Neural Substrates for Reversing Stimulus–Outcome and Stimulus–Response Associations

Gui Xue,[1,2] **Dara G. Ghahremani,**[2] **and Russell A. Poldrack**[2,3,4]

[1]Foundation for Psychocultural Research–University of California, Los Angeles Center for Culture, Brain and Development, [2]Department of Psychology, [3]Department of Psychiatry and Biobehavioral Sciences, and [4]Brain Research Institute, University of California, Los Angeles, Los Angeles, California 90095

Adaptive goal-directed actions require the ability to quickly relearn behaviors in a changing environment, yet how the brain supports this ability is barely understood. Using functional magnetic resonance imaging and a novel reversal learning paradigm, the present study examined the neural mechanisms associated with reversal learning for outcomes versus motor responses. Participants were extensively trained to classify novel visual symbols (Japanese Hiraganas) into two arbitrary classes ("male" or "female"), in which subjects could acquire both stimulus–outcome associations and stimulus–response associations. They were then required to relearn either the outcome or the motor response associated with the symbols, or both. The results revealed that during reversal learning, a network including anterior cingulate, posterior inferior frontal, and parietal regions showed extended activation for all types of reversal trials, whereas their activation decreased quickly for trials not involving reversal, suggesting their role in domain–general interference resolution. The later increase of right ventral lateral prefrontal cortex and caudate for reversal of stimulus–outcome associations suggests their importance in outcome reversal learning in the face of interference.

Key words: fMRI; cognitive control; reversal learning; interference resolution; stimulus–response association, stimulus–outcome association

Introduction

Adaptive goal-directed actions require the ability to overcome old habitual behaviors to learn new behaviors in changing environments ("reversal learning") (Miller and Cohen, 2001). Although people have the capacity to quickly switch their responses, sometimes after a single learning event, the expression of new behaviors is not stable, and it often takes time and effort to overcome prepotent behaviors and to learn the new behaviors to a satisfactory level of automaticity (Shiu and Chan, 2006). Despite its tremendous significance for adaptive behavior, the neural mechanisms involved in reversal learning of overlearned skills are not well understood.

Reversal learning has been widely used to examine how participants respond to the change of stimulus–reward or stimulus–response contingencies, in which participants must override established associations and learn new ones according to feedback (Iversen and Mishkin, 1970; Dias et al., 1996; O'Doherty et al., 2001; Cools et al., 2002; Budhani et al., 2007). Results from human lesion (Hornak et al., 2004), animal lesion (Iversen and Mishkin, 1970; Dias et al., 1996), and functional imaging

(O'Doherty et al., 2001, 2003; Cools et al., 2002; Remijnse et al., 2005) research have generally emphasized the role of ventrolateral and lateral orbital prefrontal cortex, as well as the basal ganglia, in reversal learning.

There are three aspects of typical reversal learning studies that differ from the reversal learning paradigm we investigate here. First, whereas most previous studies have used paradigms in which the subject chooses one of two stimuli and is rewarded for choosing the correct stimulus, we use a paradigm in which subjects must also learn one of two possible responses for each stimulus. This seemingly subtle change in the task allows us to separate reversal of stimulus–outcome associations from reversal of stimulus–response associations, which is not possible in the standard paradigm. Second, most previous studies of reversal learning have examined reversal after a relatively small amount of practice with a particular association. As a result, they likely involve a minimal level of conflict processing and interference resolution compared with real-world habits. By extensively training participants before reversal, we were able to examine how participants overcome habitual behaviors and to assess the role of interference resolution in reversal learning. Finally, many of these studies have adopted a serial reversal or switching paradigm, focusing their analyses on the comparison of the first successful reversal/switched trial (or last pre-reversal error) with non-reversal/switched trials. In the present study, we imaged several repetitions after reversal, which allows us to explore how the brain gradually acquires the new behavior in the face of interference from existing habits. We found that the ventrolateral prefrontal cortex (VLPFC) and caudate nucleus, which have often been associated with reversal learning, are specifically engaged by the
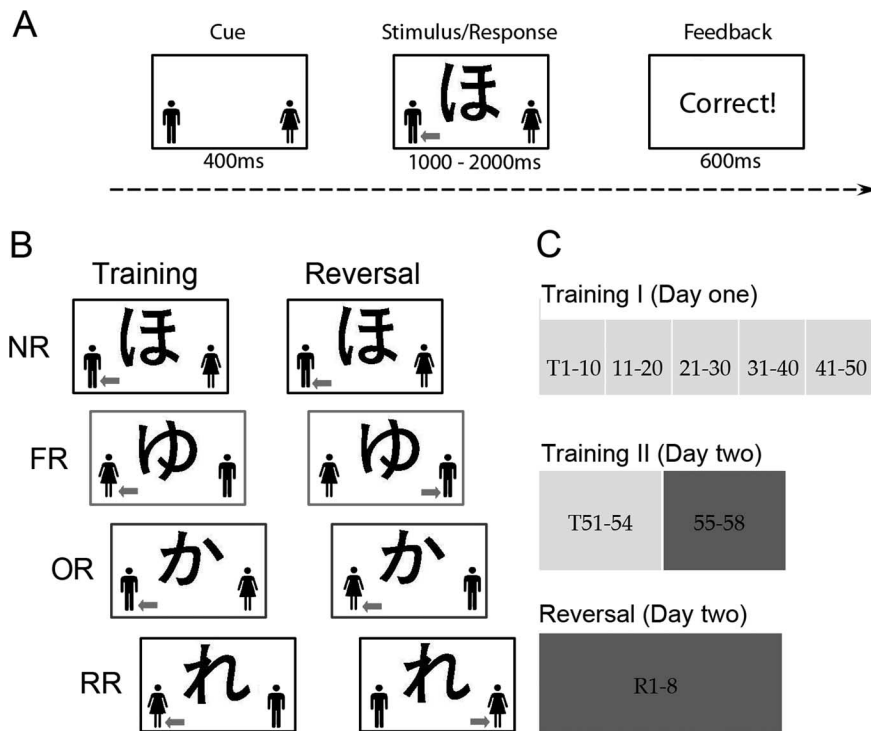
Figure 1. Schematic of experimental design. *A*, The adapted classification learning task. Participant made a manual response and received informative feedback. Note that the gender position could change from trial to trial but always fixed to a given stimulus. In this way, participants learned both the stimulus–outcome ("male" or "female") and stimulus–response ("left" or "right" key press) association. *B*, The experimental conditions. Each stimulus was associated with a gender; half the stimuli were shown with the male symbol on the left and half on the right. The gray arrow indicates the correct response during training and reversal, which is shown to help demonstrate the reversal conditions but was not printed on the screen in the actual experiment. For NR trials, nothing was changed from training to reversal. For FR trials, only the outcome was changed during reversal: participants were required to relearn the outcome as well as to switch their response. For OR trials, both the outcome and the associated gender symbols changed; participants thus only needed to relearn the outcome without switching their response. For RR trials, only the associated gender symbol was changed; participants only needed to switch their response without relearning the outcome. *C*, The training procedure. Each isolated block represents one learning block, and the numbers inside represent the number of repetitions. The light gray represents the behavioral sessions, whereas the dark gray indicates sessions in which functional scans were acquired. T, Training; R, reversal learning.

gender labels (with male and female symbols on each side) for which the spatial positions on the display were fixed for a given stimulus across training repetitions (thus requiring the same key response), but this positioning varied across stimuli; thus, for some stimuli the response "male" was always associated with the left key, whereas for others it was consistently associated with the right key. In this way, although participants still learned both the stimulus–outcome association and stimulus–response association, as in the typical classification task, we could, in the reversal learning stage, selectively change the associated outcome or gender label position to impose different types of reversal learning (see below).

The structure of a single classification learning trial is depicted in Figure 1*A*. During each trial, the gender labels (cartoon figures of a male and female) appeared on the lower left and right parts of the screen for 400 ms before the Japanese hiragana appeared in the center. Both the gender labels and Hiragana stayed on the screen until a response (left or right key corresponding to left or right index finger) was made. Participants received feedback in the form of the word "correct" or "wrong" presented in the center of the screen for 600 ms. If no response was made within the response window (to be detailed below), "no response" was presented.

Items from training were split into four conditions during the reversal phase (Fig. 1*B*). In the "no-reversal" (NR) condition, both the correct outcome and required motoric response (hereafter, "response") remained the same. In the "full reversal" (FR) condition, both the correct outcome and response changed, requiring the participants to relearn both the outcome of the stimuli (at a conceptual level) and the response. In the "outcome reversal" (OR) condition, both the correct outcome and the gender label positions were changed, such that participants only needed to relearn the outcome without switching their response. In the "response reversal" (RR) condition, the gender label positions were changed but the correct outcome remained constant; participants only needed to relearn their response (i.e., left or right key) because the outcome remained the same.

Prescan behavioral training. The overall experiment consisted of three stages, training I, training II, and reversal learning (Fig. 1*C*). One day before the scan, participants were extensively trained to become accurate and fast at making the classification (i.e., training I). Before training, participants were instructed to learn the label (i.e., outcome) for each stimulus based on feedback and that their goal should be to achieve 90% correct or higher. They were also explicitly instructed not to apply any rule because the classifications were arbitrary. Particularly, they were discouraged from associating specific visual features of the characters with male or female categories. Thus, the "classification learning" task in our study involved arbitrary associative learning, and was different from the usual category learning in which the equivalence classes for each category label share some common simple structure and subjects develop a representation of each class. The training included five sessions consisting of four mini-blocks each. Within each mini-block, 8 of the 32 characters repeated 10 times. The trials were presented in mini-blocks to help control the inter-repetition interval (IRI) for each stimulus, a variable that has been shown to influence learning difficulty as well as retention of learning (Karpicke and Roediger, 2007). This design had an average IRI of eight trials, which our pilot data suggested would produce an appropriate level of difficulty for learning. To prevent participants from developing rules based on the given set of stimuli, the same eight stimuli in one

need to override preexisting stimulus–outcome associations rather than stimulus–response associations.

Materials and Methods

Participants. Seventeen healthy, native English-speaking participants took part in this study (8 males, 9 females; average age 22.7 years, range 19–28). All participants had normal or corrected-to-normal vision and were right-handed as judged by the Edinburgh Handedness Inventory (Oldfield, 1971). None of them knew any major Asian language, including Japanese, Chinese, and Korean. They were free of neurological or psychiatric history and gave informed consent according to a procedure approved by the University of California, Los Angeles (UCLA) Human Subject Committee. One additional subject was scanned but removed from the analysis due to exceptionally poor behavioral performance in the scanner (accuracy <40%).

The reversal learning task. The present study used an adapted classification learning task, in which participants were asked to learn by trial-and-error whether each of the 32 novel Japanese Hiragana represented a male or a female name (Fig. 1). In a typical classification learning task (Poldrack et al., 2001), two conceptual classes (which we refer to here as "outcomes") are fixed to left and right button responses (e.g., outcome A-left key, outcome B-right key), and participants are required to learn both the stimulus–outcome association and the stimulus–response association. As a result, a shift in stimulus–outcome association (at a cognitive level) is coupled with a switch in motoric response (i.e., the alternative button press response). To dissociate them, the present study used

mini-block did not appear together again in the next block. As the training progressed and participants became more fluent at this task, the response window gradually decreased from 2 s to 1 s, and the interstimulus interval decreased from 1 s to 0.5 s before the next trials started.

Behavior in the functional magnetic resonance imaging session. The same task was used during the scanning session. Trial sequences were jittered (by adding null events after each trial; mean 1.4 s, range 0.5–5 s) and optimized with OPTSEQ (http://surfer.nmr.mgh.harvard.edu/optseq/) (Dale, 1999). We carefully selected sequences in which the IRIs (in terms of both time and trials between stimulus repetitions) and their SDs for the four conditions were matched. The response-time window was set to 1.5 s for all conditions. Participants made their manual responses via a magnetic resonance imaging (MRI)-compatible button box and responses were recorded by the computer. Stimulus presentation and response collection was programmed using Matlab (Mathworks) and the Psychtoolbox (www.psychtoolbox.org) on an IBM laptop.

The scanning session was divided into two stages: training II and reversal learning. During training II, participants received eight additional repetitions of the training trials divided across two runs. Because of time limitations, one run was presented during the magnetization-prepared rapid-acquisition gradient echo (MPRAGE) anatomical acquisition and another during the first functional MRI (fMRI) scan. In each run, there were four mini-blocks of eight stimuli, each repeated four times. During the two reversal learning scans, the stimulus–outcome and/or stimulus–response associations were changed for some of the trials as specified above. Unlike the training II scan, each reversal learning scan included two mini-blocks of eight stimuli (two from each condition), each repeated eight times. This allowed us to examine the time course of reversal learning within one scan without being confounded by the time factors. Both the training II scans and the reversal learning scans included 128 trials which lasted 500 s in total. The stimuli assigned to each condition and each scan were fully counterbalanced across participants. Participants were not warned about the reversals at any point before the reversal learning phase.

Postscan memory test and debriefing. After the scanning session, participants were asked to recall the outcome and response associated with each stimulus on a paper and pencil test. They were clearly asked to make their response based on the last correct response (i.e., the post-reversal response). In the outcome memory test, all 32 symbols were present on one sheet of paper in a randomized order, and they were asked to indicate whether each stimulus was male or female by putting "M" or "F" on the top-right corner of each symbol, followed by a number (1–5) to indicate their confidence, with 1 indicating "not sure at all" and 5 indicating "absolutely sure." In the response memory test, another sheet of paper with the same 32 symbols was presented, and participants were asked to indicate whether each stimulus was associated with the left or right key response by putting "L" or "R" on the top-right corner of each symbol, followed by a number (1–5) to indicate their confidence. There was no time limitation on the test and participants were free to answer the questions according to any order. In general, participants finished the task in 10 min.

MRI data acquisition. Imaging data were collected using a 3T Siemens Allegra MRI scanner at the UCLA Ahmanson-Lovelace Brain Mapping Center. For each run, 250 functional T2*-weighted echoplanar images (EPIs) were acquired using an oblique axial slice prescription with the following parameters: slice thickness, 4 mm, 33 slices; repetition time (TR), 2 s; echo time (TE), 30 ms; flip angle, 90°; matrix, 64 × 64; field of view (FOV), 200 mm. A T2-weighted matched-bandwidth high-resolution anatomical scan was acquired to aid coregistration. This scan has the same imaging bandwidth and slice prescription as the functional images (which results in matched distortions) but with a higher in-plane resolution (1 mm × 1 mm). Additionally, a high-resolution structure image (MPRAGE) was acquired. The parameters for MPRAGE were: TR, 2.3 s; TE, 2.1 ms; FOV, 256 mm; matrix, 192 × 192; sagittal plane, slice thickness, 1 mm, 160 slices.

Imaging data preprocessing and statistical analysis. Initial analysis was performed using tools from the FMRIB software library (FSL) (www.fmrib.ox.ac.uk/fsl) Version 3.3. The first two volumes were discarded to allow for T1 equilibrium effects. The remaining images were then re-aligned to compensate for small head movements (Jenkinson and Smith, 2001). Translational movement parameters never exceeded 1 voxel in any direction for any subject or session. All images were de-noised using MELODIC independent components analysis within FSL (Tohka et al., 2008). Data were spatially smoothed using a 5 mm full-width-half-maximum Gaussian kernel. The data were filtered in the temporal domain using a nonlinear high-pass filter with a 66 s cutoff. A three-step registration procedure was used whereby EPIs were first registered to the matched-bandwidth high-resolution scan, then to the MPRAGE structural image, and finally into standard (Montreal Neurological Institute) space, using affine transformations (Jenkinson and Smith, 2001).

The data were modeled at the first level using a general linear model within the FILM module of FSL. Event onsets were modeled at the time of the gender label presentations. These event onsets were convolved with canonical hemodynamic response function (double-gamma) to generate the regressors used in the general linear model. Temporal derivatives were included as covariates of no interest to improve statistical sensitivity. Null events were not explicitly modeled and, therefore, constituted an implicit baseline. For training II data, each condition was separately modeled to examine whether there was significant difference between the conditions before reversal. The linear contrast, [1 1 1 1], was used to produce an overall activation map representing the brain regions involved in the task. For the reversal learning data, the first post-reversal trial for each stimulus was modeled as a nuisance variable, separately for each condition due to the response uncertainty. The remaining seven repetitions were divided into Bin1 (repetitions 2–4) and Bin2 (repetitions 5–8), according to our initial exploration of the learning curve (see Results, Behavioral results), to examine the time course of reversal learning. Only correct responses were included in this analysis. The incorrect trials were modeled as nuisance variables separately for NR trials and all reversal learning trials. Each reversal learning condition versus baseline contrast and direct comparisons between conditions were defined for each subject and each run.

For reversal learning data, a higher-level analysis was used to combine contrasts across runs for each subject using FLAME (FMRIB's Local Analysis of Mixed Effect) stage 1 only (Beckmann et al., 2003; Woolrich et al., 2004). Runs were treated as a random effect, with the between-run variance estimate pooled across subjects. The mean contrast images (i.e., a linear combination of parameter estimate images reflecting a particular statistical contrast) across runs were then inputted into a random-effects model for group results using FLAME stage 1 only as well. Unless otherwise noted, group images were thresholded using cluster-corrected statistics, with a height threshold of $z > 2.0$ and a cluster probability of $p < 0.05$, corrected for whole-brain multiple comparisons (using Gaussian random field theory).

Regions of interest analysis. Regions showing significant reversal effects were defined functionally based on voxelwise statistical maps (all reversal learning conditions vs NR) by growing a 6 mm diameter sphere around the local maxima in each cluster. Regions specific to OR were defined by the contrast of OR − RR. Percentage signal change was calculated based on the peak height of the hemodynamic response versus the baseline level of activity [J. Mumford (2007) A Guide to Calculating Percent Change with Featquery. Unpublished Tech Report available at http://mumford.bol.ucla.edu/perchange_guide.pdf].

Results

Behavioral results for prescan training (training I)

Although participants underwent the same training conditions for all stimuli during training I, we analyzed the results according to their subsequent reversal condition assignments to ensure that no systematic differences appeared across the four reversal conditions during training. Group-averaged response times (RTs) and performance accuracy were calculated for each repetition (collapsed across eight trials) and each condition (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). To achieve appropriate statistical power, we collapsed data for each block (i.e., 10 repetitions) and entered only the first and last block into a condition-by-block ANOVA. Training significantly
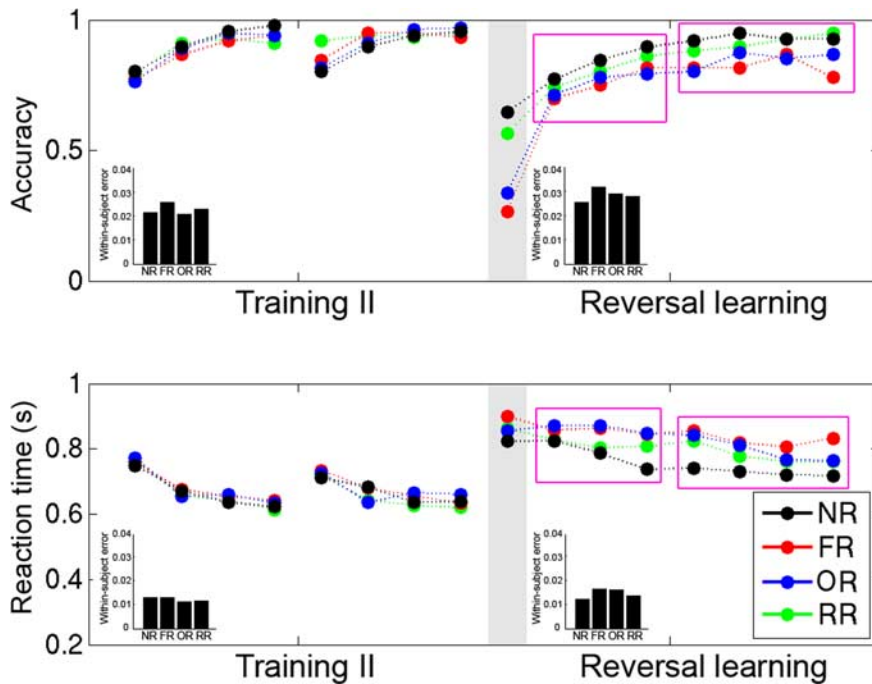
Figure 2. Behavioral performance during training II and reversal learning, separated for each condition and each repetition. Note that there were no differences between conditions during training II. The purple rectangles divide the last several reversal repetitions into two bins, Bin1 and Bin2, representing the early and late stage of reversal learning, respectively. Inset, Bar graphs represent the within-subject error for each condition.

increased accuracy ($F_{(1,16)} = 90.40$, $p < 0.0001$) and shortened RTs ($F_{(1,16)} = 77.39$, $p < 0.0001$), but there were no differences across conditions (accuracy, $F_{(3,48)} = 1.89$, $p = 0.14$; RT, $F_{(3,48)} = 0.018$, $p = 0.90$). The slight decrease in performance at the beginning of each block was attributable to the longer cross-block IRI than the within-block IRI (i.e., ~240 items vs 8 items.) On average, accuracy was >90%. Only two participants required one additional block of training to bring their accuracy up to 90% or greater.

Behavioral results of training II
On day 2, participants went through additional training (i.e., training II) during one anatomical scan and one functional scan (four repetitions in each scan). This training further improved participants' performance, as reflected by the significant accuracy increase ($F_{(7,112)} = 32.39$, $p < 0.0001$) and RT decrease ($F_{(7,112)} = 29.12$, $p < 0.0001$) across repetitions (Fig. 2). Focusing on the last pre-reversal trial, the average accuracy was ~95% for all conditions, and the RTs were approximately 620 ms, suggesting that participants had sufficiently learned the task. Moreover, there were no significant differences for RT across conditions ($F_{(3,48)} = 0.37$, $p = 0.77$), or accuracy ($F_{(3,48)} = 0.475$, $p = 0.71$), suggesting that participants had been equally trained on stimuli that were subsequently assigned to different reversal learning conditions.

Behavioral results of reversal learning
The first reversal trial
Although there were no significant differences across conditions in the last training repetition, such differences appeared in the first reversal trials ($F_{(3,36)} = 138.65$, $p < 0.0001$) (Fig. 2). Planned paired t tests indicated that the accuracy for RR and NR was higher than that for FR and OR (all p values <0.001). The accuracy for RR was not different from that for NR ($t_{(16)} = 1.37$, $p = 0.186$) and that for OR was not different from that FR ($t_{(16)} =$

1.25, $p = 0.23$), suggesting that participants relied on memory of the outcome rather than the motor response to guide their categorization. The accuracy for FR and OR did not approach zero (26.5% and 33.8% for FR and OR, respectively), and the accuracy for NR and RR was not perfect (65%), suggesting that participants may have moved into an "exploration" mode (attempting to predict reversals) after committing the first few reversal errors. There was no significant effect on reaction time ($F_{(12,36)} = 1.001$, $p = 0.40$); however, the RT should be treated cautiously because of the limited number of correct trials. Four participants were excluded in the RT analysis because of zero accuracy in one or two conditions.

Behavioral changes with reversal learning
Although previous reversal learning studies focused their analyses on the first correct postreversal trial as a single measure of reversal learning, the present study focused on how participants gradually overcame the interference and relearned the concepts and/or motoric response over time. Ideally, we would have examined the learning curve at each repetition point, but we did not have enough statistical power for this because there were only four trials (including incorrect trials) for each time point per condition. As a result, we divided the reversal learning period into two stages to improve the power at the cost of temporal resolution. This division was determined based on the examination of Figure 2, which suggests two different stages: early reversal learning (i.e., Bin1, repetitions 2–4) and late reversal learning (i.e., Bin2, repetitions 5–8). The accuracy improved quickly in Bin1 ($F_{(2,32)} = 20.4$, $p < 0.0001$), whereas it remained constant in Bin2 ($F_{(3,48)} = 1.64$, $p = 0.19$). In contrast, the RT for the three reversal learning conditions only improved in Bin2 ($F_{(3,48)} = 12.77$, $p < 0.001$), but not in Bin1 ($F_{(2,32)} = 1.42$, $p = 0.26$). The RT decrease for NR occurred in Bin1 ($F_{(2,32)} = 3.94$, $p = 0.029$), but not in Bin2 ($F_{(3,48)} = 0.18$). We further examined whether outcome reversal and response reversal were equally difficult. Planned comparisons suggested that the accuracy of RR and OR did not differ in Bin1 and was only marginally different in Bin2 ($t_{(16)} = 1.87$, $p = 0.08$). Also, although RR was faster than OR in Bin1 ($t_{(16)} = 3.10$, $p = 0.007$), this difference diminished in Bin2 ($t_{(16)} = 1.22$, $p = 0.24$). As a result, we focused our comparison between the two types of reversal learning at the later reversal learning stage, and the differences we found should be less affected by task difficulty.

Postscan memory test
When asked to explicitly recall the relearned outcome and response associated with each stimulus, participants had worse outcome memory for items for which outcome had been reversed (i.e., NR and RR) than for those for which outcome had not been reversed (i.e., OR and FR) ($F_{(3,48)} = 5.35$, $p = 0.003$; all p values for paired test <0.03), but there were no differences between NR and RR ($p = 0.41$) or OR and FR ($p = 0.89$) (supplemental Fig. S2, available at www.jneurosci.org as supplemental material). Similarly, participants had worse response memory for items for

which response had been reversed (i.e., NR and OR) than for those for which it had not (i.e., FR and RR) ($F_{(3,48)} = 13.728$, $p < 0.0001$, all p values for paired test <0.04), but there were no differences between OR and NR ($p = 0.13$) or FR and RR ($p = 0.70$). No significant differences across conditions were found for outcome memory confidence ($F_{(3,48)} = 1.87$, $p = 0.15$) or response memory confidence ($F_{(3,48)} = 1.12$, $p = 0.35$).

Post-test debriefing indicated that only three subjects noticed stimulus–response associations during training and reversal learning, but none of them intentionally ignored the outcome information in either stage. Thus, the stimulus–response association memory was implicitly and incidentally acquired during learning although it was explicitly probed in our memory test. Although subjects theoretically could have relied on outcome memory and stimulus–response association memory to perform the RR and OR tasks, respectively, during reversal, this strategy would have been inefficient and infeasible for several reasons. First, holding both an outcome and a response in mind for OR and RR items would have been an inefficient use of memory, unnecessarily increasing cognitive load. Second, subjects were not informed of the different reversal types before the experiment began; thus, they would have had to detect these differences before they could apply different item-dependent strategies. It is unrealistic to expect that subjects would have done this. Third, even if they could detect different reversal types and develop different strategies, they would have needed to know the exact manner in which a given item had been reversed. This also would have been extremely difficult given that all items were presented in a mixed order.

In summary, our behavioral results indicated that our manipulations allowed us to examine two different components of reversal learning, one that emphasized response reversal learning and another which emphasized outcome reversal learning. In the following analysis, we examined their corresponding neural mechanisms.

fMRI results
Brain regions involved in task performance
The imaging results during training II are shown in supplemental Figure S3 and supplemental Table S1, available at www.jneurosci.org as supplemental material. Because there was no difference among the four conditions at the learning stage, data were collapsed across conditions. A large bilateral frontal-striatum-thalamus-cerebellum network was involved in performing the task, including ACC/PreSMA (anterior cingulate cortex–presupplementary motor area), bilateral precentral gyri extending down to posterior inferior frontal gyrus (pIFC), right middle frontal gyrus, and subcortical regions, such as bilateral putamen, thalamus, and cerebellum. In addition, the bilateral inferior parietal lobules and visual cortex, including bilateral fusiform, inferior/middle occipital gyri, and calcarine cortex, were also active.

Common neural network for all reversal learning conditions
The first trial at the reversal stage for each stimulus was removed from this analysis to exclude activation associated with the initial
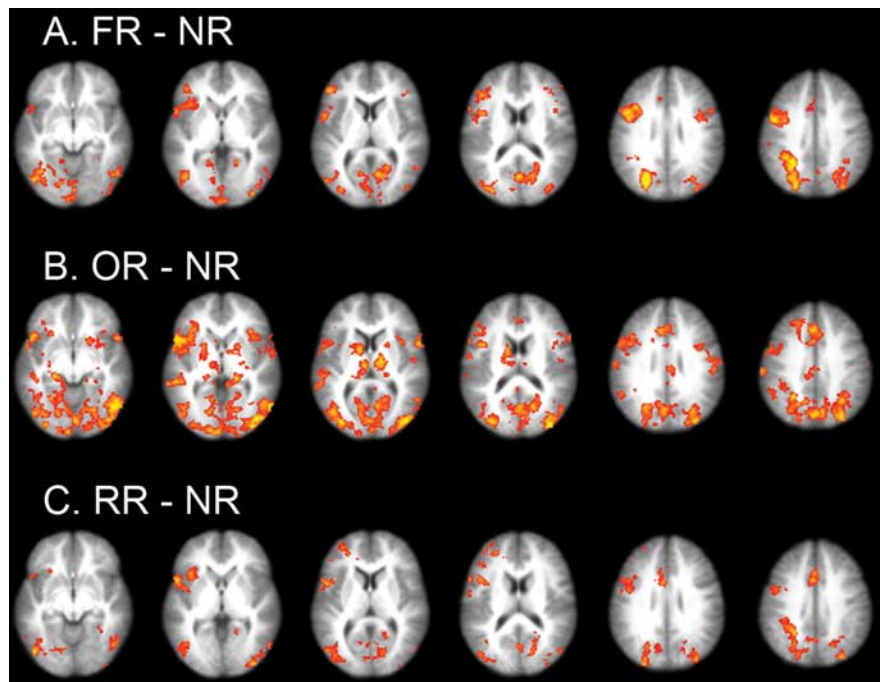

Figure 3. *A–C*, Thresholded statistical map for comparison of FR minus NR (*A*), OR minus NR (*B*), and RR minus NR (*C*) ($Z > 2.0$; $p < 0.05$, corrected for multiple comparisons at the whole-brain level).

reversal error signal and with the "prediction" of reversal that some participants attempted. We examined the reversal learning effect (reversal learning trials vs NR) for each bin and each condition separately. Only correct trials were included in this analysis to (1) examine the basis of successful reversal learning, and (2) exclude confounding factors, such as error signal processing.

For Bin1, there was no significant reversal effect (reversal vs NR) for either FR, OR, or RR at the standard threshold. This likely reflects the fact that subjects were in an exploratory mode to predict/guess which items were reversed and which were not, as reflected in the behavioral data. Because only half of the trials were reversed along a single dimension (i.e., response or outcome), the difficulty in differentiating the reversal and NR trials would have led to a general increase in response time and increase in neural activity for both NR and reversal trials.

In Bin2, the correct response had been established, but participants still needed to overcome the previously learned associations. All three reversal learning conditions elicited similar activation in the frontal-parietal network, including ACC/PreSMA, left precentral gyrus extending to left pIFC, VLPFC extending to the insula, right pIFC (although activation in this region for RR appeared at a slightly decreased threshold, $p < 0.001$, uncorrected) (supplemental Table S4, available at www.jneurosci.org as supplemental material), and bilateral superior parietal lobule (SPL) (Fig. 3; supplemental Tables S2, S3, S4, available at www.jneurosci.org as supplemental material). The common network was confirmed by the conjunction analysis across reversal conditions using the procedure suggested by Nichols et al. (2005) (supplemental Fig. S4, available at www.jneurosci.org as supplemental material). Areas responsible for visual processing, including fusiform, calcarine, and inferior and middle occipital gyri, were also activated, probably because of the increased attentional demands and top-down modulation during reversal learning. These activations will not be discussed further.
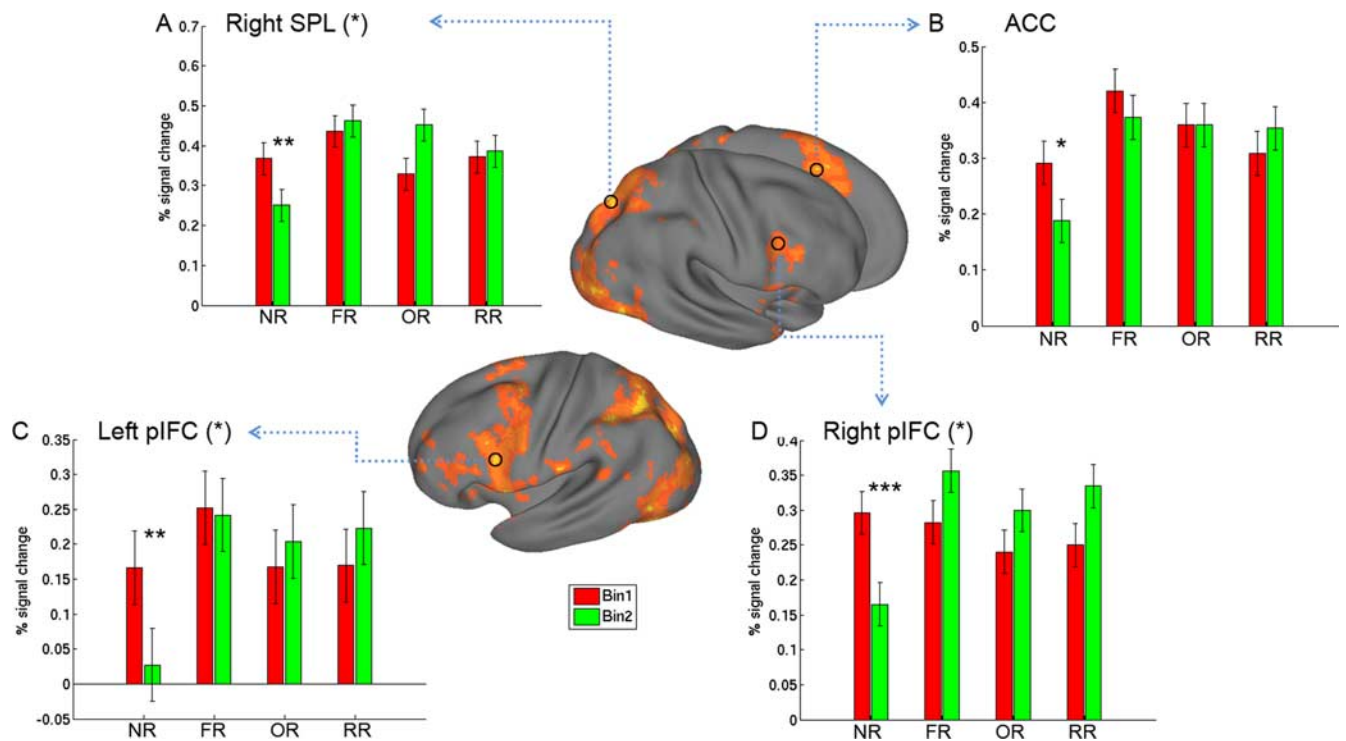
Figure 4. *A–D,* The ACC-pIFC-parietal network showed significant decreases for NR from Bin1 to Bin2 but remained stable for all reversal learning conditions. Surface rendering for the lateral view (middle) was created by mapping the overall reversal learning effect (all reversal conditions vs NR) into a population-averaged surface atlas. The small circles mark the location of the functionally defined regions of interest (ROIs). Percentage signal change is plotted for each condition and each bin in these ROIs (see Materials and Methods). Asterisks in parentheses indicate the significant effect of the bin-by-condition interaction. Asterisks near the bar indicate significant difference between Bin1 and Bin2 for that condition. Error bars represent the within-subject error. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

The striatum and VLPFC are uniquely involved in outcome reversal learning

Also for Bin2, OR vs NR elicited additional activation in the right VLPFC [Brodmann's area 44 (BA44), according to the probabilistic cytoarchitectonic map (Amunts et al., 1999)] that extended to the insula, left dorsal striatum, right ventral striatum, and bilateral thalamus (Fig. 3; supplemental Table S3, available at www.jneurosci.org as supplemental material). We directly compared OR and RR trials to further examine the different mechanisms for outcome and response reversal learning. The results indicated that OR showed stronger activation than RR in the right dorsal and ventral striatum, as well as in the right VLPFC, although the difference in VLPFC did not reach whole-brain corrected significance ($p < 0.001$, uncorrected) (see Fig. 5; supplemental Table S5, available at www.jneurosci.org as supplemental material). No regions showed more activation to RR than to OR.

Neural changes associated with reversal learning
The second major goal of the present study was to examine the neural changes associated with reversal learning. We plotted the percentage blood oxygen level-dependent signal change separately for Bin1 and Bin2 in regions showing reversal effects in Bin2. This analysis revealed two different patterns across regions, suggesting a functional dissociation within this network.

The ACC-pIFC-SPL network showed sustained activation during reversal learning
There was significant decrease from Bin1 to Bin2 for NR in ACC ($t_{(16)} = 2.44$, $p = 0.026$), bilateral pIFC (left, $t_{(16)} = 3.07$, $p = 0.007$; right, $t_{(16)} = 5.18$, $p < 0.001$), and right SPL ($t_{(16)} = 3.61$, $p = 0.002$), whereas their activations remained stable for all of the reversal learning conditions (all p values >0.20) (Fig. 4). The

bin-by-condition interaction was significant for left pIFC ($F_{(3,48)} = 3.08$, $p = 0.036$) and right pIFC ($F_{(3,48)} = 3.44$, $p = 0.026$), and marginally significant for right SPL ($F_{(3,48)} = 2.40$, $p = 0.079$), although this was not significant for ACC ($F_{(3,48)} = 1.35$, $p = 0.26$). The extended activation in ACC-pIFC-SPL network for all of the reversal learning conditions suggests that it might be involved in resolving the prolonged response and cognitive conflict imposed by the reversal conditions, as evident in behavioral data.

Right VLPFC and caudate increased for outcome reversal learning
The right VLPFC (Fig. 5A) and right caudate (Fig. 5B) showed increased activation from Bin1 to Bin2 only for OR ($t_{(16)} = 3.5$, $p = 0.003$, and $t_{(16)} = 2.3$, $p = 0.035$, respectively), but remained stable for all of the other conditions (all p values >0.15). Overall, there was a significant bin-by-condition interaction for right VLPFC ($F_{(3,48)} = 5.05$, $p = 0.004$) and a marginally significant interaction for right caudate ($F_{(3,48)} = 2.14$, $p = 0.10$). Focusing on OR and RR, there were marginally significant bin-by-condition interactions for both right VLPFC ($F_{(1,16)} = 3.30$, $p = 0.088$) and right caudate ($F_{(1,16)} = 2.76$, $p = 0.11$), suggesting that increases in activation in these regions were specific to the reversal of stimulus–outcome associations.

Discussion
Although many studies have examined cognitive control in terms of response inhibition, task set/attention switching, and reversal learning, the reversal learning of extensively trained prepotent responses or habits has been rarely studied. The present study successfully separated the stimulus–outcome and stimulus–response components in a novel associative learning task, and the results revealed both common and distinctive neural mecha-

nisms for outcome and response reversal learning. That is, whereas an ACC-pIFC-SPL network is recruited for resolving both cognitive and motoric interference, the right frontal-caudate network is specific for outcome reversal learning.

Right VLPFC and caudate support outcome reversal learning

In the OR condition, participants were required to acquire new stimulus–outcome associations while the stimulus–response association remained the same. As a result, although the previous response led to the correct performance feedback, the meaning of the response still had to be relearned, an important aspect of flexible goal-directed behavior. We found that the right VLPFC and caudate were uniquely activated for OR, and showed increased responding from Bin1 to Bin2 during reversal learning. These results suggest that the right VLPFC and caudate may be specifically involved in learning the outcome and the response–outcome contingency after reversal.

Monkey physiological studies have shown that lateral prefrontal cortex (PFC) (equivalent to human VLPFC) represents abstract categories associated with unique actions (Freedman et al., 2001, 2002), and human neuroimaging studies have consistently implicated this region in reversal learning (Cools et al., 2002; Remijnse et al., 2005). By separating the outcome and response reversal, our results extend these findings and suggest that right VLPFC

Figure 5. *A*, *B*, Neural regions specific for outcome reversal learning. Brain regions showing significantly greater activation for OR than for RR are overlaid on the group-averaged anatomical map (supplemental Table 5, available at www.jneurosci.org as supplemental material). The bar graphs show the percentage signal change in each functionally defined region of interest (see Materials and Methods). Both the right VLPFC (*A*) and the right caudate (*B*) showed (marginally) significant condition-by-bin interaction both when all conditions were included (as indicated by the symbols in the parentheses; $p = 0.004$ and 0.1, respectively) and when only OR and RR were included (as indicated by the symbol above the rectangle, $p = 0.088$ and 0.11, respectively). Asterisks near the bar indicate significant difference between Bin1 and Bin2 for that condition. Error bars represent the within-subject error. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $^+$marginally significant.

might be specifically involved in inhibiting the old outcome and response–outcome contingency (i.e., a specific form of action–outcome learning). This accords with the fact that the VLPFC has been implicated in response inhibition (Aron et al., 2004, 2007; Aron and Poldrack, 2006; Xue et al., 2008). Interestingly, other studies using linguistic material found that the left VLPFC is involved in controlled retrieval (Wagner et al., 2001) and cognitive flexibility (Badre et al., 2005; Badre and Wagner, 2006); further studies are needed to better determine the basis for lateralization of VLPFC function in cognitive flexibility.

The caudate (i.e., dorsomedial striatum) has been implicated in flexible goal-directed behavior, such as place learning and action–outcome learning (Yin et al., 2005a,b; Yin and Knowlton, 2006). Lesion or reversible inactivation of the caudate abolishes sensitivity to reward devaluation or degradation (Yin et al., 2005a), consistent with human fMRI results showing that the caudate encodes action–reward contingency (O'Doherty et al., 2004; Tricomi et al., 2004). Our data are consistent with these observations (i.e., action–outcome learning) and do not support the reward (Delgado et al., 2000; Seger and Cincotta, 2005) or salience (Zink et al., 2003) view of caudate function. The latter view cannot explain our data because only correct trials have been included and they are associated with the same positive feedback. The increased caudate activation from Bin1 to Bin2 could not reflect the salience of the stimuli or reward, which should decrease from Bin1 to Bin2.
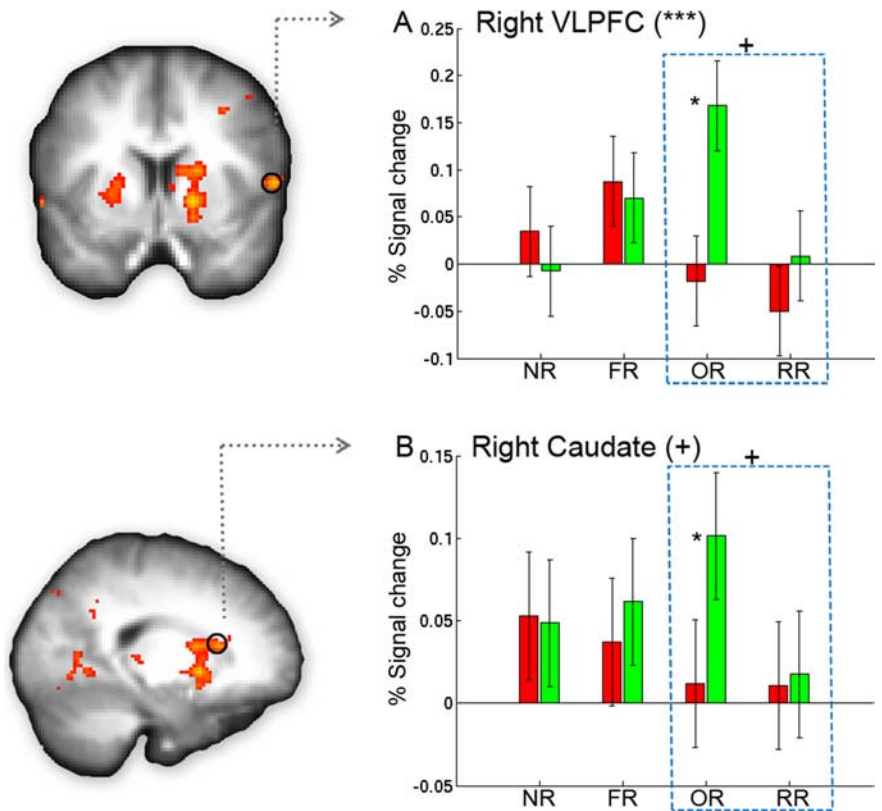
One important difference between the present study and previous reversal learning studies is that we did not observe VLPFC activation until late in the relearning period (i.e., after four repetitions), unlike other studies which showed activation in this region during the first reversal trial (or last prereversal error). The exact reasons for this difference are not clear. Presumably, presenting many trials with different reversal conditions would prevent the participants from quickly reestablishing the action–outcome contingency within the first few trials. On receiving the first few negative feedbacks, the pIFC-IPL network could quickly update and maintain the outcome in working memory and then be immediately applied to affect behavior (Frank et al., 2007) (see below). This strategy was associated with a significant increase in accuracy along with significant slowing of reaction time (Fig. 2). To further improve the fluency and reduce the demands on pIFC-SPL, participants gradually inhibited the old outcome memory and established new action–outcome contingencies (i.e., for a given OR trial, the same action is associated with a different outcome after OR reversal), which might underlie the late VLPFC and caudate increase in OR.

Although FR includes both OR and RR, the present study failed to reveal similar VLPFC and caudate activation for FR. This result suggests that the initial assumption that FR reflects the additive combination of OR and RR processes is incorrect; FR may involve processes qualitatively different from those of OR and RR combined, and the right VLPFC and caudate might be

solely involved in reversal learning of outcome without accompanying response change. Alternatively, the slow-response learning during FR [as indicated by near-chance level response memory in the post-reversal probe and slower reaction time ($t_{(16)} = 3.29$, $p = 0.005$)] likely reflects delayed reestablishment of action–outcome contingency. As a result, although we found a trend for VLPFC and caudate activation for FR, this amplitude was reduced relative to OR. Further studies are definitely required to examine these important issues. One way to test these alternative hypotheses is to examine whether extended training on FR would further increase the right VLPFC and caudate activation.

At first glance, our study seems to be inconsistent with a monkey physiological study by Pasupathy and Miller (2005), which found that the caudate exhibits earlier learning than prefrontal cortex after reversal. However, our study is different from their study in several significant ways. For example, in that study, the authors used a serial reversal task in which contingencies continuously reverse for a given stimulus; moreover, the monkeys were highly trained in performing this task. In our study, reversals occurred only once per stimulus. Second, in their analysis, they focused on the dorsal lateral PFC (BA9 and BA46), whereas our study found activation in the VLPFC and pIFC. The different time courses of learning in pIFC and VLPFC revealed by the present study, together with that found by Pasupathy and Miller (2005), are consistent with the idea that subregions of PFC might show different time courses of learning or reversal learning (Laubach, 2005).

ACC-pIFC-SPL network and interference resolution

We found that the ACC-pIFC-SPL network showed strong activation for all conditions in Bin1, and although it sharply decreased for NR, it remained high for all reversal learning conditions in Bin2. Cumulative evidence suggests that the ACC is involved in performance monitoring and provides signals that engage regulatory processes in the lateral PFC to implement performance adjustments (Ridderinkhof et al., 2004a,b). The posterior IFC and adjacent precentral gyrus are strongly connected with the superior parietal lobule (Petrides, 2005). The pIFC has been implicated in several processes associated with cognitive control, including semantic selection (Thompson-Schill et al., 1997; Badre et al., 2005), response selection (Bunge et al., 2002; Dux et al., 2006), proactive interference resolution (Badre and Wagner, 2005; Feredoes et al., 2006), and conflict resolution (Derrfuss et al., 2004, 2005). Our study extended these studies by suggesting that this network might play a domain-general role in resolving both stimulus–outcome and stimulus–response interference.

Although previous studies have suggested that the ACC is responsible for error processing (Carter et al., 1998), or error likelihood prediction (Brown and Braver, 2005), other studies suggest that it is sensitive to response conflict (Botvinick et al., 1999, 2001). It has been argued that errors are more likely to occur in the presence of response conflict, and, more crucially, response conflict alone, even if it does not lead to an actual error, is sufficient to cause a change in ACC activity (Botvinick et al., 2001; Kerns et al., 2004; Ridderinkhof et al., 2004a; Rushworth et al., 2004). Our data are well consistent with this conjecture in several regards. First, our results indicate that ACC shows an increase even when no errors are committed (i.e., all of our analyses were on correct trials). Second, we found equally strong ACC activation for RR condition relative to NR condition, where the error likelihoods (as indicated by the error rate) for both were comparable. Finally, although the error likelihood decreased sig-

nificantly from Bin1 and Bin2 for FR and OR, the ACC activation remain unchanged. The cross-domain (response vs cognitive) involvement of ACC in conflict detection also extends previous observations on its generalization across response modality (manual vs verbal) and processing domains (e.g., verbal and spatial) (Barch et al., 2001).

The prolonged activation in this network during reversal learning fits well with the behavioral observations that it takes extended effort to overcome proactive interference. Our behavioral data suggest that after eight repetitions, a significant interference effect still persists. In fact, previous work has shown that this effect remains prominent even after thousands of training trials over several days (Shiu and Chan, 2006). The heavy reliance on the executive system might account for why the expression of the relearned behavior is not stable and might often fail.

In summary, our study shows that in the face of cognitive interference, the right VLPFC and caudate are involved in relearning the outcome and response–outcome contingency, whereas the ACC-pIFC-SPL network is involved in domain–general conflict resolution. The strong activation of this network in the late stage of reversal learning might provide a neural account for the behavioral difficulties in reversal learning.

References

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HB, Zilles K (1999) Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol 412:319–341.

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

Aron AR, Robbins TW, Poldrack RA (2004) Inhibition and the right inferior frontal cortex. Trends Cogn Sci 8:170–177.

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

Badre D, Wagner AD (2005) Frontal lobe mechanisms that resolve proactive interference. Cereb Cortex 15:2003–2012.

Badre D, Wagner AD (2006) Computational and neurobiological mechanisms underlying cognitive flexibility. Proc Natl Acad Sci U S A 103:7186–7191.

Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, Wagner AD (2005) Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron 47:907–918.

Barch DM, Braver TS, Akbudak E, Conturo T, Ollinger J, Snyder A (2001) Anterior cingulate cortex and response conflict: effects of response modality and processing domain. Cereb Cortex 11:837–848.

Beckmann CF, Jenkinson M, Smith SM (2003) General multilevel linear modeling for group analysis in FMRI. Neuroimage 20:1052–1063.

Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD (1999) Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature 402:179–181.

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychol Rev 108:624–652.

Brown JW, Braver TS (2005) Learned predictions of error likelihood in the anterior cingulate cortex. Science 307:1118–1121.

Budhani S, Marsh AA, Pine DS, Blair RJ (2007) Neural correlates of response reversal: considering acquisition. Neuroimage 34:1754–1765.

Bunge SA, Hazeltine E, Scanlon MD, Rosen AC, Gabrieli JD (2002) Dissociable contributions of prefrontal and parietal cortices to response selection. Neuroimage 17:1562–1571.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280:747–749.

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.

Dale AM (1999) Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109–114.

Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000) Tracking the

hemodynamic responses to reward and punishment in the striatum. J Neurophysiol 84:3072–3077.

Derrfuss J, Brass M, von Cramon DY (2004) Cognitive control in the posterior frontolateral cortex: evidence from common activations in task coordination, interference control, and working memory. Neuroimage 23:604–612.

Derrfuss J, Brass M, Neumann J, von Cramon DY (2005) Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. Hum Brain Mapp 25:22–34.

Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380:69–72.

Dux PE, Ivanoff J, Asplund CL, Marois R (2006) Isolation of a central bottleneck of information processing with time-resolved FMRI. Neuron 52:1109–1120.

Feredoes E, Tononi G, Postle BR (2006) Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. Proc Natl Acad Sci U S A 103:19530–19534.

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

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. Science 291:312–316.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2002) Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. J Neurophysiol 88:929–941.

Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE (2004) Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. J Cogn Neurosci 16:463–478.

Iversen SD, Mishkin M (1970) Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp Brain Res 11:376–386.

Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. Med Image Anal 5:143–156.

Karpicke JD, Roediger HL 3rd (2007) Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. J Exp Psychol Learn Mem Cogn 33:704–719.

Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. Science 303:1023–1026.

Laubach M (2005) Who's on first? What's on second? The time course of learning in corticostriatal systems. Trends Neurosci 28:509–511.

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202.

Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005) Valid conjunction inference with the minimum statistic. Neuroimage 25:653–660.

O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci 4:95–102.

O'Doherty J, Critchley H, Deichmann R, Dolan RJ (2003) Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J Neurosci 23:7931–7939.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004)

Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304:452–454.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 433:873–876.

Petrides M (2005) Lateral prefrontal cortex: architectonic and functional organization. Philos Trans R Soc Lond B Biol Sci 360:781–795.

Poldrack RA, Clark J, Paré-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, Gluck MA (2001) Interactive memory systems in the human brain. Nature 414:546–550.

Remijnse PL, Nielen MM, Uylings HB, Veltman DJ (2005) Neural correlates of a reversal learning task with an affectively neutral baseline: an event-related fMRI study. Neuroimage 26:609–618.

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004a) The role of the medial frontal cortex in cognitive control Science 306:443–447.

Ridderinkhof KR, van den Wildenberg WP, Segalowitz SJ, Carter CS (2004b) Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn 56:129–140.

Rushworth MF, Walton ME, Kennerley SW, Bannerman DM (2004) Action sets and decisions in the medial frontal cortex. Trends Cogn Sci 8:410–417.

Seger CA, Cincotta CM (2005) The roles of the caudate nucleus in human classification learning. J Neurosci 25:2941–2951.

Shiu LP, Chan TC (2006) Unlearning a stimulus-response association. Psychol Res 70:193–199.

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ (1997) Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc Natl Acad Sci U S A 94:14792–14797.

Tohka J, Foerde K, Aron AR, Tom SM, Toga AW, Poldrack RA (2008) Automatic independent component labeling for artifact removal in fMRI. Neuroimage 39:1227–1245.

Tricomi EM, Delgado MR, Fiez JA (2004) Modulation of caudate activity by action contingency. Neuron 41:281–292.

Wagner AD, Paré-Blagoev EJ, Clark J, Poldrack RA (2001) Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31:329–338.

Woolrich MW, Behrens TE, Beckmann CF, Jenkinson M, Smith SM (2004) Multilevel linear modelling for FMRI group analysis using Bayesian inference. Neuroimage 21:1732–1747.

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

Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nat Rev Neurosci 7:464–476.

Yin HH, Knowlton BJ, Balleine BW (2005a) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. Eur J Neurosci 22:505–512.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005b) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 22:513–523.

Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS (2003) Human striatal response to salient nonrewarding stimuli. J Neurosci 23:8092–8097.