they continued responding to the same stimulus, independent of whether novel stimuli were introduced in the other stimulus dimension.

On average, subjects made  $9.0 \pm 1.4\%$  errors on repeat trials. Subjects did not respond within the cutoff time (see supplemental materials, available at [www.jneurosci.org](http://www.jneurosci.org)) on  $1.8 \pm 0.3\%$  of repeat trials and on  $2.5 \pm 0.5\%$  of novel trials. Importantly, the number of errors did not correlate with switch likelihood ( $r_{18} = -0.01$ ,  $p = 1.0$ ), indicating that the individual differences in switch likelihood could not be explained by individual differences in the general level of attention, arousal, or motivation.

### fMRI results

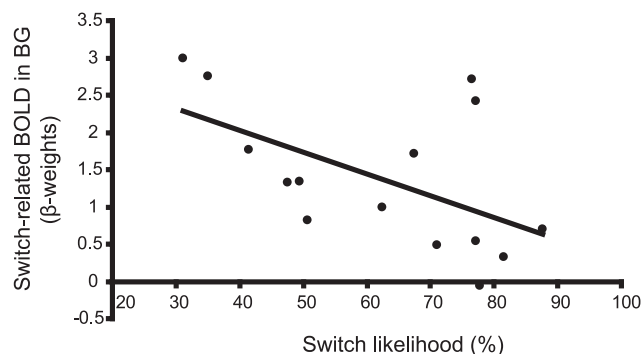
In line with previous findings showing a role for the BG in switching, we found that BOLD signal in the BG was significantly higher during novel switch trials than during repeat trials (Fig. 3A; Table 1) (see also Fig. 5A). This effect was centered on the ventral striatopallidum (VS). Furthermore, there was a significant correlation between BOLD signal in the VS during switching to a novel stimulus and the behavioral measure of switch likelihood across subjects (Fig. 4). This finding strengthens previous observations that the BG are involved in cognitive switching and extends their role in cue-based switches to salience-driven attentional switches that are not driven by instruction cues. Novel switch-related responses were also found in the IFG, V1, the FFA, and the PPA (Figs. 3A, 5A–E; Table 1).

Interestingly, the VS and the IFG showed an increase in BOLD response not only when a novel stimulus caused the subjects to switch their attention but also when a novel stimulus was introduced but not detected. In other words, the VS and the IFG responded to novelty, regardless of whether this novelty elicited an attentional switch (Figs. 3B, 5A, B; Table 1). Conversely, pos-

**Table 1. Coordinates of local maxima within volumes of interest**

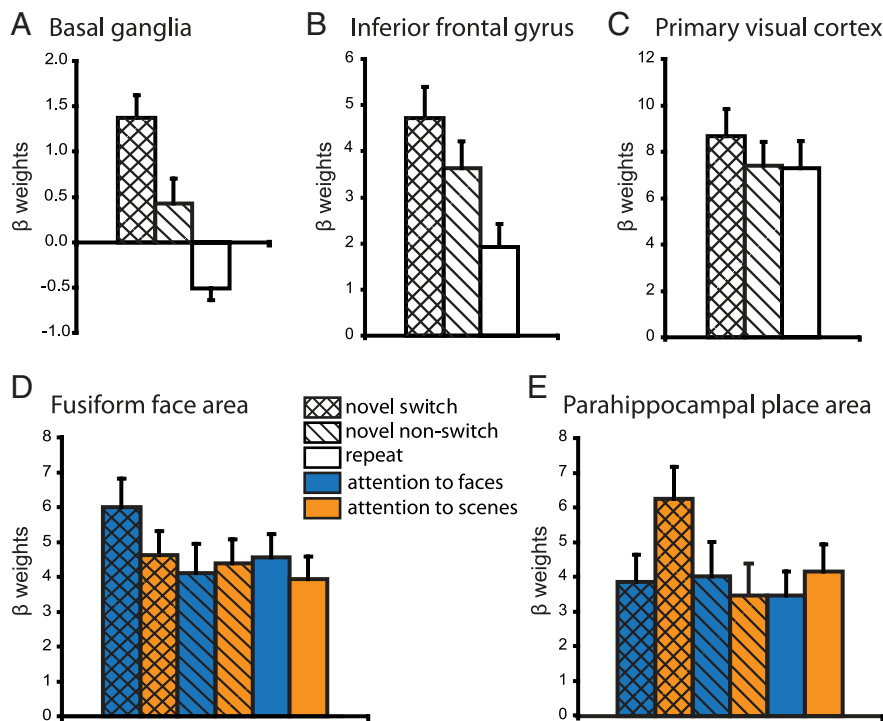
Region	Local maximum			Cluster statistics <i>t</i> value	
	<i>x</i>	<i>y</i>	<i>z</i>		
<b>Novel switch &gt; repeat</b>					
BG	L	−10	8	−2	11.40
	(insula)	26	18	−8	8.85
	R	10	10	0	8.77
	IFG	L	−46	14	20
(insula)		50	18	2	8.78
	(insula)	50	14	20	8.47
	R	48	10	28	7.47
V1	L	−8	−76	6	7.98
	R	14	−80	8	7.88
FFA (VOI)	L				5.69
	R				5.97
PPA (VOI)	L				5.10
	R				6.69
<b>Novel switch &gt; novel nonswitch</b>					
BG	L	−18	−2	18	5.67
	R	18	4	−6	6.53
IFG	L	−54	6	18	5.63
	R	62	18	16	5.27
V1	L	2	−94	4	8.90
	R	16	−88	10	6.49
FFA (VOI)	L				5.53
	R				3.90
PPA (VOI)	L				4.55
	R				4.66
<b>Novel nonswitch &gt; repeat</b>					
BG	L	−10	4	8	4.15
	R	10	8	2	5.01
IFG	L	−48	16	32	7.20
	R	42	8	30	6.71
V1	L	No suprathreshold clusters			
	R	No suprathreshold clusters			
FFA (VOI)	L	Not significant			1.39
	R	Not significant			1.45
PPA (VOI)	L	Not significant			1.75
	R	Not significant			1.63

Clusters were significant at  $p < 0.05$  cluster-level corrected. SPM maps were thresholded at  $p < 0.001$  uncorrected. FFA and PPA statistics were done on mean  $\beta$  weights, extracted from the individual localizer-defined VOIs. All reported coordinates are in Montreal Neurological Institute (MNI) space.



**Figure 4.** Individual differences in behavior could be explained by BOLD signal in the BG. More specifically, the level of BOLD signal on novel switch trials in the BG correlated negatively with the switch likelihood (left BG,  $r_{18} = -0.54$ ,  $p < 0.05$ ; right BG,  $r_{18} = -0.48$ ,  $p < 0.05$ ).  $\beta$  weights were extracted from the group peak voxel from the novel switch versus repeat contrast in the right BG [MNI coordinates  $x, y, z$  (10, 10, 0)].

terior visual regions (V1, FFA, and PPA) showed no increase in BOLD response for novelty per se (novel nonswitch – repeat) (Figs. 3B, 5C–E) but were particularly sensitive to switching as evidenced by the large increase in BOLD signal for the contrast novel switch – novel nonswitch (Figs. 3C, 5C–E; Table 1). This



**Figure 5.** To illustrate the pattern of responses in our VOIs, we extracted the  $\beta$  weights for each subject from the group peak voxels [MNI coordinates ( $x, y, z$ )] from the novel switch versus repeat contrast. Here we display the mean  $\pm$  SEM  $\beta$  weights across subjects. **A, B**, Novel stimuli increased BOLD responses in the BG [coordinates (10, 10, 0)] (**A**) and the IFG [coordinates (48, 10, 28)] (**B**) even when they were not detected. Supplementary repeated-measures ANOVA revealed a significant interaction between region (BG vs IFG) and novelty (novel switch + novel nonswitch vs repeat) ( $F_{(1,1)} = 10.2, p < 0.01$ ), suggesting that the IFG is particularly important for processing novel information. **C**, In contrast, in the primary visual cortex [coordinates (14, -80, 8)], BOLD responses increased only when the novel information elicited an attention switch. **D, E**,  $\beta$  weights for the FFA (**D**) and the PPA (**E**) were extracted from the individual localizer-defined VOIs using the supplementary GLM (model B). These areas showed stimulus-specific effects, such that the BOLD response in the FFA increased when an attention switch was elicited by a novel face but not a novel scene, whereas the reverse effect was found in the PPA.

latter contrast also showed an increase in BOLD responses in the VS, further strengthening the role of the BG in switching (Figs. 3C, 5A; Table 1).

Supplementary GLM analyses (model B) (see Materials and Methods and supplemental materials, available at [www.jneurosci.org](http://www.jneurosci.org)) revealed stimulus-specific effects in the FFA and PPA, such that BOLD responses increased significantly in the FFA on novel switch trials, only when the novel stimulus was a face and not when it was a scene (Fig. 5D). The opposite pattern was obtained in the PPA (Fig. 5E) (for details, see supplemental materials, available at [www.jneurosci.org](http://www.jneurosci.org)). Again, effects were restricted to novel switch trials and did not extend to novel nonswitch trials. This indicates that novel stimuli did not cause an overall increase in BOLD signal in posterior visual regions but that the signal was specifically upregulated in stimulus-specific areas and only when they elicited an attentional switch.

In summary, BOLD responses increased on novel switch trials in the VS, the IFG, and V1 and in a stimulus-specific manner in the FFA and PPA. In addition, novel stimuli were processed by the VS and the IFG, even when these stimuli did not trigger an attention switch. In contrast, it was only when visual information triggered flexible switching in attention that BOLD responses also increased in the primary visual cortex (V1) and stimulus-specific visual association cortices (FFA and PPA).

As outlined in the Introduction, the BG might control salience-driven attention switching by gating the influence of the

IFG to posterior visual regions. Thus, the attentional bias from the PFC on processing in posterior visual regions might be updated only when the BG open the gate in response to novel stimuli. The present results are consistent with this proposal. To test directly the hypothesis that salience-driven attention switching is mediated by VS activity on coupling between the IFG and the FFA and PPA, we used nonlinear DCM, a generic Bayesian framework for inferring hidden neuronal states from measurements of brain activity.

### Nonlinear DCM

Our hypothesis that VS activity modulates frontal-posterior coupling required the assessment of second-order modulatory effects on connectivity. Based on the GLM results, we constructed a nonlinear DCM including the VS, the IFG, the PPA, and the FFA. We constructed 16 alternative models and compared them at the group level. Each model included connections from the IFG to the FFA and the PPA (Fig. 2, black). In addition to this basic architecture, the following connections were systematically included: (1) reciprocal connections between the FFA and the PPA (red), (2) connection from the IFG to the VS (orange), (3) connection from the PPA to the VS to the IFG (blue), and (4) modulation of the connection from the IFG to the FFA and the PPA by VS activity (green), which was driven by switching. We hypothesized

that the switch signal originates in the VS, which subsequently facilitates the IFG-driven biasing of posterior visual regions processing in favor of novel stimuli.

The best model (Fig. 2, model 16) included reciprocal connections between the PPA and the FFA, reciprocal connections between the IFG and the VS, and, critically, modulation by VS activity of connectivity between the IFG and the FFA and the PPA. The exceedance probability for this model was  $\varphi = 0.83$ , surpassing the exceedance probabilities of all other models (which ranged from 0.002 to 0.1) (for details, see Table S1, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). Using model space partitioning, we could directly compare all models with and without the critical modulatory influence from the VS (Stephan et al., 2009). This comparison revealed an exceedance probability of 0.95 in favor of the set of models including this modulatory connection. Having determined the optimal model (Fig. 2, model 16), we then tested whether, in this model, the modulation of the frontal-posterior coupling by VS activity was consistently different from zero across subjects. Indeed, the parameter estimates reflecting gating effects of VS activity on frontal-posterior connections were consistently positive and significant across subjects [effect on IFG  $\rightarrow$  FFA:  $d = 0.54 \pm 0.15$  (mean  $\pm$  SEM),  $t_{(15)} = 3.55, p = 0.003$ ; effect on IFG  $\rightarrow$  PPA:  $d = 0.62 \pm 0.19, t_{(15)} = 3.11, p = 0.007$ ]. Thus, switch-related activity in the VS significantly modulated the strength of connections from the IFG to stimulus-specific visual cortices.

## Discussion

The BG have been implicated in attentional flexibility. Existing evidence indicates that the BG are activated during the performance of set-shifting, reversal learning, and task-switching paradigms (Rogers et al., 2000; Cools et al., 2002, 2004; Leber et al., 2008) and that lesions in this region impair the ability to flexibly switch attention in response to changes in the environment (Cools et al., 2006). However, the mechanism by which the BG control attentional flexibility is unclear. Here we investigate a potential mechanism using a new attention-switching paradigm in which subjects flexibly switched attention only when they detected a change in the unattended dimension of two-dimensional stimuli. The results demonstrate that BOLD responses in the BG, in particular in the ventral parts of the striatum and pallidum (VS), as well as in the PFC were increased when novel stimuli triggered switches in attention. Strikingly, the BOLD signal in these regions also increased during novel stimuli that did not elicit flexible attention switching. In contrast, the primary visual cortex and stimulus-specific visual association cortices responded only when those novel stimuli elicited switches in attention.

The finding that the main effects of stimulus in V1, FFA, and PPA were driven by attention rather than by novelty per se is consistent with many previous studies, reporting similar attentional gain effects in posterior visual regions (Moran and Desimone, 1985). The absence of signal in V1, FFA, and PPA during the novel nonswitch trials relative to the repeat trials is particularly striking and suggests that BOLD in these regions might be driven by top-down signals to a greater extent than by bottom-up signals (Maier et al., 2008). It is precisely the combination of, on the one hand, absence of signal in posterior visual regions, and, on the other hand, presence of signal in the VS that led us to test the hypothesis that the VS might control attentional flexibility by modulating the processing of visual information in posterior visual regions. Given extensive connections between the VS and the inferior parts of the PFC (Alexander et al., 1986) and known attentional influences from the PFC on the FFA/PPA, we hypothesized that such an influence would most likely occur via modulation of inferior prefrontal inputs to posterior regions.

This hypothesis concurs with the basic architecture of current action selection and center surround models of the BG (Hikosaka and Wurtz, 1989; Redgrave et al., 1999b; Nambu et al., 2002), which highlight their role in gating task-relevant cortical programs via the focal release of extensive inhibition mediated by connections between the output nuclei of the BG and the thalamus. This gating function of the BG in the motor domain has been suggested to extend to the domain of attention, selection of eye movement, and the selective updating of task-relevant representations in the PFC (Braver and Cohen, 2000; Frank et al., 2001; Dodds et al., 2009).

Nonlinear DCM enabled us to test the hypothesis that the VS functions as a gate to modulate top-down attentional biasing by the PFC on processing in stimulus-specific posterior visual areas. Consistent with this prediction, we found that our data were best explained by a model that included a modulatory influence of the VS on connectivity between the PFC and stimulus-specific visual regions.

The present finding that attentional flexibility is mediated by influences from VS activity on frontal-posterior coupling was obtained in the context of a paradigm that required switching in response to the introduction of novel exemplars of an unattended stimulus dimension. We hypothesize that the degree of salience of the novel stimuli determined whether they were detected or not. Only changes that reached a certain salience threshold

caused a switch in attention. In other words, the stimulus changes on novel nonswitch trials were not salient enough to trigger an attention switch, but changes on novel switch trials were. This hypothesis is reminiscent of a mechanism suggested for action selection, in which evidence for a certain action accumulates until a threshold is reached, on which the action is executed (Lo and Wang, 2006; Yang and Shadlen, 2007; Forstmann et al., 2008). The BG have been implicated in this process, and, based on the literature and our results, we suggest here that the BG might play similar roles in the domain of attention and action.

Although the VS is often associated with the processing of reward (Schultz, 2007), several studies have revealed a more general role for the VS in the processing of salient information. For example, several fMRI studies have shown increased BOLD responses in the VS in response to novel or surprising nonreward stimuli (Bunzeck and Düzal, 2006; Wittmann et al., 2008; den Ouden et al., 2010). Specifically, Zink et al. (2003, 2006) have found that BOLD signal increased in the VS (and more dorsal parts of the striatum) in proportion to the degree to which an unexpected novel sound interfered with an ongoing task. These data suggest that the salience of a stimulus is reflected in the BOLD responses in the VS. Indeed we found the BOLD signal in parts of the VS to gradually increase over different trial types (Fig. 5A), such that novel switch trials showed an increase in BOLD signal compared with novel nonswitch trials, which in turn showed an increase in BOLD signal compared with repeat trials. Thus, novel stimuli that were not detected caused an increase in BOLD response in the VS. However, they did not affect ongoing behavior in terms of RTs, nor in terms of BOLD responses in the posterior visual regions. We suggest that, although novel stimuli elicited a response in the VS, the evidence on nonswitch trials was not sufficient to trigger attention switching, i.e., did not reach the salience threshold. Note that the present study did not enable us to disentangle whether salience was driven by exogenous (e.g., stimulus properties) or endogenous (e.g., intrinsic brain states) factors. This question is currently under investigation in a follow-up study.

Our finding of a between-subject correlation of switch-related BOLD signal in the VS and behavioral switch likelihood can be reconciled with this hypothesis. Low BOLD signal in the VS during switching was accompanied by high switch likelihood, whereas subjects who showed relatively high BOLD signal during switching were less likely to detect the novel stimulus. At first, a negative correlation may seem counterintuitive, but the critical observation is that this is a between-subject correlation. Although on average the BOLD signal was higher on novel switch trials than on novel nonswitch trials, here we look at individual differences in the height of the switch-related BOLD signal. The observed negative correlation to individual behavioral differences could be explained by the following hypothesis: if attention switching occurs when neural activity in the VS reaches a particular salience threshold, then in subjects with a low salience threshold, a switch will be caused even by a moderately salient stimulus. In these subjects, the average of neural activity across switch trials will be lower than in subjects with a high salience threshold. Salience should be manipulated parametrically in future study to test this hypothesis.

One mechanism by which salient stimuli might influence the selective gating of attention is the regulation of VS activity by the neuromodulator dopamine (DA), which is released in the VS during salient events (Redgrave and Gurney, 2006; Schultz, 2007). This hypothesis is in line with suggestions that short-latency DA signals mediate the switching of attention to unex-



pected, behaviorally relevant stimuli (Redgrave et al., 1999a,b) and concurs with pharmacological functional imaging studies showing that dopaminergic manipulations modulate BOLD signals in the VS (Cools, 2006; Dodds et al., 2008) and its connectivity with the PFC during attention switching (Nagano-Saito et al., 2008). Future studies will test this hypothesis by assessing whether BOLD responses in the VS during the performance of the present paradigm are modulated by administration of dopaminergic drugs.

The finding that responses within the BG and PFC were centered on their ventral inferior parts concurs with proposals that a ventral attentional network is involved when reorienting to behaviorally relevant stimuli, especially when they are salient or unexpected, whereas a dorsal attentional network is involved when selecting stimuli in a more goal-directed manner (Corbetta and Shulman, 2002). The present finding does not imply that all forms of attention switching are mediated by the BG. Indeed, there is evidence that different forms of switching are subserved by distinct cortical and subcortical mechanisms (Cools et al., 2004, 2006; Kehagia et al., 2009).

In summary, we combined the use of a new attention-switching paradigm with fMRI and DCM to test a hypothesized mechanism by which the BG might control attentional flexibility. Our results integrate two hitherto disparate literatures on the role of the PFC in top-down biasing of attention and the role of the BG in selective gating by demonstrating that salience-driven attention switching is accompanied by modulatory influences of activity in the VS on connectivity between the PFC and stimulus-specific visual association cortex.

## References

- Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9:357–381.
- Braver TS, Cohen JD (2000) On the control of control: the role of dopamine in regulating prefrontal function and working memory. In: *Attention and performance XVIII; Control of cognitive processes* (Monsell S, Driver J, eds), pp 713–737. Cambridge, MA: MIT.
- Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. Presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan, June 2–6.
- Bundesden C (1990) A theory of visual-attention. *Psychol Rev* 97:523–547.
- Bunzeck N, Düzal E (2006) Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* 51:369–379.
- Cools R (2006) Dopaminergic modulation of cognitive function—implications for L-DOPA treatment in Parkinson's disease. *Neurosci Biobehav Rev* 30:1–23.
- Cools R, Barker RA, Sahakian BJ, Robbins TW (2001a) Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex* 11:1136–1143.
- Cools R, Barker RA, Sahakian BJ, Robbins TW (2001b) Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain* 124:2503–2512.
- Cools R, Clark L, Owen AM, Robbins TW (2002) Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* 22:4563–4567.
- Cools R, Barker RA, Sahakian BJ, Robbins TW (2003) L-Dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. *Neuropsychologia* 41:1431–1441.
- Cools R, Clark L, Robbins TW (2004) Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *J Neurosci* 24:1129–1135.
- Cools R, Ivry RB, D'Esposito M (2006) The human striatum is necessary for responding to changes in stimulus relevance. *J Cogn Neurosci* 18:1973–1983.
- Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215.
- den Ouden HE, Daunizeau J, Roiser J, Friston KJ, Stephan KE (2010) Striatal prediction error modulates cortical coupling. *J Neurosci* 30:3210–3219.
- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
- Dodds CM, Müller U, Clark L, van Loon A, Cools R, Robbins TW (2008) Methylphenidate has differential effects on blood oxygenation level-dependent signal related to cognitive subprocesses of reversal learning. *J Neurosci* 28:5976–5982.
- Dodds CM, Clark L, Dove A, Regenthal R, Baumann F, Bullmore E, Robbins TW, Müller U (2009) The dopamine D2 receptor antagonist sulpiride modulates striatal BOLD signal during the manipulation of information in working memory. *Psychopharmacology (Berl)* 207:35–45.
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601.
- Forstmann BU, Dutilh G, Brown S, Neumann J, von Cramon DY, Ridderinkhof KR, Wagenmakers EJ (2008) Striatum and pre-SMA facilitate decision-making under time pressure. *Proc Natl Acad Sci U S A* 105:17538–17542.
- Frank MJ, Loughry B, O'Reilly RC (2001) Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cogn Affect Behav Neurosci* 1:137–160.
- Friston KJ, Stephan KE (2007) Free-energy and the brain. *Synthese* 159:417–458.
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19:1273–1302.
- Gazzaley A, Rissman J, D'Esposito M (2004) Functional connectivity during working memory maintenance. *Cogn Affect Behav Neurosci* 4:580–599.
- Gazzaley A, Rissman J, Cooney J, Rutman A, Seibert T, Clapp W, D'Esposito M (2007) Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. *Cereb Cortex* 17 [Suppl 1]:i125–i135.
- Hampshire A, Owen AM (2006) Fractionating attentional control using event-related fMRI. *Cereb Cortex* 16:1679–1689.
- Hampshire A, Duncan J, Owen AM (2007) Selective tuning of the blood oxygenation level-dependent response during simple target detection dissociates human frontoparietal subregions. *J Neurosci* 27:6219–6223.
- Hampshire A, Thompson R, Duncan J, Owen AM (2009) Selective tuning of the right inferior frontal gyrus during target detection. *Cogn Affect Behav Neurosci* 9:103–112.
- Hikosaka O, Wurtz RH (1989) The basal ganglia. *Rev Oculomot Res* 3:257–281.
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.
- Kehagia AA, Cools R, Barker RA, Robbins TW (2009) Switching between abstract rules reflects disease severity but not dopaminergic status in Parkinson's disease. *Neuropsychologia* 47:1117–1127.
- Leber AB, Turk-Browne NB, Chun MM (2008) Neural predictors of moment-to-moment fluctuations in cognitive flexibility. *Proc Natl Acad Sci U S A* 105:13592–13597.
- Lo CC, Wang XJ (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat Neurosci* 9:956–963.
- Maier A, Wilke M, Aura C, Zhu C, Ye FQ, Leopold DA (2008) Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nat Neurosci* 11:1193–1200.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
- Mink JW (1996) The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* 50:381–425.
- Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–784.
- Nagano-Saito A, Leyton M, Monchi O, Goldberg YK, He Y, Dagher A (2008) Dopamine depletion impairs frontostriatal functional connectivity during a set-shifting task. *J Neurosci* 28:3697–3706.
- Nambu A, Tokuno H, Takada M (2002) Functional significance of the cortico-subthalamo-pallidal “hyperdirect” pathway. *Neurosci Res* 43:111–117.
- Penny WD, Stephan KE, Mechelli A, Friston KJ (2004a) Comparing dynamic causal models. *Neuroimage* 22:1157–1172.
- Penny WD, Stephan KE, Mechelli A, Friston KJ (2004b) Modelling functional integration: a comparison of structural equation and dynamic causal models. *Neuroimage* 23 [Suppl 1]:S264–S274.
- Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, Frith CD (2008) Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59:561–567.

- Petrides M, Pandya DN (2009) Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol* 7:e1000170.
- Redgrave P, Gurney K (2006) The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci* 7:967–975.
- Redgrave P, Prescott TJ, Gurney K (1999a) Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* 22:146–151.
- Redgrave P, Prescott TJ, Gurney K (1999b) The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* 89:1009–1023.
- Rogers RD, Andrews TC, Grasby PM, Brooks DJ, Robbins TW (2000) Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J Cogn Neurosci* 12:142–162.
- Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC (2005) Prefrontal cortex and flexible cognitive control: rules without symbols. *Proc Natl Acad Sci U S A* 102:7338–7343.
- Schultz W (2007) Behavioral dopamine signals. *Trends Neurosci* 30:203–210.
- Stephan KE, Harrison LM, Kiebel SJ, David O, Penny WD, Friston KJ (2007a) Dynamic causal models of neural system dynamics: current state and future extensions. *J Biosci* 32:129–144.
- Stephan KE, Marshall JC, Penny WD, Friston KJ, Fink GR (2007b) Inter-hemispheric integration of visual processing during task-driven lateralization. *J Neurosci* 27:3512–3522.
- Stephan KE, Kasper L, Harrison LM, Daunizeau J, den Ouden HE, Breakspear M, Friston KJ (2008) Nonlinear dynamic causal models for fMRI. *Neuroimage* 42:649–662.
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *Neuroimage* 46:1004–1017.
- Stephan KE, Penny WD, Moran RJ, den Ouden HE, Daunizeau J, Friston KJ (2010) Ten simple rules for dynamic causal modeling. *Neuroimage* 49:3099–3109.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. *Nature* 411:953–956.
- Wittmann BC, Daw ND, Seymour B, Dolan RJ (2008) Striatal activity underlies novelty-based choice in humans. *Neuron* 58:967–973.
- Yang T, Shadlen MN (2007) Probabilistic reasoning by neurons. *Nature* 447:1075–1080.
- Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS (2003) Human striatal response to salient nonrewarding stimuli. *J Neurosci* 23:8092–8097.
- Zink CF, Pagnoni G, Chappelow J, Martin-Skurski M, Berns GS (2006) Human striatal activation reflects degree of stimulus saliency. *Neuroimage* 29:977–983.