Cerebral Cortex January 2013;23:230–240
doi:10.1093/cercor/bhs013
Advance Access publication January 31, 2012

Abstract Rule Learning: The Differential Effects of Lesions in Frontal Cortex

Andrew S. Kayser[1,2,3] and Mark D'Esposito[3,4,5]

[1]Department of Neurology, University of California, San Francisco, CA 94143, USA, [2]Ernest Gallo Clinic and Research Center, Department of Neurology, University of California, San Francisco, Emeryville, CA 94608, USA, [3]Department of Neurology, VA Northern California Health Care System, Martinez, CA 94553, USA, [4]Helen Wills Neuroscience Institute and [5]Department of Psychology, University of California, Berkeley, CA 94720, USA

Address correspondence to Andrew S. Kayser, Ernest Gallo Clinic and Research Center, Department of Neurology, University of California San Francisco, 5858 Horton Street, Suite 200, Emeryville, CA 94608, USA. Email: akayser@gallo.ucsf.edu.

Learning progressively more abstract stimulus–response mappings requires progressively more anterior regions of the lateral frontal cortex. Using an individual differences approach, we studied subjects with frontal lesions performing a hierarchical reinforcement-learning task to investigate how frontal cortex contributes to abstract rule learning. We predicted that subjects with lesions of the left pre-premotor (pre-PMd) cortex, a region implicated in abstract rule learning, would demonstrate impaired acquisition of second-order, as opposed to first-order, rules. We found that 4 subjects with such lesions did indeed demonstrate a second-order rule-learning impairment, but that these subjects nonetheless performed better than subjects with other frontal lesions in a second-order rule condition. This finding resulted from both their restricted exploration of the feature space and the task structure of this condition, for which they identified partially representative first-order rules. Significantly, across all subjects, suboptimal but above-chance performance in this condition correlated with increasing disconnection of left pre-PMd from the putative functional hierarchy, defined by reduced functional connectivity between left pre-PMd and adjacent nodes. These findings support the theory that activity within lateral frontal cortex shapes the search for relevant stimulus–response mappings, while emphasizing that the behavioral correlate of impairments depends critically on task structure.

Keywords: connectivity, hierarchy, prefrontal cortex, stroke

Introduction

The capacity to adapt rapidly and flexibly to novel circumstances represents a fundamental feature of higher cognitive function. As shown by multiple investigators, this ability to abstract—for example, to discover higher order relationships (Robin and Holyoak 1995), to chunk lower level items (Chase and Simon 1973; Newell 1990), and to analogize/transfer information to new situations (Gick and Holyoak 1980 1983)—depends critically on the frontal and prefrontal cortex (Koechlin et al. 2003; Bunge 2004; Petrides 2006; Bor and Owen 2007; Christoff and Keramatian 2007; Koechlin and Summerfield 2007; Badre 2008; Badre and D'Esposito 2009). In addition, multiple theoretical and empirical accounts have suggested that more anterior regions of the frontal lobe support more abstract representations (reviewed in Badre 2008), though these accounts differ in important ways. One theory emerging from the study of working memory argues that increasingly anterior regions of prefrontal cortex represent domain-general, as opposed to domain-specific, information (Christoff and Gabrieli 2000; Buckner 2003; Courtney 2004). In this formulation, more caudal regions of the frontal cortex maintain the location of an object across a delay, for example, while more rostral regions maintain both location and object identity. Another theory argues that increasingly rostral areas of prefrontal cortex represent greater relational complexity (Robin and Holyoak 1995; Christoff et al. 2001; Christoff and Keramatian 2007). Under this formulation, the number of relationships required to generate a response—that is, whether the response depends upon the color of an object (one relation) or the match between colors of different objects (two relations)—determines the locus of brain activity (Christoff and Keramatian 2007; and discussed in Badre 2008). Importantly, neither of these theories necessitates that processing in lower order brain regions is dependent upon processing in higher order regions, or vice versa.

Recently, a number of studies have suggested that our capacity to apply rules to new situations, and to identify new rules based on current circumstances, may be supported by a specifically hierarchical organization of lateral frontal cortex—that is, by an organization in which progressively more anterior regions process progressively more abstract representations, and in which superordinate frontal regions modulate responses in subordinate ones (Koechlin et al. 2003; Koechlin and Jubault 2006; Badre and D'Esposito 2007; Badre et al. 2009). Support for this idea (discussed further below) has come from studies based on policy abstraction, a form of abstraction in which first-order stimulus–response mappings are potentially contingent upon more abstract second-order (and higher order) rules (Badre and D'Esposito 2007; Badre et al. 2009). Notably, in these theories, interactions between representations at different levels of abstraction are critical.

To evaluate whether a policy abstraction account might explain how we learn abstract action rules, we recently designed a novel reinforcement-learning task (Badre et al. 2010). In this task, participants were required to learn 2 sets of rules, in separate epochs, that linked each of 18 different stimuli uniquely and deterministically to 1 of 3 button-press responses (Fig. 1). For each rule set, an individual stimulus consisted of 1 of 3 shapes, at 1 of 3 orientations, inside a box that was 1 of 2 colors for a total of 18 unique stimuli (3 shapes × 3 orientations × 2 colors). Participants initially learned the correct one of the 3 possible button-press responses, for each stimulus, based on trial and error responding (Fig. 1*A*). For 1 of the 2 rule sets—the "Flat" set—each of the 18 rules had to be learned individually as one-to-one mappings (first-order policy) between a conjunction of color, shape, and orientation and a response (Fig. 1*B,D*). In the other set—the "Hierarchical" set—stimulus display parameters and instructions were identical to those for the Flat set. In fact, the Hierarchical set could also be learned as 18 individual first-order rules. However, the stimulus–response mappings

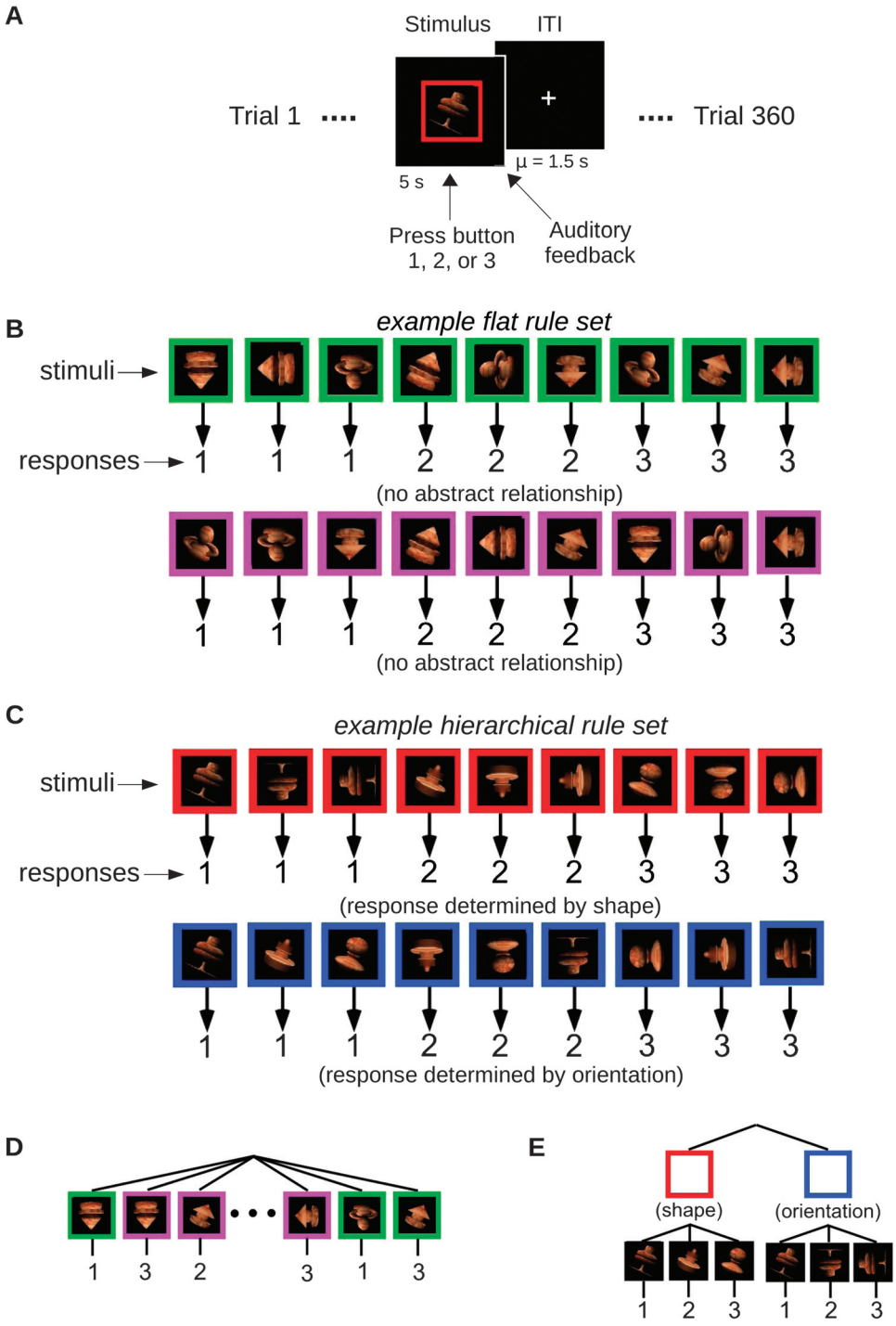

Figure 1. Schematic depiction of trial events, example stimulus-to-response mappings, and policy for Hierarchical and Flat rule sets. (*A*) Trials began with presentation of the stimulus for 5 s, during which subjects could respond with a button press at any point. Immediately after stimulus presentation, participants received auditory feedback indicating whether the response they had chosen was correct given the presented stimulus. Trials were separated by a variable intertrial interval with a mean of 1.5 s. (*B*) Example stimulus-to-response mappings for the Flat set. The arrangement of mappings for the Flat set was such that no higher order relationship was present; thus, each rule had to be learned individually. (*C*) Example stimulus-to-response mappings for the Hierarchical set. Response mappings are grouped such that in the presence of a red square, only shape determines the response, while in the presence of a blue square only orientation determines the response. (*D*) The Flat set of many first-order rules can be represented as a large flat policy structure with only 1 level and 18 alternatives. (*E*) The Hierarchical set can be represented as a 2-level policy structure with a second-order rule selecting between the shape or orientation mapping sets and a set of first-order rules linking specific shapes or orientations to responses.

were defined such that a second-order relationship could be learned instead, thereby, reducing the number of first order rules to be learned (Fig. 1*C*). Specifically, in the context of 1 colored box, only the shape dimension was relevant to the response, with each of the 3 unique shapes mapping to one of the 3 button responses irrespective of orientation. In contrast, in the context of the other colored box, the orientation dimension fully determined the response. Thus, the Hierarchical rule set permitted learning of abstract, second-order rules mapping color-to-feature along with 2 sets of first-order rules (i.e., specific

shape-to-response and orientation-to-response mappings; Fig. 1*E*). Critically, this simplifying second-order structure did not permit subjects to simply ignore one of the stimulus features. For example, learning only the first-order mapping of orientation to response in the Hierarchical case would fully identify one half of the stimulus space—that in which only orientation determined the response—but not the other half of the stimulus space, in which only shape determined the response.

In this work (Badre et al. 2010), we demonstrated that previously identified first- and second-order cortical regions in the left lateral frontal cortex are associated with learning first- and second-order stimulus–response mappings (left dorsal premotor [PMd] and left dorsal pre-premotor [pre-PMd] cortex; Picard and Strick 2001), respectively. Along with higher order mappings associated with even more anterior regions—activations within inferior frontal sulcus and a region within frontopolar cortex, for example, have been associated with third- and fourth-order mappings, respectively (Badre and D'Esposito 2007)—these regions define an anatomical hierarchical ordering of first- through fourth-order representations. However, arguing that this hierarchical organization is important for abstract rule acquisition, and that it might provide a more complete explanation for the gradient of abstraction in frontal cortex than other theories, would benefit greatly from an approach that addresses disruption of the system. For this reason, here we investigate the effects of lesions in relevant cortices on neural function.

If a hierarchical organization impacts second-order rule identification, at least 3 specific predictions follow. First, disruption of left pre-PMd, but not closely adjacent cortical areas, should disrupt acquisition of second-order rules. Second, because this area is enmeshed in a hierarchical structure, the degree to which this region is disconnected from other (i.e., adjacent first- and third-order) nodes in the hierarchy—that is, its residual intrinsic connectivity within the hierarchy—should correlate with task performance. Neither of these predictions is required for other theories of prefrontal cortical (PFC) performance, as these theories either specify different (or broader) cortical regions, or do not necessarily depend upon functional connections with other areas. Third, the dependence on these functional connections should itself vary with the structure/level of abstraction of the task. To address the above hypotheses about this putative lateral frontal hierarchy, we followed an individual differences approach to test subjects with frontal lobe lesions on our hierarchical reinforcement-learning task.

Materials and Methods

Participants

Eighteen English-speaking subjects (mean age 61.7 ± 9.3 years, range 44–75 years) with single lesions due to ischemic stroke (n = 11), intracerebral hemorrhage (n = 3), traumatic brain injury (n = 3), or tumor resection (n = 1) were studied (Table 1 and Supplementary Information). All visible lesions were limited to the frontal lobe. Subjects were at least 1.5 years postevent (mean 11.3 ± 9.0 years; range 1.5–34 years) and were prescreened to exclude individuals with a history of other neurological or psychiatric conditions. In particular, stroke patients with a history of cardioembolic stroke were selected in order to minimize the contribution of known atherosclerotic cerebrovascular disease to the neuroimaging data. A neuropsychological battery was administered to all subjects (see Supplementary Information). One subject with a left frontal lesion had a right face and arm hemiparesis that required him to respond with the left rather than the right hand. Written informed consent was obtained from subjects in accordance with procedures approved by the Committee for Protection of Human Subjects at the University of California, Berkeley.

Logic and Design

In order to investigate the discovery of abstract rules, we used a reinforcement-learning task that required the learning of 2 rule sets, one of which contained a higher order rule structure (Hierarchical rule set) and one that could only be learned as one-to-one mappings between stimuli and responses (Flat rule set; both rule sets shown in Fig. 1). Participants were not given an indication through an instruction or any other cue that a higher order structure existed in one of the rule sets. Moreover, trials for both rule sets were identical in terms of all stimulus presentation parameters, instructions, and response-reward contingencies.

Each rule set was learned over the course of 360 individual learning trials. Each trial commenced with the presentation of a stimulus display consisting of a nonsense object (i.e., without a real-world counterpart) appearing in 1 of 3 orientations (up [0°], left [–90°], or oblique [23°]) and bordered by a colored square. For each rule set, 2 colors, 3 object shapes, and 3 orientations were used, giving rise to 18 unique stimulus displays (i.e., 3 shapes × 3 orientations × 2 colors). Each of the 18 unique displays occurred 20 times for each rule set (Hierarchical and Flat). The specific colors and shapes differed across the 2 rule sets within subject and were counterbalanced for rule set across subjects.

The object and square appeared together for 5 s and were then replaced by a white fixation cross for a variable intertrial interval (mean of 1.5 s, range 0–8 s). While the stimulus display was present, the participant could respond with 1 of 3 buttons using the index, middle, or ring fingers of the right hand. (For one subject with significant right-hand weakness, the left hand was used.) Once a response was made or 6 s had passed without a response, the green fixation cross turned red and no further responding was allowed. A lack of response was scored as an incorrect trial. Subjects then received auditory feedback: a high tone (750 Hz) indicated a correct response and a buzzing tone (combination of 300 and 400 Hz pure tones) indicated an incorrect response. A running total of correct responses was displayed at the end of each run of 60 trials. The order of trials within a block was determined in pseudorandom fashion, as described in our previous study (Badre et al. 2010), and the order of rule set learning (i.e., whether Hierarchical or Flat was learned first) was counterbalanced across participants.

For both rule sets, participants were given the same instructions. No indication was given that a higher order relationship existed or that they should search for an abstract rule. Participants did not practice the task but they were allowed to fully familiarize themselves with all 18 stimuli they would encounter for a given rule set prior to conducting the learning trials for that rule set. As a result, the 2 rule sets differed only in the arrangement of mappings between stimulus displays and responses (Fig. 1*B–E*).

Behavioral Analysis

Learning curves were calculated using a state-space modeling procedure (Smith et al. 2004) that estimates the probability of a correct response on each trial as a function of a latent Gaussian state process (i.e., the state of knowledge of the subject) and an observable Bernoulli response process (i.e., the responses of the subject). In other words, the model uses the learner's trial-by-trial responses (either correct or incorrect) to estimate his/her knowledge about the task over time. In contrast with "sliding average" or other methods of computing learning curves, this approach allows one to define a confidence interval associated with the estimate of learning on each trial. Thus, this method produces a "learning trial," or the trial at which the confidence interval no longer encompasses chance performance. Because this method estimates a single value for the variance of the Gaussian state process across learning, it does not incorporate details of the task or make assumptions about hierarchical learning (for further details, see Smith et al. 2004). Learning curves using this procedure were calculated for each subject for the entire rule set, as well as for each of the 18 individual rules based on the 20 presentations of a particular stimulus.

Table 1
Demographic information for each of the 18 subjects

Subject number	Age	Education (years)	Lesion site	Lesion size (cc)	Time since onset (years)	Etiology
1	61	14	R lateral PFC	49	6	Ischemic stroke
2	75	13	R lateral PFC	105	4.5	Ischemic stroke
3	60	14	R basal ganglia	7	8	Hemorrhagic stroke
4	73	14	L basal ganglia	3	15	Hemorrhagic stroke
5	68	14	L basal ganglia	6	6	Hemorrhagic stroke
6	56	9	B OFC (R > L)	137	34	Trauma
7	63	13	L rostromedial PFC	65	4	Ischemic stroke
8	60	16	B OFC	247	16	Tumor resection
9	72	12	R inferolateral PFC	20	10	Ischemic stroke
10	44	16	B OFC	18	1.5	Trauma
11	67	16	R lateral OFC	12	33	Trauma
12	63	15	L lateral PFC	92	3.5	Ischemic stroke
13	58	18	L lateral PFC	62	9	Ischemic stroke
14	65	11	L lateral PFC	116	13	Ischemic stroke
15	75	16	L lateral PFC	147	10.5	Ischemic stroke
16	51	20	L lateral PFC	150	12	Ischemic stroke
17	47	18	L lateral PFC	122	7.5	Ischemic stroke
18	52	20	L lateral PFC	237	10.5	Ischemic stroke

Note: R, right; L, left; B, bilateral; OFC, orbitofrontal cortex.

We focused our behavioral analysis on 2 measures of learning for both Hierarchical and Flat sessions: 1) the terminal accuracy (i.e., the probability of a correct response on the final trial), which is related to the degree of learning at the conclusion of each of the Hierarchical and Flat conditions and 2) the learning trial for each of the 18 stimulus–response mappings in each of the Hierarchical and Flat conditions—that is, that trial, if any, at which there was a 95% or greater probability that responding for a given mapping was different from chance performance. Individual object-response mappings were considered to be learned if a learning trial could be defined, or if the number of actual correct responses for an object deviated significantly from the expected number of correct responses for chance responding, based on Bernoulli assumptions (>13 correct responses of the 20 presentations of each object; $P < 0.05$, Bonferroni corrected for the number of objects). We also compared the number of learned objects for each colored square, in both Hierarchical and Flat sessions, in order to determine whether learning was specific to a subset of feature combinations. Terminal accuracy values were variance stabilized via an arcsine square root transform prior to statistical analyses. All parametric tests were performed under the assumptions of potentially unequal numbers and variances, with the Welch–Satterthwaite approximation used to estimate the degrees of freedom. As a consequence, the degrees of freedom varied from test to test, despite unchanging numbers of subjects.

Magnetic Resonance Imaging Acquisition Procedures

T_2^*-weighted echo planar images (EPIs) were collected on a whole body 3-T Siemens MAGNETOM Trio magnetic resonance imaging (MRI) scanner using a 12-channel head coil. Structural images were acquired using an axial MP-RAGE 3D T_1-weighted sequence (time repetition [TR] = 2300 ms, time echo [TE] = 2.98 ms, flip angle = 9°, voxel size = 1 mm^3) and a fluid attenuated inversion recovery image to assist with lesion visualization, as in our previous work (Nomura et al. 2010). Resting state images consisted of 28 slices acquired with a gradient echoplanar imaging protocol (300 time points for each of 2 runs, TR = 2000 ms, TE = 30 ms, field of view = 225 mm, matrix size = 128 × 128, voxel size = 1.75 × 1.75 × 3.3 mm) for each of the 18 subjects. Prior to resting state scans, participants were instructed simply to remain awake with their eyes open. All scans were obtained at least 6 months after the index event for each subject.

MRI Preprocessing and Lesion Definition

In keeping with our previous work (Nomura et al. 2010), the software package AFNI (Analysis of Functional NeuroImages) was used for slice timing correction, image realignment, and removal of nonbrain structures from the EPI volumes prior to spatial smoothing with a 6-mm full-width at half-maximum Gaussian kernel. The high-resolution T_1-weighted image was co-registered with the mean functional data and segmented using SPM5 (Wellcome Department of Cognitive Neurology, London) via a template derived from 152 normal subjects (MNI152; Montreal Neurological Institute, Montreal, Quebec, Canada). All analyses were then performed on the native-space functional images. The extra segmentation step was necessary for accurate registration of images demonstrating structural brain damage. To address the effect of subjects' lesions on regions of interest (ROIs) implicated in the learning and application of abstract rules, we calculated the percentage of voxels in each ROI that overlapped with the lesion mask (Fig. 2). Lesion masks were constructed in our previous study (Nomura et al. 2010), with all individual subject masks shown in normalized space in Supplementary Figure S1.

Functional connectivity

Twelve ROIs were derived from our previous work (Badre et al. 2010), including 4 left lateral frontal cortical areas representing first- through fourth-order levels of policy abstraction—dorsal premotor cortex (PMd), pre-PMd cortex, the inferior frontal sulcus (IFS), and frontopolar cortex, respectively—as well as bilateral caudate and putamen ROIs. Activity within these basal ganglia ROIs was noted to share Granger causal influences with both PMd and pre-PMd in our previous study (Badre et al. 2010). Given the potential importance of cortico-striato-thalamic loops in cognitive processing (Alexander et al. 1986; Houk and Wise 1995; Graybiel 1998), we included these basal ganglia ROIs in our analyses. To derive right-sided frontal ROIs, we chose areas homologous to the left-sided regions by simply inverting the x-coordinate for each left-sided ROI. These 12 ROIs (Fig. 2*A*) were then reverse normalized to each subject's native space, utilizing the normalization parameters obtained from the SPM5 segmentation tool.

After preprocessing, each time series for the two 5-min resting-state runs was windowed with a 4-point split-cosine bell and concatenated with the other segment to produce a subject-specific 600 time-point series for every voxel in the brain. Time series within each ROI were then averaged across voxels to generate a single time series for each ROI. Coherency values were obtained by applying a fast Fourier transform (Matlab 6.5, http://www.mathworks.com) to the data for each pair of ROIs, implemented via Welch's periodogram averaging method using a 64-point discrete Fourier transform, Hanning window, and overlap of 32 points (Kayser et al. 2009). Coherence values for each ROI were then computed using the band-averaged coherence. To compute correlations between coherence results and other values, we first Fisher transformed the coherence values to generate an approximately normal distribution (Rosenberg et al. 1989) that permitted us to apply parametric statistical tests.

Results

We tested 18 subjects with brain lesions, involving the frontal cortex and/or basal ganglia and potentially affecting 12 ROIs

implicated in the learning and execution of hierarchical rules (Fig. 2, Supplementary Figures S1 and S2; see also Materials and Methods). In keeping with our previous data and with other results from both control and subject populations that the left lateral prefrontal cortex may be implicated in rule processing (Goel and Dolan 2004; Reverberi et al. 2009), subjects were ordered by lesion location, such that low subject numbers were associated with right-predominant lesions of pre-PMd cortex and high numbers with left-predominant lesions. Four subjects demonstrated lesions that significantly involved left pre-PMd (Fig. 2*C*, subjects 15–18). Both neuropsychological testing and demographic variables were assessed (see Supplementary Materials).

As demonstrated by their learning curve trajectories, none of the subjects reached perfect performance in either the Hierarchical or the Flat condition (Fig. 3). Notably, across the group of subjects, there were no significant differences between terminal accuracies or the learning trial in the Hierarchical and Flat conditions (Ps > 0.22), suggesting that subjects did not uncover a second-order rule in the

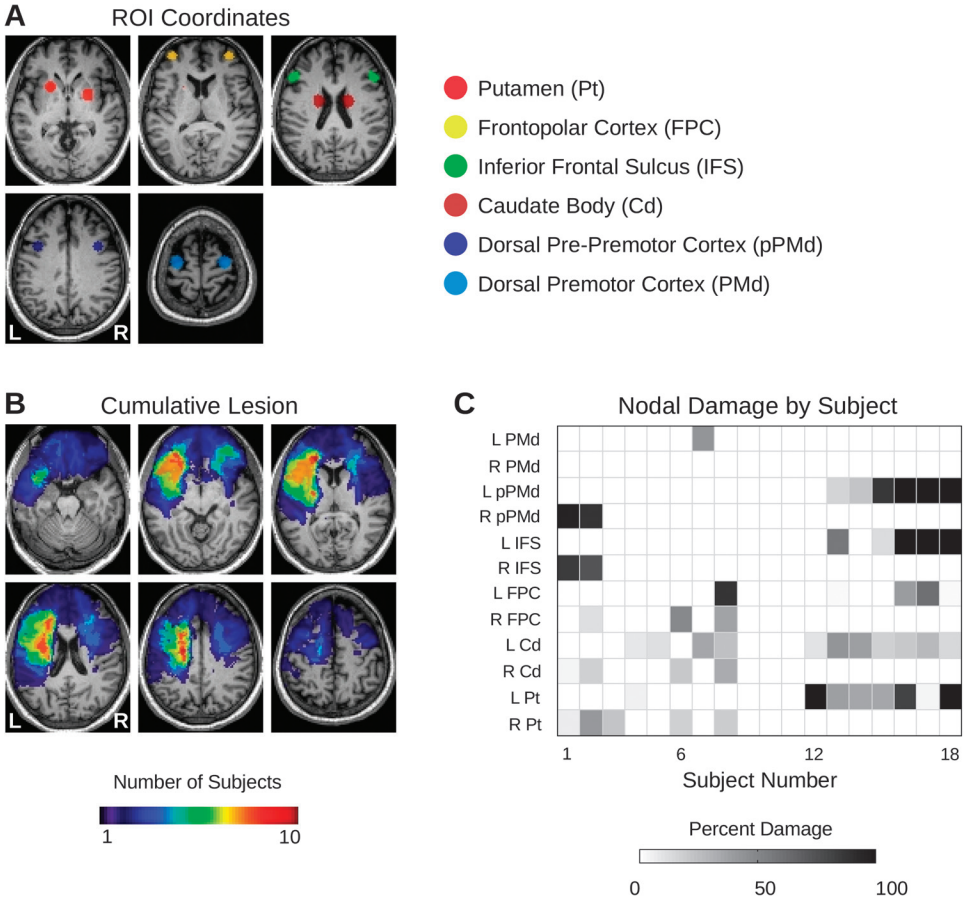

Figure 2. (*A*) Locations of ROIs. (*B*) The cumulative lesion burden across all 18 subjects. The number of subjects with overlapping lesion locations is indicated by the color bar at bottom. (*C*) The percentage of voxels in each ROI affected by the single lesions in each of the 18 subjects. Note that the lesions in subjects 9–11 did not involve any of the prespecified ROIs. (Please see Supplementary Figure S1 for individual subject lesions.)

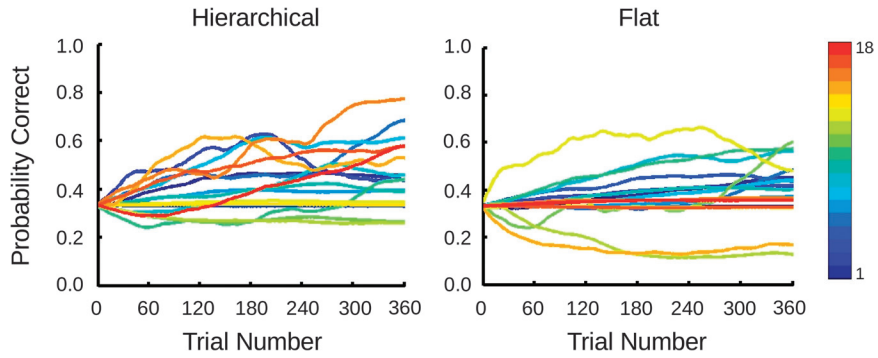

Figure 3. Learning curves for each of the (color coded) 18 subjects across the 360 trials for the Hierarchical (left) and Flat (right) rule sets. A probability correct of 0.33 represents chance performance. The probability correct after the last trial is defined as the terminal accuracy.

Hierarchical condition (Badre et al. 2010). To investigate whether performance in the Hierarchical condition was differentially affected by lesion location, we evaluated the Hierarchical and Flat conditions in 4 subjects with complete or near-complete lesions of the second-order region, left pre-PMd. As evidenced by Fig. 4*A*, there was a strongly significant effect of lesion location on differential learning in the Hierarchical versus Flat cases ($F_{1,14}$ = 22.7, P = 0.0003). Despite failing to learn the full second-order rule space, left pre-PMd subjects showed significantly better differential accuracy than subjects with other frontal lesions: Hierarchical – Flat difference = 0.10 versus 0.00, T_7 = 5.35, P = 0.001. (Importantly, this result remained significant if group membership was weighted by the extent of left pPMd involvement by the lesion, thereby incorporating the minor influence of subjects 13–14: $F_{1,14}$ = 17.8, P = 0.0009; weighted differences 0.08 versus 0.00, T_6 = 4.5, P = 0.002.) The differences between terminal accuracies in the Hierarchical and Flat conditions for these subjects were driven primarily by differences in terminal accuracy for the Hierarchical (second-order) rule set (left pre-PMd group = 0.62; other group = 0.41; T_7 = 5.3, P = 0.001; Fig. 4*B*). There were also concordant between-group differences for these values,

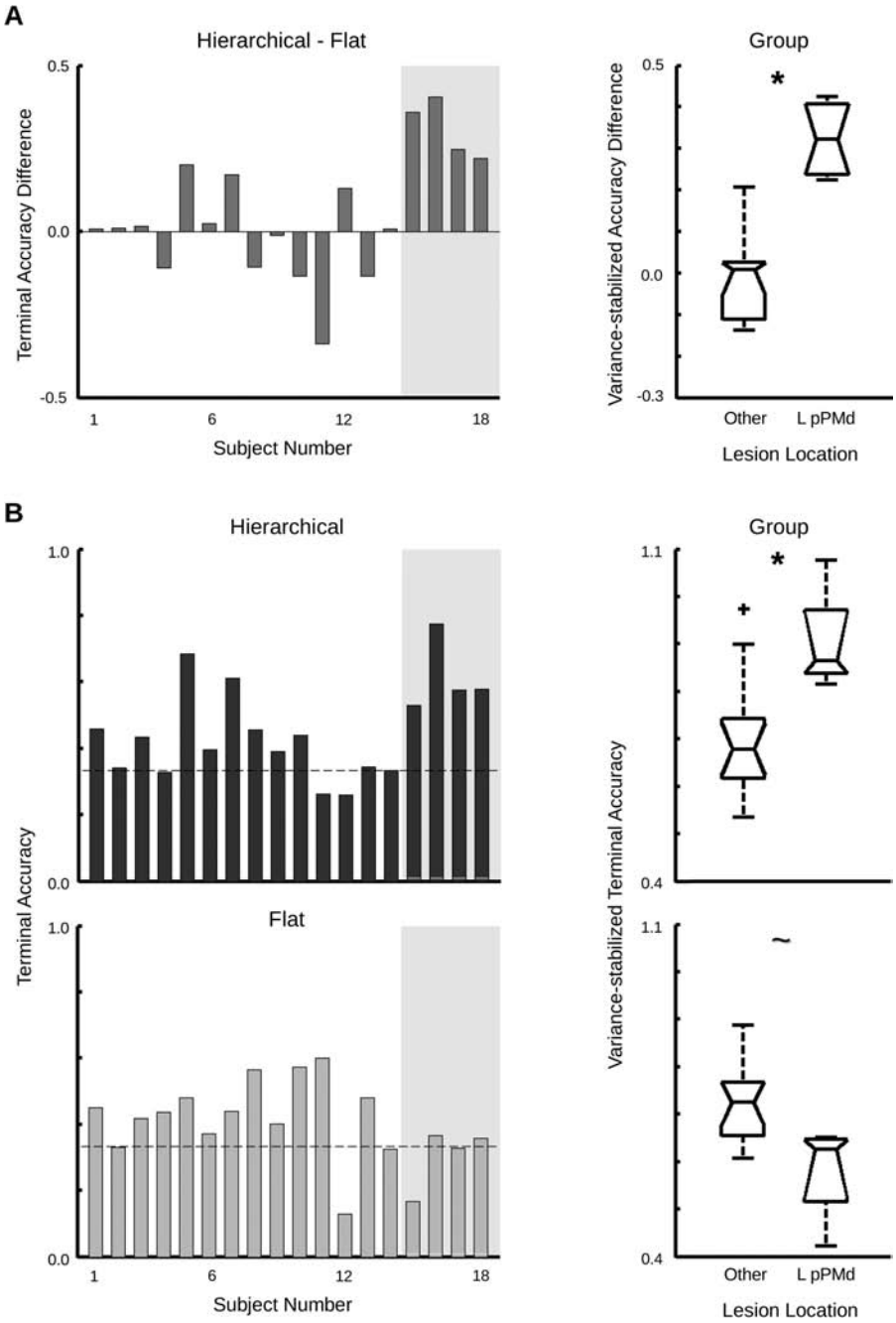

Figure 4. (*A*) The difference between the terminal accuracies in the Hierarchical and Flat conditions for each of the 18 subjects. Those subjects with complete or near-complete lesions of left pre-PMd cortex (left pre-PMd; subjects #15–18) are highlighted by the gray shading. The box-whisker plot to the right summarizes these differences for the left pre-PMd and other-lesion groups. (*B*) Terminal accuracies for the subjects in the Hierarchical condition (top) and Flat condition (bottom). *indicates $P \leq 0.001$; ~ indicates $0.05 < P < 0.10$.

though only at trend significance, for the Flat rule (left pre-PMd group = 0.30; other group = 0.42; $T_6 = -2.0$, $P = 0.09$; Fig. 3*C*). Thus, despite failing to learn the second-order rule structure, subjects with left pre-PMd lesions performed significantly better in the Hierarchical condition than subjects with lesions elsewhere, even closely adjacent ones.

To understand why subjects with left pre-PMd lesions performed better than other subjects in the Hierarchical than Flat condition, we compared the number of individual stimulus–response mappings successfully learned in subjects with and without left pre-PMD lesions. Across both the Hierarchical and Flat conditions, the left pre-PMd and other lesion groups learned equal numbers of stimulus–response mappings (11 vs. 9.1, $T_5 = 0.74$, $P = 0.5$ [not significant, ns]). However, there were trends for left pre-PMd subjects to learn more Hierarchical rules (8.3 vs. 4.2, $T_5 = 2.4$, $P = 0.06$) and fewer Flat rules (2.8 vs. 4.9, $T_{10} = -2.0$, $P = 0.07$) than subjects with other lesions. Importantly, the mappings learned by left pre-PMd subjects in the Hierarchical case were not divided equally across the "shape" and "orientation" rules that together comprised the second-order condition; rather, subjects learned one, but not both, of these mappings. The absolute difference between number of "shape" mappings and number of "orientation" mappings learned in the Hierarchical case for subjects with left pre-PMd lesions was both significantly greater than in the Flat case (5.25 vs. 1.25, $T_3 = 6.9$, $P = 0.006$) and significantly greater than the performance of the other lesion group in either rule condition (both Ps $< 10^{-5}$; Fig. 5*A*). As this finding suggests, and as evidenced by the left pre-PMd subject with median performance (#15, Fig. 5*B*), these subjects learned significantly more than the other-lesion subjects, but on only a restricted portion of the rule space: that is, learning for 3 (1) subjects occurred on the "shape" ("orientation") rule and not on the "orientation" ("shape") rule.

To understand the neural correlates of these performance differences, we evaluated resting state functional connectivity between left pre-PMd and other nodes implicated in policy abstraction (Badre and D'Esposito 2007; Badre et al. 2010). We hypothesized that search restricted to a portion of the rule space—that is, search in which second-order rules were not evaluated—should be correlated with disconnection of the second-order region from the hierarchy—that is, decreased baseline connectivity between left pre-PMd and superordinate (left IFS) and subordinate (left PMd) nodes. Consistent with this hypothesis, a significant inverse relationship could be seen between terminal accuracy in the Hierarchical condition and both left pre-PMd ↔ left IFS connectivity ($R = -0.69$, $P = 0.001$; Fig. 6*A*, left) and left pre-PMd ↔ left PMd connectivity ($R = -0.52$, $P = 0.03$; Fig. 6*A*, right). In other words, decreased connectivity between the second-order region (pre-PMd) and both first- and third-order areas was correlated with improved performance on a subset of the rule space in the Hierarchical condition. Significantly, neither relationship to connectivity held for the Flat condition ($R = 0.00$, ns and $R = -0.25$, ns; Fig. 6*B*)—that is, in the Flat case, the task structure does not reward search that neglects one feature.

Importantly, the above correlation results were not confined only to the 4 subjects with left pre-PMd lesions. A concern might be that these correlations arose because subjects known to perform better in the Hierarchical condition had lesions in the ROI, and therefore such significant correlations resulted from the performance of these subjects alone. However, when

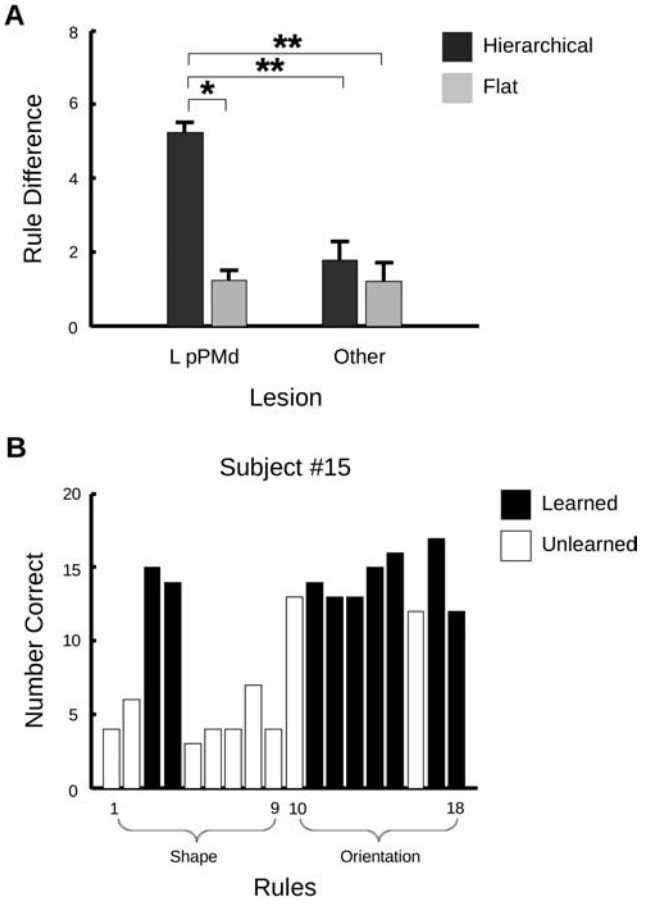

Figure 5. (*A*) The difference between the number of learned rules associated with one colored square versus the other (in the Hierarchical condition, representing the "shape" and "orientation" rules). The left pre-PMd group shows asymmetric learning across colored squares in the Hierarchical condition only (*$P < 0.01$; **$P < 0.0001$). (*B*) Number of correct responses for each of the 18 stimuli during the Hierarchical condition for subject #15, who showed median terminal accuracy within the left pPMd group. Learned rules are indicated by the black shading, while unlearned rules are indicated by white shading. The first 9 stimuli are part of the "shape" rule, while the second 9 are part of the "orientation" rule.

subjects #15–18 were removed from the calculation (dark gray circles, Fig. 6), both correlations remained strongly significant ($R = -0.71$, $P = 0.005$ and $R = -0.66$, $P = 0.01$, respectively). Moreover, when partial correlations were taken with respect to demographic variables—age, education, and lesion size—for all subjects, these correlations remained ($R = -0.73$, $P = 0.002$ and $R = -0.54$, $P = 0.04$, respectively). Similar correlations were not seen with respect to performance in the Flat condition (all Ps > 0.32; Fig. 6*B*). Importantly, these correlations between connectivity and behavior also captured the performance of subjects #5 and #7 (left caudate lesions; Fig. 2*C*), who performed differentially well in the Hierarchical rule case relative to the Flat case (Fig. 4*A*). Thus, the inverse correlation between reduced left pre-PMd connectivity and Hierarchical performance held across all subjects, not just those with left pre-PMd lesions.

Discussion

Our capacity for generalizing previously learned behaviors to changed circumstances, a hallmark of intelligent behavior, is critical to our day-to-day ability to navigate the world. Its

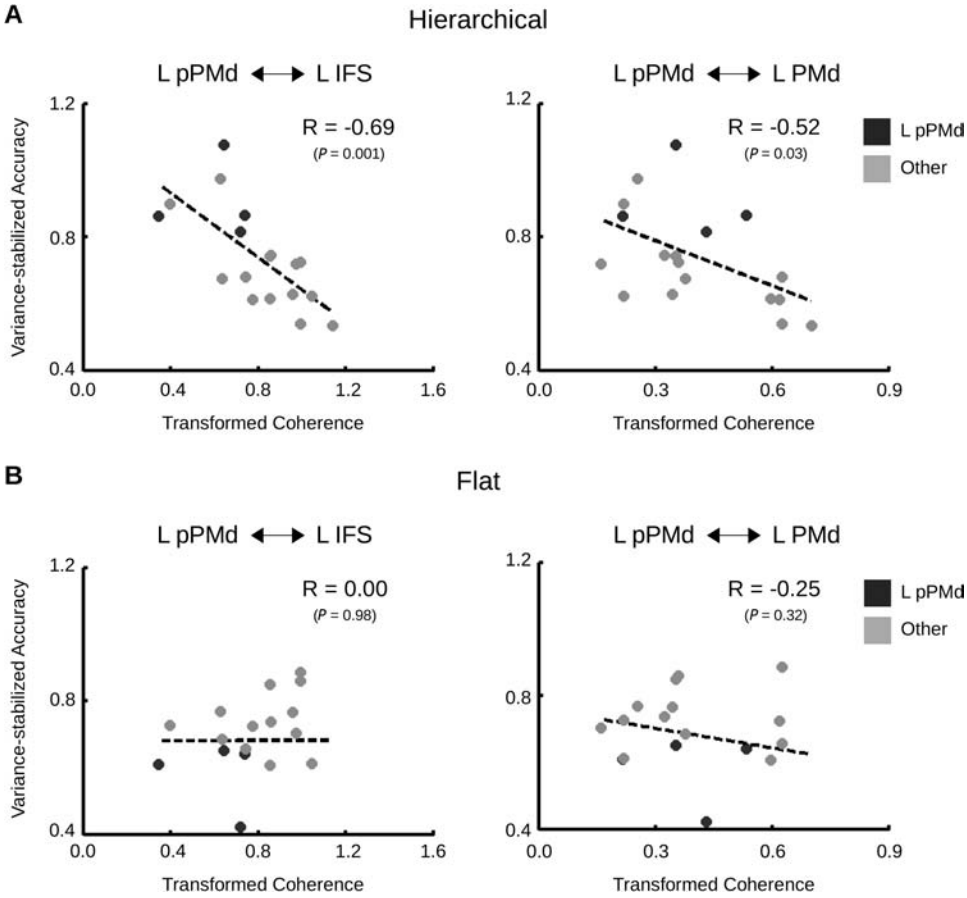

Figure 6. (*A*) The correlation of terminal accuracy in the Hierarchical condition with resting-state coherence between left pre-PMd and left IFS (left panel) and between left pre-PMd and left premotor cortex (PMd; right panel). Subjects in the left pre-PMd group are indicated by dark gray circles; subjects in the other-lesion group are indicated by light gray circles. Both correlations remain significant if left pre-PMd subjects are excluded. (*B*) The correlation of terminal accuracy in the Flat condition with resting-state coherence between these same ROIs. Neither correlation value is significant.

importance is demonstrated by behavioral changes in subjects with lesions of the lateral prefrontal cortex, following which cognition is often described as more concrete, perseverative, and/or stimulus bound (Devinsky and D'Esposito 2004). Our recent work investigating the neural basis for such adaptive behavior in healthy individuals demonstrated that the ability to learn an abstract rule structure correlates with activity in hierarchically organized lateral frontal regions: specifically, with activity in the left dorsal pre-PMd cortex (left pre-PMd) for the acquisition of a second-order rule structure (Badre et al. 2010). Here, we extend these findings to subjects with frontal lesions, and show that, while no subject learned the full rule structure when a more abstract rule was available, 4 subjects with lesions of left pre-PMd performed better in the Hierarchical condition than those with frontal lesions elsewhere. Consistent with the hierarchy hypothesis, learning for left pre-PMd subjects was restricted to a portion of the rule space in the Hierarchical condition. Moreover, across "all" subjects with lesions, the degree to which left pre-PMD was disconnected from the hierarchy—that is, from the brain regions functionally above (IFS) and below (PMd) it—correlated with improved performance in the Hierarchical task only.

Other possible explanations do not account as well for our data. These results, for example, are not simply a consequence of lesions that affect the left hemisphere, irrespective of hierarchy. Subjects 12–14, for example, demonstrate significant left-sided lesions (Supplementary Figure S1), but did not show the behavior of patients 15–18, whose left-sided lesions involved most or all of left pre-PMd. Another question concerns the role of lesion extent, as the lesions involving left pre-PMd also involved the third-order region in the left inferior frontal sulcus (Fig. 2). We cannot exclude that the lesion of IFS is also critical to this behavior, but note that functional connectivity between regions outside of IFS (i.e., between pre-PMd and PMd) was also negatively correlated with performance (discussed further below). While this change could arise secondary to the change in coherence between IFS and pre-PMd, a more parsimonious explanation is that both coherence values involving pre-PMd (i.e., PMd–pre-PMd and pre-PMd–IFS) correlate with behavior because of the disconnection of one area: pre-PMd.

By what cognitive mechanism, then, do lesions to left pre-PMd but not other frontal regions lead to suboptimal, but relatively advantageous, performance? An attractive idea is that higher order hierarchical regions have an important role in resolving competition in lower order regions (Badre and D'Esposito 2007, 2009; Badre et al. 2009), allowing different lower order stimulus–response relationships to be active at different times (Bunge 2004). If a higher order region is no longer capable of resolving this competition, the ability to flexibly employ different lower order rules would be lost. In a learning experiment, lesions of left pre-PMd could thereby

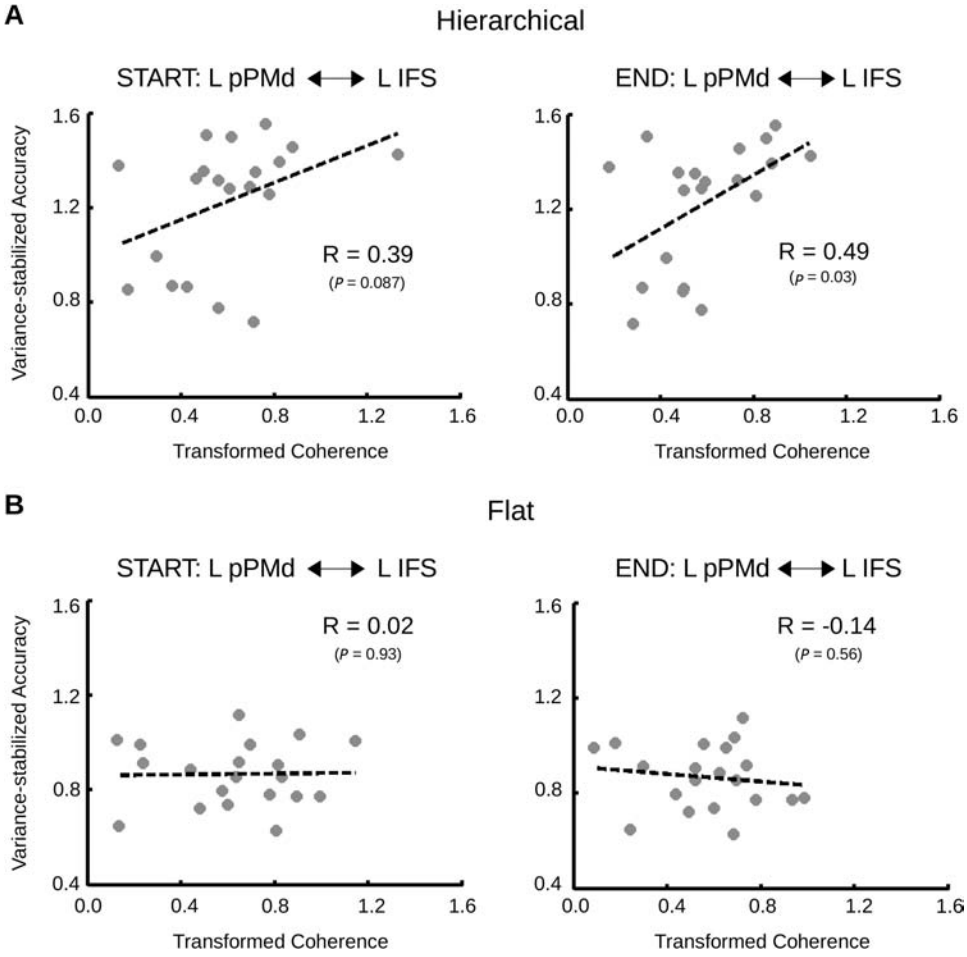

Figure 7. (*A*) The correlation of terminal accuracy in the Hierarchical condition with task-state coherence between left pre-PMd and left IFS during both early learning (i.e., during the first third of trials—left panel) and later learning (i.e., during the last third of trials—right panel) for healthy subjects from our previous study (Badre et al. 2010). (*B*) The correlation of terminal accuracy with functional connectivity between these same regions in the Flat condition. Neither correlation value is significant.

limit the rule search space to a subset of possible rules driven by function in the (intact) first-order region (PMd).

However, this reduction in the search space of possible stimulus–response mappings was clearly helpful in patients with pre-PMd lesions, permitting suboptimal but above-chance performance on the task. In theory, the reduction in search space is potentially quite significant; given the 18 unique stimuli and 3 responses, the number of possible combinations of S–R mappings is quite large: 3^{18} = 387 420 489. While a number of these mappings are unlikely—for example, ones in which all stimuli map to the same response or to unbalanced numbers of the different responses—patients who effectively neglected one of the features (i.e., who selected combinations of color and shape or combinations of color and orientation, rather than combinations of all 3 features) would reduce the number of effective stimuli to 6, rather than 18, and decrease the search space to 3^6 = 729 combinations. Given further assumptions about a relatively equal distribution of button-press responses, for example (which all subjects expressed), a much smaller space could be searched; and the strategy, while not optimal, would be relatively successful in the Hierarchical task (Frank and Badre 2011) given that one-half of the S–R mappings in this rule set effectively "ignore" one feature. Consistent with this explanation, such a strategy would be quite ineffective in the Flat condition, where all 3 features are important.

In support of these ideas, the above behavioral changes were inversely correlated with a neurophysiological measure of left pre-PMd connectivity across all subjects. Specifically, coherence between left pre-PMd and super- (IFS) and subordinate (PMd) cortical areas within the policy abstraction hierarchy was lower in subjects who learned more accurately. Thus, even in subjects without lesions of left pre-PMd, increasing disconnection of this second-order region—accompanied by a corresponding restriction in the search space of second-order rules—improved performance only in the Hierarchical condition. This finding highlights the specificity of these results for the "function" of left pre-PMd, whether the functional deficit is due to an overt lesion or not, and supports the notion of a hierarchical relationship between these regions. As noted in the Results, for example, 2 subjects with lesions of the left caudate (subjects #5 and #7) demonstrated similar patterns of performance as did the left pre-PMd subjects and had similar levels of disconnection of left pre-PMd. Anatomically, this region of the caudate was shown to be functionally connected to the left pre-PMd in our previous study (Badre et al. 2010), suggesting that the disconnection in these subjects may be related to dysfunction in the cortico-striato-thalamic loops

linking them (Alexander et al. 1986). These subjects thus reinforce the idea that dysfunction in brain regions is the ultimate determinant of behavior, and that this dysfunction can be mediated via injury to remote anatomical sites—that is, via diaschisis (Monakow 1914; Finger et al. 2004; Nomura et al. 2010). Moreover, this finding only held for the relevant task structure—that is, the Hierarchical case.

In this context, a final prediction of our model is that reduced functional connectivity between left pre-PMd and super-/subordinate lateral cortical regions should only correlate inversely with terminal accuracy when performance is suboptimal. In other words, the tendency to explore a reduced portion of the rule space should correlate inversely with engagement of left pre-PMd. However, this same argument suggests that learning the "full" rule space should be positively, not negatively, correlated with functional connectivity of left pre-PMd, as seen by others in the execution of higher order rules (Koechlin et al. 2003; Kouneiher et al. 2009). To address this question, we reanalyzed task-related functional MRI data from our previous study (Badre et al. 2010), in which many healthy subjects successfully learned the full rule structure. As predicted, in this case, greater connection of left pre-PMd within the hierarchy during task performance showed a positive, not negative, correlation with terminal accuracy at trend significance at the beginning of learning ($r = 0.39$, $P = 0.087$) that reached significance by the end of learning ($r = 0.49$, $P = 0.03$; Fig. 7*A*). As expected, this relationship did not hold for the Flat condition (Fig. 7*B*).

One final question concerns the relatively poor performance of the subjects with lesions outside left pre-PMd. Given the variable nature of the lesions in other subjects, there are likely to be multiple explanations. Consistent with theories in which there are multiple behavioral "controllers" subject to reinforcement learning, for example (Doya et al. 2002; Holroyd and Coles 2002; Frank and Badre 2011), it has been suggested that motivational processes in hierarchically-organized regions of the medial prefrontal cortex "energize" processing in corresponding areas of the lateral prefrontal cortex (Kouneiher et al. 2009). Additionally, striatal regions are thought to acquire first-order stimulus–response associations over longer time periods (Houk and Wise 1995; O'Reilly et al. 2007; Grahn et al. 2009), and reward-related striatal and ventral prefrontal regions may mediate learning from unexpected outcomes (Schoenbaum et al. 2009; Takahashi et al. 2009). Subjects with lesions affecting these areas may therefore be doubly disadvantaged in their efforts to search the S–R space, in that they confront the full search space with impaired search mechanisms. More broadly, these findings reinforce the idea that lesions in other sites can disrupt component processes supported by the networks in which the left pre-PMd is embedded. While left pre-PMd may be preferentially engaged in second-order learning, it is "specialized" for this function only in the context of a network of active brain regions. Further work to test lesion subgroups will clearly be important to better define how these other lesions impact both learning and network behavior.

In summary, these results are consistent with the existence of a rostrocaudal hierarchical organization within lateral frontal cortex that supports learning at various levels of abstraction. In keeping with our previous work (Badre et al. 2009; Badre et al. 2010), lesions of the left pre-PMd, implicated in the discovery of second-order rule structures, disrupt learning of second-order rules. Importantly, the nature of this disruption depends critically on task structure. Because the Hierarchical condition includes 2 simpler rules that ignore one feature (shape and orientation, respectively), this disruption improves performance in subjects relative to those with lesions elsewhere in the frontal cortex and basal ganglia, although overall learning remains impaired. In addition to operationalizing concepts that subjects with such lesions can be more concrete or stimulus bound (Devinsky and D'Esposito 2004), these findings suggest that under some conditions, a restricted rule space can be relatively advantageous. More generally, they emphasize that the impact of lesions on behavior can be modified by task structure and context—a concept ultimately important in using this knowledge to advance rehabilitation efforts for these and similar subjects.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/

Funding

The State of California to A.S.K.; the National Institutes of Health (grants MH63901 and NS40813 to M.D.).

Notes

We thank J. Hoffman for assistance with data collection, D. Badre and M. Frank for helpful discussions of preliminary results, E. Nomura and C. Gratton for access to resting state data, R. Knight, J. Black, and D. Scabini for assistance with subject recruitment, and the subjects for their participation. *Conflict of Interest*: None declared.

References

Alexander GE, DeLong MR, Strick PL. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci. 9:357–381.

Badre D. 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn Sci. 12:193–200.

Badre D, D'Esposito M. 2007. Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. J Cogn Neurosci. 19:2082–2099.

Badre D, D'Esposito M. 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? Nat Rev Neurosci. 10:659–669.

Badre D, Hoffman J, Cooney JW, D'Esposito M. 2009. Hierarchical cognitive control deficits following damage to the human frontal lobe. Nat Neurosci. 12:515–522.

Badre D, Kayser AS, D'Esposito M. 2010. Frontal cortex and the discovery of abstract action rules. Neuron. 66:315–326.

Bor D, Owen AM. 2007. A common prefrontal-parietal network for mnemonic and mathematical recoding strategies within working memory. Cereb Cortex. 17:778–786.

Buckner RL. 2003. Functional-anatomic correlates of control processes in memory. J Neurosci. 23:3999–4004.

Bunge SA. 2004. How we use rules to select actions: a review of evidence from cognitive neuroscience. Cogn Affect Behav Neurosci. 4:564–579.

Chase WG, Simon HA. 1973. Perception in chess. Cognitive Psychology. 4:55–81.

Christoff K, Gabrieli JDE. 2000. The frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. Psychobiology. 28:168–186.

Christoff K, Keramatian K. 2007. Abstraction of mental representations: theoretical considerations and neuroscientific evidence. In: Bunge SA, Wallis JD, editors. Perspectives on rule-guided behavior. New York: Oxford University Press.

Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, Gabrieli JD. 2001. Rostrolateral prefrontal cortex involvement in relational integration during reasoning. Neuroimage. 14:1136–1149.
Courtney SM. 2004. Attention and cognitive control as emergent properties of information representation in working memory. Cogn Affect Behav Neurosci. 4:501–516.
Devinsky O, D'Esposito M. 2004. Neurology of cognitive and behavioral disorders. New York: Oxford University Press.
Doya K, Samejima K, Katagiri K, Kawato M. 2002. Multiple model-based reinforcement learning. Neural Comput. 14:1347–1369.
Finger S, Koehler PJ, Jagella C. 2004. The Monakow concept of diaschisis: origins and perspectives. Arch Neurol. 61:283–288.
Frank MJ, Badre D. 2011. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cereb Cortex. PMID: 21693490.
Gick ML, Holyoak KJ. 1980. Analogical problem solving. Cognitive Psychology. 12:306–355.
Gick ML, Holyoak KJ. 1983. Schema induction and analogical transfer. Cognitive Psychology. 15:1–38.
Goel V, Dolan RJ. 2004. Differential involvement of left prefrontal cortex in inductive and deductive reasoning. Cognition. 93:B109–B121.
Grahn JA, Parkinson JA, Owen AM. 2009. The role of the basal ganglia in learning and memory: neuropsychological studies. Behav Brain Res. 199:53–60.
Graybiel AM. 1998. The basal ganglia and chunking of action repertoires. Neurobiol Learn Mem. 70:119–136.
Holroyd CB, Coles MG. 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev. 109:679–709.
Houk JC, Wise SP. 1995. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. Cereb Cortex. 5:95–110.
Kayser AS, Sun FT, D'Esposito M. 2009. A comparison of Granger causality and coherency in fMRI-based analysis of the motor system. Hum Brain Mapp. 30:3475–3494.
Koechlin E, Jubault T. 2006. Broca's area and the hierarchical organization of human behavior. Neuron. 50:963–974.
Koechlin E, Ody C, Kouneiher F. 2003. The architecture of cognitive control in the human prefrontal cortex. Science. 302:1181–1185.
Koechlin E, Summerfield C. 2007. An information theoretical approach to prefrontal executive function. Trends Cogn Sci. 11:229–235.
Kouneiher F, Charron S, Koechlin E. 2009. Motivation and cognitive control in the human prefrontal cortex. Nat Neurosci. 12:939–945.
Monakow Cv. 1914. Die lokalisation im grosshirn und der abbau der funktion durch kortikale herde. Wiesbaden (Germany): J. F. Bergmann.
Newell A. 1990. Unified theories of cognition. Cambridge (MA): Harvard University Press.
Nomura EM, Gratton C, Visser RM, Kayser A, Perez F, D'Esposito M. 2010. Double dissociation of two cognitive control networks in patients with focal brain lesions. Proc Natl Acad Sci U S A. 107:12017–12022.
O'Reilly RC, Frank MJ, Hazy TE, Watz B. 2007. PVLV: the primary value and learned value Pavlovian learning algorithm. Behav Neurosci. 121:31–49.
Petrides M. 2006. The rostro-caudal axis of cognitive control processing within lateral frontal cortex. In: Dehaene S, Duhamel J-R, Hauser MD, Rizzolatti G, editors. From monkey brain to human brain: a Fyssen foundation symposium. Cambridge (MA): MIT Press. p. 293–314.
Picard N, Strick PL. 2001. Imaging the premotor areas. Curr Opin Neurobiol. 11:663–672.
Reverberi C, Shallice T, D'Agostini S, Skrap M, Bonatti LL. 2009. Cortical bases of elementary deductive reasoning: inference, memory, and metadeduction. Neuropsychologia. 47:1107–1116.
Robin N, Holyoak KJ. 1995. Relational complexity and the functions of prefrontal cortex. In: Gazzaniga MS, editor. The cognitive neurosciences. Cambridge (MA): MIT Press. p. 987–997.
Rosenberg JR, Amjad AM, Breeze P, Brillinger DR, Halliday DM. 1989. The Fourier approach to the identification of functional coupling between neuronal spike trains. Prog Biophys Mol Biol. 53:1–31.
Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK. 2009. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nat Rev Neurosci. 10:885–892.
Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, Graybiel AM, Suzuki WA, Brown EN. 2004. Dynamic analysis of learning in behavioral experiments. J Neurosci. 24:447–461.
Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. 2009. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. Neuron. 62:269–280.

Supplementary Material

This supplement consists of detailed descriptions of neuropsychological, counterbalancing, and demographic data, and two figures (Figures S1 and S2) depicting the locations of the frontal and basal ganglia lesions for each of the 18 subjects and a conjunction analysis, respectively. Figures S1 and S2 relate to figure 2 in the main text, which includes a summary image.

Neuropsychological Data

Twelve of our subjects, including three subjects with left pre-PMd lesions, underwent the following focused neuropsychological battery: the Wide Range Achievement Test (WRAT-4), a measure of general reading, spelling, and math performance; the Wechsler Adult Intelligence Scale – Revised digit symbol subtest (WAIS-R), a test of complex visual attention; the Stroop-color and Stroop-interference tasks, the former more sensitive to simple visual attention and the latter to executive dysfunction; and Trails A & B, a test sensitive to task switching performance. No significant differences were seen in the normed WRAT scores (96.3 versus 97.7; $T(9) = 0.19$, $p = 0.85$) or in any of the WRAT subscales (all p's > 0.6), arguing that general performance differences did not distinguish our lesion subgroups. Differences in the WAIS-R digit symbol test were likewise non-significant (raw scores 48 versus 39; $T(10) = 1.1$, $p = 0.29$), suggesting that differences in task performance could not be easily explained by deficits in complex visual attention. Stroop – Color scores were significantly better in patients without pPMd lesions (37.3

versus 82.7; $T(9) = -4.2$, $p = 0.002$) but these differences were not found in the Stroop – Interference scores thought to more heavily tax executive function (30.3 versus 43.3; $T(7) = -1.5$, $p = 0.19$).

Task switching, as assessed by Trails A & B testing, was a particularly important part of these analyses. Specifically, the failure of left pre-PMd subjects to fully explore the rule space could have resulted from a failure to switch between different task sets in the Hierarchical condition – i.e. to switch from a shape- to an orientation-based rule set. If so, one would expect these subjects to perform more poorly on independent tasks of task switching. However, left pre-PMd subjects actually showed an overall tendency for better task switching performance (Trails B – Trails A) than other subjects (57.3 seconds versus 120.6 seconds; $T(9) = 2.1$, $p = 0.06$). On closer inspection, other-lesion subjects showed a bimodal distribution, in which 5 of 9 showed differences greater than 100 seconds and 4 of 9 performed comparably to left pre-PMd subjects (range 34-62 seconds). Thus, these results argue against an explanation that relies solely on task-switching performance or other baseline differences in cognitive processes such as complex visual attention.

Task Counterbalancing

Although there was no main effect of the order in which the tasks were performed (i.e. Hierarchical rule first versus second: $F(1,14) = 0.68$, $p = 0.42$ (ns)), a borderline interaction between lesion location and the order of the task was evident ($F(1,14) = 4.5$, p

= 0.051). Because of the potential interaction with order of presentation, we compared the difference between terminal accuracies in the Hierarchical and Flat conditions for those subjects who learned the Hierarchical rule second (3 subjects in the left pre-PMd group, 5 subjects in the other lesion group). Despite the smaller numbers of subjects, the left pre-PMd group again showed a trend toward better performance even when rule order was directly addressed (difference = 0.088 versus 0.013, $T(4) = 2.45$, $p = 0.075$). (We did not evaluate this same result for the case in which the Hierarchical rule was first, as only one of the left pre-PMd subjects met this criterion.)

Demographic Data

Differences were seen between the left pre-PMd and other-lesion patients in two of our demographic variables. Lesion size was significantly greater in the left pre-PMd patients (164 cc versus 67 cc; $T(7) = 3.1$, $p = 0.02$) and the number of years of education was also greater (18.5 years versus 13.9 years; $T(6) = 4.0$, $p = 0.007$). There were no significant differences in age (56.3 years old versus 63.2 years old; $T(4) = -1.0$, $p = 0.36$) or in time since lesion onset (10.1 years versus 11.7 years; $T(15) = -0.54$, $p = 0.60$).

Education is potentially the more critical confound. Hypothetically, education might lead to general improvements in cognition that enhance hierarchical performance, independent of lesion location and the aforementioned neuropsychological data. To directly assess this possibility, we divided our subjects by education level >= 16 years, the

lowest educational level attained by the subjects within our left pre-PMd group. However, this change rendered our results no longer significant. In particular, the differential terminal accuracy (Hierarchical minus Flat) for the more highly educated subjects did not differ from the more poorly educated group (mean difference 0.005 versus 0.002; $T(8) = 0.2$, $p = 0.84$). If, instead of relying on a somewhat arbitrary division of groups, we correlated educational level directly with differential terminal accuracy, the result was also non-significant ($r = 0.29$, $p = 0.25$), as it was for separate correlations with Hierarchical ($r = 0.42$, $p = 0.09$) and Flat ($r = 0.04$, $p = 0.88$) terminal accuracy alone. Thus, we do not believe that educational level in itself can explain our data.

With respect to lesion size, we note that the effect would be somewhat counterintuitive: i.e. that a larger lesion, irrespective of location, produces better, not worse, performance in this demanding task. Nonetheless, if this hypothesis is correct, lesion size independent of the frontal regions involved could drive our results. As above, we divided our subjects into groups based on a lesion size >= 122 cc, the smallest lesion found in our left pre-PMd group. In so doing, we eliminated the strongly significant ($p = 0.001$) effect for differential terminal accuracy (mean difference 0.04 versus 0.00; $T(8) = 2.2$, $p = 0.06$). To avoid relying on an arbitrary division of groups, we also correlated lesion size directly with differential terminal accuracy ($r = 0.40$, $p = 0.10$), Hierarchical terminal accuracy ($r = 0.27$, $p = 0.27$), and Flat terminal accuracy ($r = -0.3$, $p = 0.24$); and obtained non-significant results. These findings are perhaps less surprising if we note, for example, that

the largest lesion belongs to a non-prePMd patient (subject #8) who performs somewhat poorly (figure 3). Thus, we do not believe that lesion size in itself can explain our data.

Figure Legends

Figure S1: The locations of the lesions, indicated in red, for each of the 18 subjects. The left side of each image corresponds to the left side of the brain.

Figure S2: A conjunction map showing the lesioned areas, indicated in red, shared by all 4 left pre-PMd subjects but not found in any of the other 14 subjects.

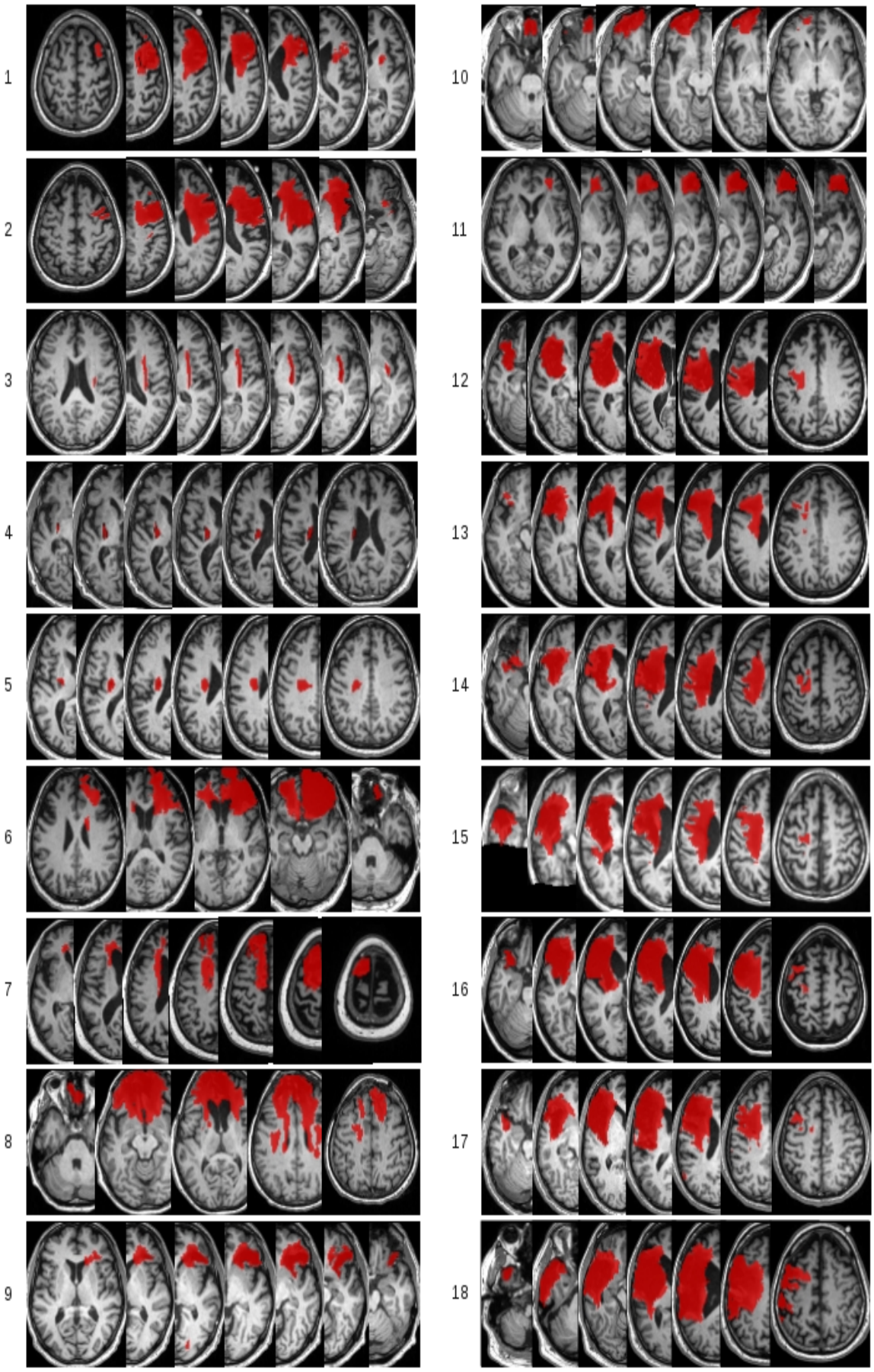
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

Figure S2