DOI:
10.1162/jocn_a_01130

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Kehagia, A., Ye, R., Joyce, D., Doyle, O., Rowe, J. B., & Robbins, T. W. (2017). Parsing the roles of the frontal lobes and basal ganglia in task control using multivoxel pattern analysis. *Journal of Cognitive Neuroscience* , *29*(8), 1390-1401. https://doi.org/10.1162/jocn_a_01130

Parsing the Roles of the Frontal Lobes and Basal Ganglia in Task Control Using Multivoxel Pattern Analysis

Angie A. Kehagia[1], Rong Ye[1], Dan W. Joyce[1], Orla M. Doyle[1], James B. Rowe[2,3], and Trevor W. Robbins[2]

Abstract

■ Cognitive control has traditionally been associated with pFC based on observations of deficits in patients with frontal lesions. However, evidence from patients with Parkinson disease indicates that subcortical regions also contribute to control under certain conditions. We scanned 17 healthy volunteers while they performed a task-switching paradigm that previously dissociated performance deficits arising from frontal lesions in comparison with Parkinson disease, as a function of the abstraction of the rules that are switched. From a multivoxel pattern analysis by Gaussian Process Classification, we then estimated the forward (generative) model to infer regional patterns of activity that predict Switch/Repeat behavior between rule conditions. At 1000 permutations, Switch/Repeat classification accuracy for concrete rules was significant in the BG, but at chance in the frontal lobe. The inverse pattern was obtained for abstract rules, whereby the conditions were successfully discriminated in the frontal lobe but not in the BG. This double dissociation highlights the difference between cortical and subcortical contributions to cognitive control and demonstrates the utility of multivariate approaches in investigations of functions that rely on distributed and overlapping neural substrates. ■

INTRODUCTION

Since the inception of cybernetics, the application of control theory to biological systems (Wiener, 1949), neuroscience and medicine have identified control as a dimension of cognition that is critical to adaptive behavior in changing environments. Control deficits have traditionally been associated with damage to the frontal lobe, as a result of disease or trauma, based on demonstrations of rule-shifting impairments in patients with frontal lesions (Reitan & Wolfson, 1994; Stuss, Eskes, & Foster, 1994; Stuss & Benson, 1986; Benton, 1968). Feedback processes are a key characteristic of control and are typified in reinforcement learning as well as rule or policy discovery (e.g., Frank & Badre, 2012). However, to isolate control, it is necessary to unconfound it from learning.

Task-switching paradigms systematically investigate transitions between operational states or task sets, which can be defined as online representations of stimulus–response mappings governed by well-learned rules. From this perspective, task switching lends itself to the operationalization of cognitive control (Allport & Wylie, 2000; Rogers & Monsell, 1995), which is mirrored in the switch cost, embodied in the contrast between task switches and task repetitions. The switch cost theoretically controls for task-specific cognitive components, leaving a pure index of transition that reflects control. Studies on the neural implementation of control have used neuropsychology and neuroimaging to isolate its neural locus in terms of a region or network. In this paper, we challenge the orthodox view that the frontal lobe is necessary for the implementation of control. We use neuroimaging in a design that originates in neuropsychological dissociations (Kehagia, Barker, & Robbins, 2014), indicating that control can be subserved by the BG or frontal regions, depending on the nature of the rules that are switched.

Neuropsychological studies addressing the effects of brain damage and disease on task switching present an inconsistent picture. Patients with frontal lesions exhibit switching deficits only under certain conditions: Switch costs are either generally inflated, indicating inefficiency in the case of left-sided lesions (Mayr, Diedrichsen, Ivry, & Keele, 2006; Aron, Monsell, Sahakian, & Robbins, 2004; Keele & Rafal, 2000; Rogers et al., 1998), or inflated only when control over irrelevant responses is required following right-sided lesions (Mayr et al., 2006; Aron et al., 2004). Other studies however have demonstrated intact switching in these patients when concrete attentional rules are employed (Kehagia et al., 2014). Preserved switching performance in the face of frontal damage seen with such rules is observed after left (Mecklinger, von Cramon, Springer, & Matthes-von Cramon, 1999) or right lesions (Rogers et al., 1998). Moreover, patients with neuropsychiatric disorders that compromise frontostriatal function such as schizophrenia or obsessive-compulsive disorder exhibit intact task switching, despite deficits on other tasks associated with frontal lobe damage (Channon, Gunning, Frankl, & Robertson, 2006; Karayanidis et al., 2006; Moritz, Hubner, & Kluwe, 2004; Manoach

[1]King's College London, [2]University of Cambridge, [3]MRC Cognition and Brain Sciences Unit, Cambridge, UK

Journal of Cognitive Neuroscience 29:8, pp. 1390–1401
doi:10.1162/jocn_a_01130

et al., 2002). Similarly, neurodegenerative disorders that impact on the same circuitry such as Huntington disease (Middleton & Strick, 2000; Robbins & Rogers, 2000; Alexander, DeLong, & Strick, 1986) and Parkinson disease (PD) (e.g., Kehagia, Cools, Barker, & Robbins, 2009; Aron et al., 2003; Cools, Barker, Sahakian, & Robbins, 2003) also reveal a mixed picture of intact and impaired control.

To understand these impairment patterns and tease apart the cortical and subcortical contributions to cognitive control, we previously investigated the comparative neuropsychology of patients with frontal lesions and PD (Kehagia et al., 2014). We compared a condition in which control was implemented through a concrete rule, switching between deterministic rules pertaining to stimuli, with a condition in which control was implemented through abstract judgments of a stimulus property. The demands of switching between concrete and abstract rules highlight differences in terms of cognitive operations, which have significant implications for the organization of control hierarchies: Concrete rules (if the target is on the left, press left key; if it is on the right, press the right key) dictate deterministic responses to stimuli, whereas abstract rules (if the target is an odd number, press the left key; if it is even, press the right key) engender categorical responses to classes of stimuli, enabling response generalization. Crucially, the difference between these rules is not one of monotonic difficulty, where additional load is placed on the same cognitive operation; instead, the new, well-characterized cognitive component of activating a new set of stimulus–response mappings is introduced when switching between abstract rules (see also, Kehagia et al., 2009). Patients with frontal lesions were impaired only when switching between abstract but not concrete rules. PD patients demonstrated abstract switching deficits only when their disease had progressed (Hoehn & Yahr Stage II), but not when it was in its earliest stages with predominantly subcortical pathology (Hoehn & Yahr Stage I). This neuropsychological dissociation holds also at the neurochemical level: dopaminergic manipulations in PD affect concrete (e.g., Cools, Barker, Sahakian, & Robbins, 2001) but not abstract switches (Kehagia et al., 2009), or switches between abstract stimulus dimensions (for a review, see Kehagia, Murray, & Robbins, 2010). Thus, rule abstraction is a useful tool in systematizing disparate findings and characterizing possible differential roles of frontal regions and the BG in cognitive control.

This issue of cortical and subcortical contributions to cognitive control or flexibility as a function of rule abstraction has not yet been addressed in a neuroimaging context. Distributed, left-lateralized activations in frontoparietal areas are often reported for the critical switch versus repeat contrast, which include lateral pFC, insula, anterior cingulate, SMA, pre-SMA (e.g., Badre & Wagner, 2006; Slagter et al., 2006; Yeung, Nystrom, Aronson, & Cohen, 2006; Brass, Ullsperger, Knoesche, von Cramon, & Phillips, 2005; Ruge et al., 2005; Brass & von Cramon, 2004; Braver, Reynolds, & Donaldson, 2003; Rushworth, Hadland, Paus, & Sipila, 2002; Rushworth, Passingham, & Nobre, 2002), and, sometimes, their subcortical connections to the striatum (Crone, Wendelken, Donohue, & Bunge, 2006; Woodward, Ruff, & Ngan, 2006; Barber & Carter, 2005). Attempts to reconcile the precise contributions of different regions to discrete aspects of control have focused on cortical regions: for example, a rostrocaudal gradient of frontoparietal activation has been proposed to reflect degrees of representational abstraction in the rules that are switched (Kim, Johnson, Cilles, & Gold, 2011). Alternatively, the inferior frontal junction and posterior parietal cortex have been promoted as a common locus of control irrespective of abstraction (Kim, Cilles, Johnson, & Gold, 2012). However, the frontostriatal components of control have yet to be dissociated.

Given the intrinsic anatomical and functional complexity of frontal and subcortical regions during cognitive control and the former's concomitant, frequently observed association with variable task-related information, we adopted a multivoxel pattern analysis (MVPA) to resolve voxel-wise spatial patterns by accounting for simultaneous regional activation patterns. There is good agreement that the activity of one voxel often represents a mixed functionality, confounding the interpretation of BOLD signal changes in voxel-wise patterns. For example, monkey electrophysiology relates activation of frontal and parietal neurons to various task components and rules (Sigala, Kusunoki, Nimmo-Smith, Gaffan, & Duncan, 2008; Muhammad, Wallis, & Miller, 2006; Stoet & Snyder, 2004; Niki & Watanabe, 1976). In human neuroimaging, MVPA is predicated on acknowledging variation in neuronal function across or within voxels (Haynes & Rees, 2005), because domain-specific subpopulations of neurons within an area associated with different task-related representations (Chiu & Yantis, 2009) will be included as a mixture in any given voxel. With MVPA, it is possible to exploit the aggregation of otherwise weak information across multiple locations, which, only when considered in tandem but not individually, may be meaningfully related to cognition. This is achieved by addressing differences in terms of the activation patterns evoked by different events within the whole brain or specific regions. Inference can proceed at the individual subject level as well as at the group level. MVPA has been used to "decode" the nature of rule representation in the brain, specifically in regions such as the frontal cortex (e.g., Reverberi, Gorgen, & Haynes, 2012a) and even implicated the striatum in the representation of abstract rules (Reverberi, Gorgen, & Haynes, 2012b). This is a related, but different, question to the one addressed in the current investigation on cognitive control. We sought to interrogate not the representational content of the frontal lobes and BG in relation to rules per se but rather the implementation of control over these by addressing how each of these regions contributes to flexible behavior in the context of different, well-learned rules.

We extend our previous neuropsychological findings to the neuroimaging domain and interrogate the existence of a double dissociation in control over abstract and concrete rules, as dissociable processes subserved by frontal and subcortical regions using MVPA. Gaussian process classifiers (GPCs) are a robust Bayesian approach to MVPA. By training GPCs to classify individual participants' patterns of frontal lobe and BG voxel activity for switches and repeats, across abstract and concrete rules, we elicited the neural substrates of cognitive control distributed over these regions. Classification was predicted to be at chance level in both rule conditions in the temporal lobe, which was selected as a region with no central involvement in control.

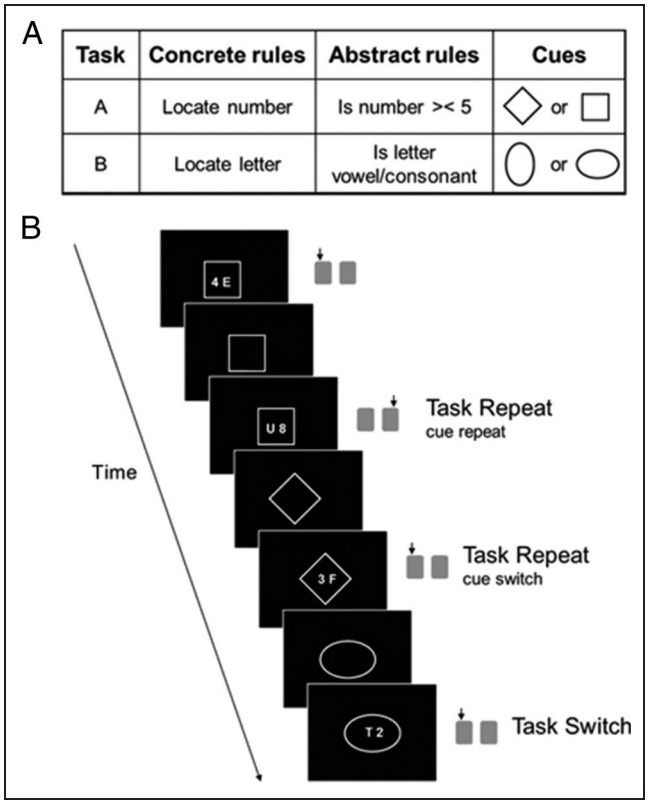

Figure 1. (A) Table representing rule conditions, tasks, and cue mappings. (B) Example of a trial sequence in the concrete rule condition, where Task A (locate the number) repeats twice and a cue switch occurs on the second repeat, before switching to Task B (locate the letter).

METHODS

Participants

Eighteen right-handed healthy volunteers (mean age = 25.2 years, range = 23–33 years, nine women) completed a single testing session. Data from one participant were excluded from analysis because of excessive head movement (>3 mm). Exclusion criteria were a history of neurological or psychiatric disease, drug abuse, and concurrent medication. All participants gave written informed consent and received monetary compensation of £20. The protocol was approved by the Cambridgeshire research ethics committee (LREC 07/H0311/213), and the study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

fMRI Task

The task was run on E-Prime 2.0 (Psychology Tools, Inc., Pittsburgh, PA) during a single EPI session of 25 min. There were eight blocks in total: four blocks were governed by abstract categorization rules and four blocks were governed by concrete attentional rules. An instruction screen at the start of each block informed participants whether they would be switching between abstract categorizations or concrete rules, with a brief summary of the stimulus and response mappings for each task set. On each trial, a digit (1, 2, 3, 4, 6, 7, 8, or 9) and a letter (A, E, I, U, C, F, T, or X) were presented as a pair in the centre of the screen. Within each rule block, participants were required to switch between two tasks: one was based on the letter, and one was based on the numerical digit. Each block comprised 49 trials of these pseudorandomly alternating tasks, yielding a balanced mix of task Switch and task Repeat trials.

The task on each trial was cued by the shape surrounding the character pair. To unconfound task switching from cue switching, we employed two cues to signal each task. An angular shape (a square or a diamond) cued the digit task, and a curved shape (a vertical or a horizontal ellipse) cued the letter task (Figure 1). For example, within the pseudorandomized task sequence of any given block, the digit task could be cued on one trial with a square and on another with a diamond, and similarly for the letter task using a vertical and a horizontal ellipse. Whether a given trial is classed as a Repeat or a Switch is a function of the task that preceded it: In a sequence of BAA, the first Task A (digit) trial is a Switch, as it immediately follows a Task B (letter) trial, and the second is a Repeat, as it follows a trial of the same task kind. By definition, the cue always also switches on a Switch trial, confounding this switch of task with a switch of cue, a controversial issue in the task-switching literature (Schneider & Logan, 2005). By assigning one of two cues to each task, it is possible to generate Repeat trials where the cue has also switched, which can then be legitimately compared to Switch trials (where the cue has also switched) to yield a measure of task-switching cost unconfounded by cue switching.

In the concrete rule condition, the stimuli were presented side by side. Participants switched between two tasks in which they were instructed to respond to the location of the digit (Task A) or letter (Task B), using the index (L) and middle (R) finger of their right hand, in a spatially compatible manner to the target, which could be on the left or right. We define concrete attentional rules ("if the stimulus is on the left/right, press the left/right key")

as comprising 1:1, deterministic responses to stimuli: In the two localization tasks, a digit on the left maps to left button press, as does a letter on the left. Thus, on a switch of task, the same deterministic relationship between stimuli and responses (1:1) applies to different stimulus sets after, for example, a task switch from letters to digits has taken place.

In the Abstract rule condition, participants performed a numeric judgment task for digits (Task A is the digit greater or less than 5) and a grammatical judgment task for letters (Task B is the letter a vowel or consonant). To minimize response selection interference or the Simon effect (Simon, 1969), stimuli were presented one above the other. "Less than 5" and "vowel" responses were mapped to the index finger of the right hand, and "greater than 5" and "consonant" were mapped to the middle finger. In what we define as abstract rules, an *n*:1 stimulus–response correspondence applies to both task sets as multiple stimuli map to a single response (e.g., numbers 1–4 map to "less than 5," or A, E, I, U map to "vowel"). Switching between abstract rules engenders an equivalent stimulus set reconfiguration, but also a new set of responses, such that the left/right key press acquires a new meaning (from "vowel" to "less than five"). Switching between abstract rule-governed tasks engenders different cognitive and neural processes given the difference in the degree of response set reconfiguration.

At the beginning of each trial, the cue was presented on the screen, and 300 msec later the stimulus pair appeared within the cue. The stimulus was cleared as soon as a response was made or after a maximum of 2000 msec. Following the response, a central fixation cross appeared during a variable intertrial interval of 1700, 1825, or 1950 msec. No feedback was given.

The stimulus sequence was permuted to achieve complete counterbalancing of all stimulus types within and between conditions and blocks, avoiding repetition of stimuli on successive trials. Within each rule condition, trials were presented in a fixed pseudorandom order to ensure that (a) a task switch, a task repetition with a cue switch, and a task repetition with a cue repetition were equally probable; (b) target and distracter stimuli were counterbalanced across these conditions but did not repeat over consecutive trials; (c) cues were counterbalanced across conditions; and (d) response repetitions and response switches were counterbalanced across switch and repeat trials within each task type. Within these constraints, a new trial sequence was generated using Matlab 7.1 (www.mathworks.com) for each participant. The cue–task and target–response mappings were counterbalanced across participants who were trained in a practice session within 3 days of the fMRI session.

fMRI Acquisition

Participants were scanned at the Medical Research Council Cognition and Brain Sciences Unit, UK, on a 3-T Tim Trio Magnetic Resonance Imaging scanner (Siemens, Berlin, Germany) with an eight-channel head coil. Whole-brain data were acquired with echo-planar T2-weighted imaging (EPI), sensitive to BOLD signal contrast (36 sequential oblique axial 3-mm slices, distance factor = 25%; repetition time = 2000 msec; echo time = 30 msec; flip angle = 78°; field of view = 192 mm; in-plane resolution = 3 × 3 × 3 mm). Stimuli were back-projected onto a mirror on the head coil. Responses were made using the index or middle finger on a button box. High-resolution MPRAGE anatomical images (repetition time = 2250 msec, echo time = 2.99 msec, flip angle = 9°, inversion time 900 msec, 256 × 256 × 192 isotropic 1 mm voxels) were collected for anatomical localization and coregistration.

Behavioral Data Analyses

RTs for error and post-error trials were excluded. Error rates were arcsin-transformed, as the variance was proportional to the mean (Howell, 1997). RT and error rates were subjected to repeated measures ANOVA, with Rule (Abstract vs. Concrete) and Switch (Repeat vs. Switch) as the within-subject factors. Greenhouse–Geisser corrections were applied to adjust for non-sphericity where necessary.

fMRI Data Analyses

Data were preprocessed using SPM8 (www.fil.ion.ucl.ac.uk/spm/). The first six volumes were discarded to allow for T1-equilibrium effects. The EPI images were realigned to the first scan by rigid body transformations to correct for head movement and sinc interpolated in time to correct for slice time differences. EPI and structural images were coregistered and then normalized to the T1 standard template in MNI space (Montreal Neurological Institute International Consortium for Brain Mapping) and smoothed with an 8-mm FWHM Gaussian kernel, as a compromise between the larger and smaller kernels typically used for the pFC and BG in neuroimaging studies. Low-frequency signal drift was removed using a high-pass filter (cutoff 128 sec).

To model the task, we implemented a variable epoch model within the general linear model framework (Henson, 2007), which is widely used for designs including complex modulation of task conditions with both univariate (Crittenden & Duncan, 2014) and multivariate (Woolgar, Hampshire, Thompson, & Duncan, 2011) frameworks. This models each trial with a boxcar epoch function whose duration is equal to the RT on that trial, which assumes a constant hemodynamic response at the voxel level throughout the time course of the task. Because it takes into account RT variability, variable epoch modeling enhances statistical power and improves the interpretability of changes in neural activation (Grinband, Wager, Lindquist, Ferrera, & Hirsch, 2008). Moreover, explicitly

modeling variables such as individual RT differences is one way of removing within condition variance to maximize sensitivity in multivariate analyses (Haynes, 2015).

Each experimental regressor in the general linear model was constructed using RT as trial duration for task Switch and task Repeat (only those which also contained a cue switch were included) in the two rule conditions. Motion parameters were included as nuisance covariates. The first trial in each block and errors were modeled separately. Stimulus events were convolved with the canonical hemodynamic response function.

Classifier Model and Training

Individual participants' beta maps for each of the four experimental regressors (Switch and Repeat in the Abstract and Concrete rule conditions) were used as inputs to different GPCs for the abstract and concrete rule conditions, respectively (see below). Gaussian Process Classification is a Bayesian method that learns a mapping between a multivariate input, in this case, beta maps, and a discrete output—a class label of interest, in this case, "Switch" versus "Repeat." For a detailed exposition of GPCs, we refer the reader to Rasmussen and Williams (2006). GPC has been well validated in neuroimaging (see Schrouff, Kusse, Wehenkel, Maquet, & Phillips, 2012; Pereira, Mitchell, & Botvinick, 2009). An important consideration for any multivariate modeling where the number of "inputs" to the model (i.e., the size of the beta maps) is orders of magnitude larger than the number of observations (i.e., participants) is "overfitting." When learning a classification mapping, the model is trained on a subset of the data and then validated using a "hold-out" set to ensure the learned model can generalize beyond the training set, ensuring that a robust model of the underlying joint distribution of inputs and outputs has been acquired rather than simply "memorizing" the data. An advantage of GPCs is that they incorporate regularization that prevents overfitting by virtue of their Bayesian foundations. Given the limited number of participants, the GPCs were also trained using leave-one-out cross-validation (LOOCV), whereby the data of all but one participant (i.e., a hold-out set of one) are used to train the model and the data from this remaining single participant are used to independently validate the model's classification performance. This method results in more conservative estimates of classifier performance (Hastie, Tibshirani, & Friedman, 2003) than more liberal approaches such as *k*-fold cross-validation.

GPCs were implemented using a linear kernel within the PIPR toolbox (www.kcl.ac.uk/ioppn/depts/neuroimaging/research/imaginganalysis/Software/PIPR.aspx), which uses the GPML library implementation of GPCs (www.GaussianProcess.org/gpml/code). To assess the statistical significance of classification performance, permutation testing was used where the LOOCV training procedure was repeated 1000 times, with randomly permuted training labels (Switch vs. Repeat) to generate a null distribution for testing the hypothesis that the Switch and Repeat beta images cannot discriminate between the Switch and Repeat conditions above chance (i.e., no better than 50% accuracy).

Classifier Construction and ROIs

This proceeded on the basis of our earlier neuropsychological findings in patients with frontal lobe lesions and PD and our hypothesized double dissociation between the frontal lobe and the BG. These ROIs, along with the temporal lobe as the "control" region, were defined anatomically using the Harvard–Oxford cortical and subcortical atlases (Desikan et al., 2006) using FSL (version 4.1.8, FMRIB software library, www.fmrib.ox.ac.uk/fsl; see Table A1 for individual regions comprising the BG, frontal, and temporal lobe masks). The total number of voxels ($2 \times 2 \times 2$ mm) for the BG, frontal, and temporal lobe masks was 3585, 82,127, and 41,728, respectively.

We trained different classifiers to discriminate between classes (Switch and Repeat) separately for the Abstract and Concrete rules. For each classifier, beta maps were extracted from the frontal lobe mask and BG mask. Thus, in total, we constructed six GPCs. For the Abstract condition, we trained one GPC that takes as input frontal lobe beta images and learns Switch and Repeat classifications, and similarly, a second GPC but for the BG beta images and a third for the temporal lobe beta images. For the Concrete condition, one GPC was trained with inputs from the frontal lobe beta images (learning Switch and Repeat classifications), a second from the BG, and a third from the temporal lobe. Following training, the classifiers report the probability that any given participant's beta image (from the frontal lobe, BG, temporal lobe with either abstract or concrete rules) belong to a Switch or Repeat condition. Moreover, we directly compared these classification accuracies within each ROI to demonstrate that, within each mask, classification accuracy was significantly different between the abstract and concrete conditions. The predicted class (Switch vs. Repeat) probabilities were obtained for each participant when that participant was the hold-out sample in LOOCV training to ensure class probabilities represent independent samples of the more conservative classifier generalization performance rather than optimistic biased estimates attributable to having being "seen" by the classifier on multiple LOOCV training sets. The distributions of these predictive probabilities for each of the four classifiers for each subject were compared using Kolmogorov–Smirnov tests, as these data obtained non-parametrically were neither normal nor unimodal.

To visualize the voxels driving the classification performance for each of the four trained GPCs for the regions pertaining to our hypothesis (the frontal lobe and BG), the parameters (e.g., weight maps) of a forward (generative)

model were estimated from the parameters of each trained discriminative GPC (backward) model. Forward models produce a pattern of weights for each voxel's contribution to the trained GPC classifier that are more closely related to changes in the underlying neural activity, avoiding difficulties resulting from directly interpreting the parameters learned by multivariate classifiers (Haufe et al., 2014) such as the sensitivity of the backward model to the noise components in the data. The resulting forward model (see equation 6 in Haufe et al., 2014) provides an interpretable visualization of the voxels driving the classifier, where the magnitude and sign of each voxel in the weight map denotes its contribution to predicting the class. In other words, the maps presented in Figures 2 and 3 represent voxels whose direction and magnitude contribute to classification accuracy, rather than simply voxels whose collective magnitudes drive classification. This renders our analysis directly interpretable in the same way as a mass-univariate analysis, and enables us to support our claims on the basis of maps of voxels whose direction and magnitude contribute significantly. The resulting forward maps for the two GPCs that achieved statistically significant classification performance were overlaid onto high-resolution T1 structural images using MRIcroGL (www.mccauslandcenter.sc.edu/mricrogl/) and MRIcron (Rorden & Brett, 2000).

RESULTS

Behavior

There were significant main effects of Rule, $F(1, 16) = 98.99$, $p < .0001$, and Switch, $F(1, 16) = 55.95$, $p < .0001$, confirming switch costs. Switch costs were observed with both rule types (Abstract: $F(1, 16) = 60.63$, $p < .0001$; Concrete: $F(1, 16) = 15.82$, $p < .001$), but the Rule × Switch interaction, $F(1, 16) = 40.57$, $p < .0001$, indicated greater switch costs with abstract rules. Although cue repetition was not the primary focus of this investigation, a one-way repeated measures ANOVA on repeat trials with cue switch (repeat vs. switch) as the within-subject factor revealed no significant effect on RT ($F < 1$). Error rates were on average 5.5% and were subjected to the same two-way within-subject ANOVA. There was a significant effect of Switch, $F(1, 16) = 29.79$, $p < .0001$, paralleling the RT effects, but no effects of Rule or an interaction ($F < 1$). The behavioral data are presented in Table 1.

Neuroimaging

Examining the classification performance within and between the frontal lobe and BG masks yielded a double dissociation. The GPC trained for Switch/Repeat classification

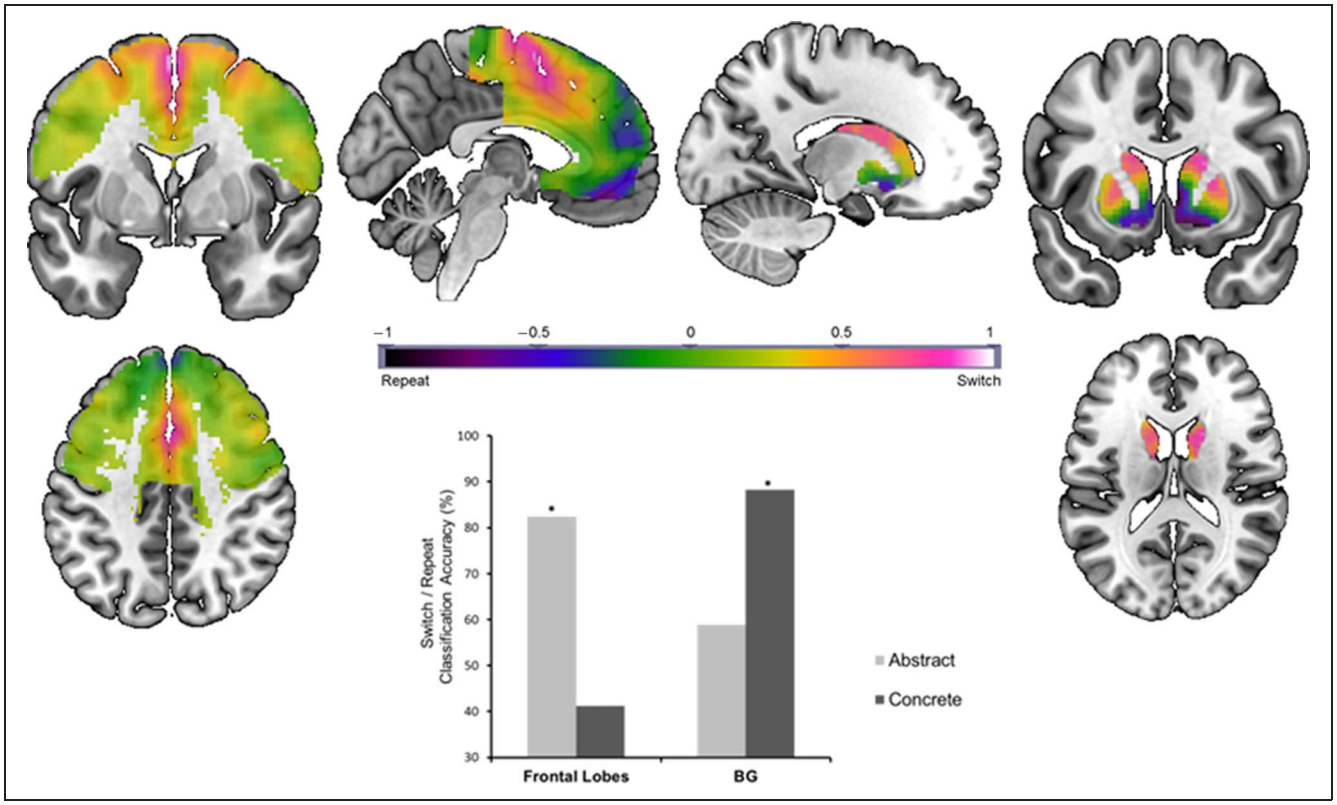

Figure 2. Multiplanar forward model maps showing relative contribution of voxels for Switch/Repeat classes with abstract (left hand panel) and concrete rules (right hand panel). Scale bar: Reddish/pink colors (>0) indicate voxels more implicated in classifying Switch. Bluish/purple colors (<0) indicate voxels more implicated in classifying Repeat. The central figure represents the double dissociation between Switch/Repeat classification accuracy for the abstract and concrete rule conditions in the frontal lobes and the BG, respectively.

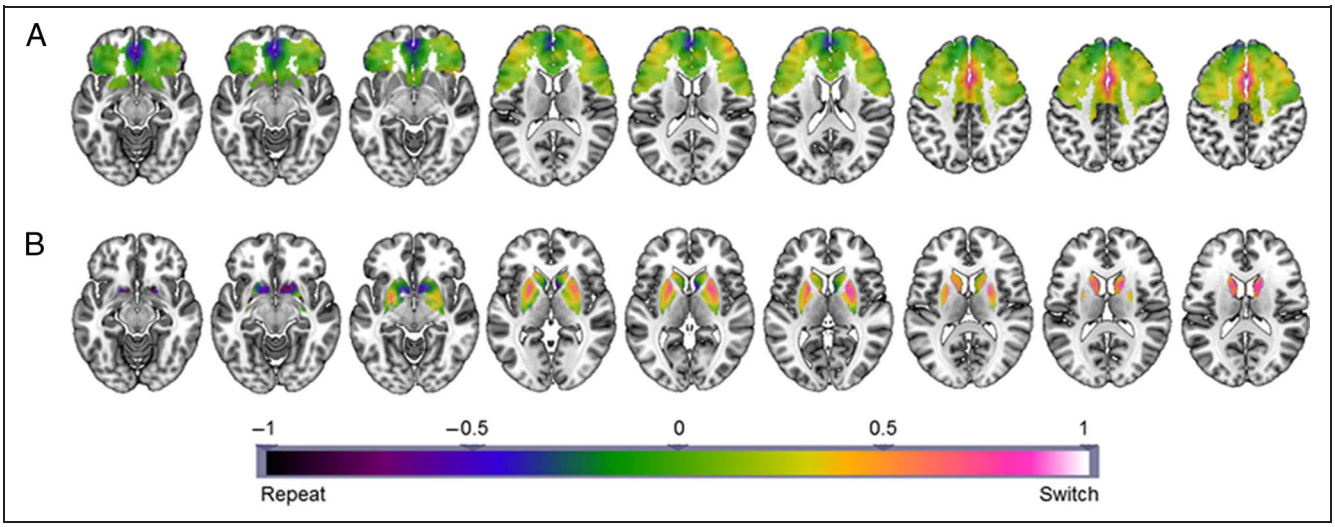

Figure 3. Axial slice renderings illustrating forward model maps. Scale bar as for Figure 2. For abstract rules in the frontal lobe (A), voxels favoring Switch were greatest in the medial superior frontal gyrus whereas voxels favoring Repeat clustered around medial orbitofrontal regions. For concrete rules in the BG (B), voxels favoring Switch concentrated on the caudate nucleus and putamen whereas those favoring Repeat were clustered in the nucleus accumbens.

for abstract rules within the frontal lobe yielded statistically significant classification accuracy of 82.35% (p = .022, 1000 permutations), but the GPC for the BG did not (classification accuracy 41.18%, p = .791). Conversely, the GPC trained for Switch/Repeat classification for concrete rules in the BG achieved statistically significant classification accuracy at 88.24% (p = .003), but the corresponding GPC for the frontal lobe did not (classification accuracy 58.82%, p = .255). No significant classification accuracy was achieved by the GPCs trained for Switch/Repeat classification in the temporal lobe with either abstract (classification accuracy 64.17%, p = .17) or concrete rules (classification accuracy 52.94%, p = .5). The plot of the double dissociation in classification accuracies and voxel contributions visualized from the forward model within the frontal lobe and the BG are presented in Figure 2.

The double dissociation was directly tested by comparing the predictive probabilities for all subjects in each of the four classification conditions using Kolmogorov–Smirnov tests. The abstract–concrete predictive probability distributions were significantly different both in the frontal lobe (D = 0.588, p = .005) as well as the BG (D = 0.529, p = .016) masks.

The forward model voxel weight maps created by forward modeling and driving Switch/Repeat classification are presented in Figures 2 and 3A, B. In the Abstract rule condition (Figure 3A), the frontal lobe classification weight pattern favoring the Switch class highlights the mid cingulum and the medial frontal gyri, with the most prominent weights concentrated around the superior frontal gyrus medially. Conversely, lateral and medial orbitofrontal regions exhibited greater weights for the Repeat class. In the Concrete rule condition, the BG classification weight pattern (Figure 3B) favoring Switch is focused on dorsal regions of the putamen and caudate, particularly the head, whereas the greatest weights for Repeats were concentrated ventrally in the nucleus accumbens.

Table 1. Behavioral Mean (*SEM*) RT and Error Rates

	RT (msec)	*Errors (%)*
Concrete Rules		
Repeat	819 (61)	3 (0.6)
Switch	868 (64)	8.1 (1.5)
Abstract Rules		
Repeat	989 (66)	3.6 (0.6)
Switch	1173 (76)	7.5 (1.2)

DISCUSSION

We have demonstrated a double dissociation between cortical and subcortical contributions to the implementation of cognitive control at different levels within a hierarchy of rule-based behavior. We used GPC to discriminate BOLD signal changes associated with switching and repeating abstract and concrete rules, in two distinct domains, the frontal lobes and the BG. In our study, Switch/Repeat classification specifically seeks to identify whether voxel-wise activation patterns seen on Switches are different from those seen on Repeats in the frontal lobe and the BG, respectively. MVPA succeeded in isolating the differences between these regional patterns,

which are both associated with the same task set specific processes but differ in terms of the state they had assumed on the immediately preceding trial. For a Switch trial, this state reflected a different task rule, whereas the immediately preceding state of a Repeat trial reflected the same task rule. In this way, the "signature" of control can be isolated from the cognitive operations associated with executing any one particular task. Switch/Repeat classification accuracy for concrete rules was significant in the BG, but at chance level in the frontal lobe. The inverse pattern was obtained for abstract rules, where classification was significant in the frontal lobe, but at chance level in the BG.

This regional dissociation cannot be explained either by differences in stimulus materials or by the cue switching confound present in many other task-switching studies, differences in preparation time, strategy, or learning, due to the training session, which ensured stable performance with these well-learned rules. It is also not attributable to a main effect of differences in task-specific activity, which is subtracted out in the comparison of switch versus a repetition of the same task. In evaluating the emergent double dissociation, it is nonetheless important to discount the possibility that the current pattern of results is artifactual, stemming for example from differences in the dimensionality of the frontal lobe and BG data, whereby a weaker "signal" associated with Concrete Switch/Repeat classification was missed in the larger frontal lobe mask. However, GPC performance was at chance level for both rule conditions in the temporal lobe, a large control region of the same order of magnitude as the frontal lobe, where no significant classification was predicted. Moreover, we might have predicted abstract Switch/Repeat classification to reach significance in the smaller BG mask; the opposite was observed.

We predicted this double dissociation on the basis of earlier neuropsychological findings of abstract switching deficits in patients with frontal lesions and advanced PD, in the face of intact performance with concrete rules seen in the frontal lesion group and patients with early PD (Kehagia et al., 2014). Our results demonstrate that rule abstraction is a powerful way of simultaneously engaging and dissociating the contributions of distinct brain regions during cognitive control in task switching. They carry significant implications for neuropsychological assessment, as they indicate that the concept of a generic prefrontal substrate for executive control is of limited utility. Frontal or subcortical contributions to control may vary as a function of the abstraction of task set representations.

Our results also confirm the promising role for multivariate techniques in neuroimaging, especially in parsing the neural substrates of executive functions, which are notoriously heterogeneous and variably assessed. Because of its inherent sensitivity in detecting intercorrelated systems among or within a mixture of voxel-wise functional distributions, MVPA is ideally suited to probe the features of frontostriatal function during cognitive control. This study did not focus on neural representation of rules or rule decoding, which has been addressed in a number of studies previously (see Reverberi et al., 2012b; Badre & D'Esposito, 2009; Badre, 2008; Koechlin & Summerfield, 2007), but rather addressed the question of control over these. Our observation that frontal or BG networks participate differentially in both switch and repeat trials, depending on the abstraction of the rules that govern behavior, represents a parsimonious solution to the problem of control. Our findings are consistent with the proposed role of frontal regions in the timely and adaptive coordination of behavior, while also providing evidence that the BG play a parallel coordinating role for hierarchically simpler behaviors.

The current study highlights the complementary contributions of the BG and the frontal lobes to cognitive control. Efforts to distil a meaningful neural basis for control from the complex picture that arises from a burgeoning literature characterized by differences between paradigms and heterogeneity of findings have also focused on rule abstraction, yet a role for the BG has hitherto not been proposed. For example, Kim et al. (2011) directly contrasted three types of switches that varied according to rule abstraction, mirroring our defined concrete and abstract conditions in what they correspondingly termed stimulus and context switches, as well as an intermediate response switch condition. Distinct cortical regions emerged in a rostrocaudal gradient depending on rule abstraction, with concrete switches associated with posterior prefrontal regions and more abstract switches with more anterior regions. Frontal, parietal, thalamic, and cerebellar activations were observed during concrete switches, but no activation emerged in the striatum. Another study placed the origins of an abstraction-independent control signal in the superior parietal lobule (Chiu & Yantis, 2009). In a meta-analysis of 36 univariate studies and using an activation likelihood estimation, Kim et al. (2012) reported that two regions, the inferior frontal junction and posterior parietal cortex, represent the common denominators of all types of switch but found no evidence for activation in the striatum, even during concrete stimulus switches. The robust and dissociable caudate activation during concrete switches we report here indicates that this region is critical to understanding lower levels of the control hierarchy.

The highlighted contribution of the caudate nucleus during concrete switches compared to repeats is consistent also with the striking observations from a study in which rule abstraction was elegantly recapitulated as a distinction between intention (between categorical response rules) and attention (between concrete stimuli; Rushworth, Hadland, et al., 2002). In that study, although a frontal cortical region during a concrete switch was identified by mass univariate fMRI, disrupting its function with repetitive TMS did not impair performance, calling into question the functional significance of this activation altogether. Only switching intention (or abstract rules)

was associated with BOLD signal changes in rostrocaudal cingulate and pre-SMA/SMA and correspondingly behaviorally impaired by repetitive TMS application.

Notably, converging evidence from work in attentional shifting between stimuli and abstract rules is consistent with our current findings on the role of the BG (Cools, Clark, & Robbins, 2004). This study reported significant ventral striatal activation associated with attentional object compared to rule shifts but failed to isolate activation associated with the converse contrast, rule shifting over object shifting, and revealed only lateral pFC activation when contrasting all shift and nonshift trials. On the same paradigm, patients with BG lesions exhibited object but not attentional rule-shifting deficits, whereas patients with frontal lesions were unimpaired; surprisingly, they were also unimpaired at attentional rule shifting (Cools, Ivry, & D'Esposito, 2006), in contrast to our own neuropsychological findings of abstract task-switching deficits in this group. Thus, although our paradigm and analysis differs in a number of ways, including the nature of stimuli, responses, and cognitive operations (see Kehagia et al., 2010, for comparison between attentional set shifting and task switching), there is agreement on the role of the BG in coordinating concrete, stimulus-relevant behavior, with frontal regions apparently not contributing differentially to cognitive flexibility in this context. Presumably, the procedural and analytical characteristics of the current task-switching study are better suited to doubly dissociating the frontal cortex from the BG during cognitive control, extending our own findings as well as those by Cools et al.

Finally, we note that, in our study, the weight maps of the forward model derived from the GPC classifiers (Figure 3) show a ventral to dorsal gradient that characterized a Repeat and Switch of task, respectively. The ventral to dorsal weight map distributions in the frontal lobes with abstract rules paralleled those seen in the BG with concrete rules. Task repetition highlighted ventral regions, specifically the OFC in the case of abstract rules and the nucleus accumbens for concrete rules. Conversely, switching was associated with dorsal GPC weight clustering: superior medial frontal regions for abstract rules and the caudate nucleus for concrete. Importantly, these corresponding ventral and dorsal patterns mirror wellestablished corticostriatal connections (Alexander et al., 1986), in agreement with the known structural and functional reciprocal connections between ventral (OFC and nucleus accumbens) and corresponding dorsal loops (dorsal and dorsolateral pFC and caudate nucleus). They are consistent in a general sense with increasing complexity in the cognitive operations carried out in the context of cognitive control. Our finding of chance level Switch/Repeat classification in the BG in the context of well-learned abstract rules appear discrepant from the predictions of computational and algorithmic work (e.g., Collins & Frank, 2014; Hazy, Frank, & O'Reilly, 2006), which holds that dopaminergic activity in the BG acts as a major gating signal over updating cortically subserved working memory representations or task sets, in this case. These models, which are rooted in reinforcement learning and have been extended to simulate executive tasks such as the WCST and Stroop, revolve around dopamine. However, in the context of switching between established abstract rules, subcortical dopamine may not be critical. We have shown previously that dopaminergic withdrawal had no effect on switching between such rules in PD patients (Kehagia et al., 2009). Our neuroimaging findings indicate that during abstract rule switches frontal cortical input has a significant bearing, which we have theorized elsewhere may be noradrenergically mediated (Kehagia et al., 2010).

Collectively, our methods and results indicate the need for a shift in the conceptual framework of task switching, away from its representation as an unequivocally "frontal" lobe paradigm. The dissociation presented here offers a nuanced perspective on control by linking low level, deterministic control systems to the BG (Redgrave, Prescott, & Gurney, 1999), a structure conserved in the phylogenetically oldest vertebrate, the lamprey (Stephenson-Jones, Ericsson, Robertson, & Grillner, 2012). Conversely, control systems that allow flexible response generalization to stimulus classes and hence embody abstract rules highlight the involvement of the more recently evolved frontoparietal cortex in man (Passingham, 1973).

APPENDIX

Table A1. Anatomical Definition of the Basal Ganglia, Frontal and Temporal Lobe Masks (Bilateral), Comprising Structures from the Subcortical and Cortical Harvard–Oxford Atlases (Labels and Atlas Indices in Parentheses)

Basal Ganglia	*Frontal Lobe*	*Temporal Lobe*
Left caudate (11)	Frontal pole (1)	Temporal pole (8)
Left putamen (12)	Superior frontal gyrus (3)	Superior temporal gyrus, anterior (9) and posterior (10) divisions
Left pallidum (13)	Middle frontal gyrus (4)	Middle temporal gyrus, anterior (11) and posterior (12) divisions
Left accumbens (26)	Inferior frontal gyrus, pars triangularis (5)	Middle temporal gyrus, temporo-occipital part (13)
Right caudate (50)	Inferior frontal gyrus, pars opercularis (6)	Inferior temporal gyrus, anterior (14) and posterior (15) divisions
Right putamen (51)	Precentral gyrus (7)	Inferior temporal gyrus, temporo-occipital part (16)
Right pallidum (52)	Medial frontal cortex (25)	Parahippocampal gyrus, anterior (34) and posterior (35) divisions
Right accumbens (58)	Supplementary motor cortex (26)	Fusiform cortex, anterior (37) and posterior (38) divisions
	Subcallosal cortex (27)	Temporal occipital fusiform cortex (39)
	Paracingulate gyrus (28)	Planum polare (44)
	Cingulate gyrus, anterior division (29)	Heschl's gyrus (45)
	Frontal orbital cortex (33)	Planum temporale (46)
	Frontal operculum cortex (41)	

Acknowledgments

We gratefully acknowledge L. Aksman for assistance with forward modeling, O. O'Daly and J. Ó Muircheartaigh for helpful discussion, and two anonymous reviewers for their constructive comments. This research was sponsored by a Wellcome Trust grant (076274/Z/04/Z) to T. W. R., B. J. Everitt, A. C. Roberts, and B. J. Sahakian and was completed at the University of Cambridge Behavioural and Clinical Neuroscience Institute, supported by a joint award from the Medical Research Council and the Wellcome Trust, as well as a Wellcome Trust Senior Investigator Award to T. W. R. (104631/Z/14/Z) and a Senior Fellowship award to J. B. R. (103838). A. A. K. was supported by an Isaac Newton fellowship and the Wellcome Trust (077029). R. Y. is grateful for support from the China Scholarship Council (CSC). D. W. J. is an NIHR academic clinical lecturer. O. M. D. gratefully acknowledges support from the Engineering and Physical Sciences Research Council (EPSRC grant EP/L000296/1).

Reprint requests should be sent to Angie A. Kehagia, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, London, SE5 8AF, UK, or via e-mail: angie.kehagia@gmail.com.

REFERENCES

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9,* 357–381.

Allport, A., & Wylie, G. S. (2000). Task switching, stimulus–response bindings, and negative priming. In S. Monsell & J. Driver (Eds.), *Attention & performance XVIII: Control of cognitive processes* (pp. 35–70). Cambridge, MA: MIT Press.

Aron, A. R., Monsell, S., Sahakian, B. J., & Robbins, T. W. (2004). A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain, 127,* 1561–1573.

Aron, A. R., Watkins, L., Sahakian, B. J., Monsell, S., Barker, R. A., & Robbins, T. W. (2003). Task-set switching deficits in early-stage Huntington's disease: Implications for basal ganglia function. *Journal of Cognitive Neuroscience, 15,* 629–642.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences, 12,* 193–200.

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience, 10,* 659–669.

Badre, D., & Wagner, A. D. (2006). Computational and neurobiological mechanisms underlying cognitive flexibility. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 7186–7191.

Barber, A. D., & Carter, C. S. (2005). Cognitive control involved in overcoming prepotent response tendencies and switching between tasks. *Cerebral Cortex, 15,* 899–912.

Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia, 6,* 53–60.

Brass, M., Ullsperger, M., Knoesche, T. R., von Cramon, D. Y., & Phillips, N. A. (2005). Who comes first? The role of the prefrontal and parietal cortex in cognitive control. *Journal of Cognitive Neuroscience, 17,* 1367–1375.

Brass, M., & von Cramon, D. Y. (2004). Decomposing components of task preparation with functional magnetic resonance imaging. *Journal of Cognitive Neuroscience, 16,* 609–620.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39,* 713–726.

Channon, S., Gunning, A., Frankl, J., & Robertson, M. M. (2006). Tourette's syndrome (TS): Cognitive performance in adults with uncomplicated TS. *Neuropsychology, 20,* 58–65.

Chiu, Y. C., & Yantis, S. (2009). A domain-independent source of cognitive control for task sets: Shifting spatial attention and switching categorization rules. *Journal of Neuroscience, 29,* 3930–3938.

Collins, A. G., & Frank, M. J. (2014). Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological Review, 121,* 337–366.

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W., (2001). Enhanced or impaired cognitive function in Parkinson disease as a function of dopaminergic medication and task demands. *Cerebral Cortex, 11,* 1136–1143.

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (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, T. W. (2004). Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *Journal of Neuroscience, 24,* 1129–1135.

Cools, R., Ivry, R. B., & D'Esposito, M. (2006). The human striatum is necessary for responding to changes in stimulus relevance. *Journal of Cognitive Neuroscience, 18,* 1973–1983.

Crittenden, B. M., & Duncan, J. (2014). Task difficulty manipulation reveals multiple demand activity but no frontal lobe hierarchy. *Cerebral Cortex, 24,* 532–540.

Crone, E. A., Wendelken, C., Donohue, S. E., & Bunge, S. A., (2006). Neural evidence for dissociable components of task-switching. *Cerebral Cortex, 16,* 475–486.

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31,* 968–980.

Frank, M. J., & Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: Computational analysis. *Cerebral Cortex, 22,* 509–526.

Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P., & Hirsch, J. (2008). Detection of time-varying signals in event-related fMRI designs. *Neuroimage, 43,* 509–520.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2003). *The elements of statistical learning* (2nd ed.). New York: Springer.

Haufe, S., Meinecke, F., Gorgen, K., Dahne, S., Haynes, J. D., Blankertz, B., et al. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage, 87,* 96–110.

Haynes, J. D. (2015). A primer on pattern-based approaches to fMRI: Principles, pitfalls, and perspectives. *Neuron, 87,* 257–270.

Haynes, J. D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience, 8,* 686–691.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2006). Banishing the homunculus: Making working memory work. *Neuroscience, 139,* 105–118.

Henson, R. (2007). Chapter 15—Efficient experimental design for fMRI. In K. Friston, J. Ashburner, S. Kiebel, T. Nichols, & W. Penny (Eds.), *Statistical parametric mapping* (pp. 193–210). London: Academic Press.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury Press.

Karayanidis, F., Nicholson, R., Schall, U., Meem, L., Fulham, R., & Michie, P. T. (2006). Switching between univalent task-sets in schizophrenia: ERP evidence of an anticipatory task-set reconfiguration deficit. *Clinical Neurophysiology, 117,* 2172–2190.

Keele, S. W., & Rafal, B. (2000). Deficits of attentional set in frontal patients. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 627–651). Cambridge, MA: MIT Press.

Kehagia, A. A., Barker, R. A., & Robbins, T. W. (2014). Revisiting the effects of Parkinson's disease and frontal lobe lesions on task switching: The role of rule reconfiguration. *Journal of Neuropsychology, 8,* 53–74.

Kehagia, A. A., Cools, R., Barker, R. A., & Robbins, T. W. (2009). Switching between abstract rules reflects disease severity but not dopaminergic status in Parkinson's disease. *Neuropsychologia, 47,* 1117–1127.

Kehagia, A. A., Murray, G. K., & Robbins, T. W. (2010). Learning and cognitive flexibility: Frontostriatal function and monoaminergic modulation. *Current Opinion in Neurobiology, 20,* 1–6.

Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis. *Human Brain Mapping, 33,* 130–142.

Kim, C., Johnson, N. F., Cilles, S. E., & Gold, B. T. (2011). Common and distinct mechanisms of cognitive flexibility in prefrontal cortex. *Journal of Neuroscience, 31,* 4771–4779.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences, 11,* 229–235.

Manoach, D. S., Lindgren, K. A., Cherkasova, M. V., Goff, D. C., Halpern, E. F., Intriligator, J., et al. (2002). Schizophrenic subjects show deficient inhibition but intact task switching on saccadic tasks. *Biological Psychiatry, 51,* 816–826.

Mayr, U., Diedrichsen, J., Ivry, R., & Keele, S. W. (2006). Dissociating task-set selection from task-set inhibition in the prefrontal cortex. *Journal of Cognitive Neuroscience, 18,* 14–21.

Mecklinger, A. D., von Cramon, D. Y., Springer, A., & Matthes-von Cramon, G. (1999). Executive control functions in task switching: Evidence from brain injured patients. *Journal of Clinical and Experimental Neuropsychology, 21,* 606–619.

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: Motor and cognitive circuits. *Brain Research Reviews, 31,* 236–250.

Moritz, S., Hubner, M., & Kluwe, R. (2004). Task switching and backward inhibition in obsessive-compulsive disorder. *Journal of Clinical and Experimental Neuropsychology, 26,* 677–683.

Muhammad, R., Wallis, J. D., & Miller, E. K. (2006). A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *Journal of Cognitive Neuroscience, 18,* 974–989.

Niki, H., & Watanabe, M. (1976). Prefrontal unit activity and delayed response: Relation to cue location versus direction of response. *Brain Research, 105,* 79–88.

Passingham, R. E. (1973). Anatomical differences between the neocortex of man and other primates. *Brain, Behavior and Evolution, 7,* 337–359.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage, 45(1 Suppl.),* S199–S209.

Rasmussen, C., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. Cambridge, MA: MIT Press.
Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience, 89,* 1009–1023.
Reitan, R. M., & Wolfson, D. (1994). A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review, 4,* 161–198.
Reverberi, C., Gorgen, K., & Haynes, J. D. (2012a). Compositionality of rule representations in human prefrontal cortex. *Cerebral Cortex, 22,* 1237–1246.
Reverberi, C., Gorgen, K., & Haynes, J. D. (2012b). Distributed representations of rule identity and rule order in human frontal cortex and striatum. *Journal of Neuroscience, 32,* 17420–17430.
Robbins, T. W., & Rogers, R. D. (2000). Functioning of frontostriatal anatomical "loops" in mechanisms of cognitive control. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive performance* (pp. 475–509). Cambridge, MA: MIT Press.
Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.
Rogers, R. D., Sahakian, B. J., Hodges, J. R., Polkey, C. E., Kennard, C., & Robbins, T. W. (1998). Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain, 121,* 815–842.
Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12,* 191–200.
Ruge, H., Brass, M., Koch, I., Rubin, O., Meiran, N., & von Cramon, D. Y. (2005). Advance preparation and stimulus-induced interference in cued task switching: Further insights from BOLD fMRI. *Neuropsychologia, 43,* 340–355.
Rushworth, M. F., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: A combined fMRI and TMS study. *Journal of Neurophysiology, 87,* 2577–2592.
Rushworth, M. F., Passingham, R. E., & Nobre, A. C. (2002). Components of switching intentional set. *Journal of Cognitive Neuroscience, 14,* 1139–1150.
Schneider, D. W., & Logan, G. D. (2005). Modeling task switching without switching tasks: A short-term priming account of explicitly cued performance. *Journal of Experimental Psychology: General, 134,* 343–367.
Schrouff, J., Kusse, C., Wehenkel, L., Maquet, P., & Phillips, C. (2012). Decoding semi-constrained brain activity from fMRI using support vector machines and Gaussian processes. *PLoS One, 7,* e35860.
Sigala, N., Kusunoki, M., Nimmo-Smith, I., Gaffan, D., & Duncan, J. (2008). Hierarchical coding for sequential task events in the monkey prefrontal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 105,* 11969–11974.
Simon, J. R. (1969). Reactions towards the source of stimulation. *Journal of Experimental Psychology, 81,* 174–176.
Slagter, H. A., Weissman, D. H., Giesbrecht, B., Kenemans, J. L., Mangun, G. R., Kok, A., et al. (2006). Brain regions activated by endogenous preparatory set shifting as revealed by fMRI. *Cognitive and Affective Behavioral Neuroscience, 6,* 175–189.
Stephenson-Jones, M., Ericsson, J., Robertson, B., & Grillner, S. (2012). Evolution of the basal ganglia: Dual-output pathways conserved throughout vertebrate phylogeny. *Journal of Comparative Neurology, 520,* 2957–2973.
Stoet, G., & Snyder, L. H. (2004). Single neurons in posterior parietal cortex of monkeys encode cognitive set. *Neuron, 42,* 1003–1012.
Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.
Stuss, D. T., Eskes, G. A., & Foster, J. K. (1994). Experimental neuropsychological studies of frontal lobe functions. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 149–185). Amsterdam: Elsevier.
Wiener, N. (1949). *Cybernetics: Or control and communication in the animal and the machine*. New York (John Wiley & Sons) and Paris (Hermann et Cie): The Technology Press.
Woodward, T. S., Ruff, C. C., & Ngan, E. T., (2006). Short- and long-term changes in anterior cingulate activation during resolution of task-set competition. *Brain Research, 1068,* 161–169.
Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience, 31,* 14592–14599.
Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between-task competition and cognitive control in task switching. *Journal of Neuroscience, 26,* 1429–1438.