

Transient Activation of Superior Prefrontal Cortex during Inhibition of Cognitive Set

Seiki Konishi, Koji Jimura, Tomoki Asari, and Yasushi Miyashita

Department of Physiology, The University of Tokyo School of Medicine, Tokyo 113, Japan

The prefrontal cortex implements a set-shifting function that includes inhibition of a previously acquired cognitive set. The impairment of the inhibitory function results in perseverative behavior that forms one characteristic feature of frontal lobe dysfunction. Previous neuroimaging studies have revealed inhibitory mechanisms in the inferior prefrontal cortex. The present functional magnetic resonance imaging study devised “dual-match” stimuli in a set-shifting paradigm that allowed us to temporally isolate the inhibitory processes recruited during exposure to a previously acquired set. Transient activation time-locked to the isolated inhibition was revealed in the left middle frontal gyrus near the superior frontal sulcus. In a control experiment conducted after subjects had been informed and made aware of the exposure, however, the superior prefrontal activation disappeared, and prominent activation was revealed in a set of brain regions that included the left posterior inferior frontal sulcus. These double dissociation results indicate inhibitory mechanisms in the superior prefrontal cortex, alternative to the inferior prefrontal ones, that are activated depending on the subjects’ strategy for inhibition of cognitive set.

Key words: prefrontal; fMRI; set; shifting; inhibition; human

Introduction

The prefrontal cortex permits a flexible transition between competing behavioral patterns in accordance with changing environments, especially by inhibiting a mental set that supported previously appropriate behavior. Such inhibitory function manifests itself most typically in set-shifting paradigms such as the Wisconsin card sorting test (WCST), during which one behavioral pattern based on a particular principle, or a “dimension,” that is maintained for a prolonged period is updated to another at the time of dimensional changes. Previous neuropsychological findings for both humans (Milner, 1963; Owen et al., 1993; Rogers et al., 1998) and monkeys (Passingham, 1972; Dias et al., 1996) have shown that patients with damage to the lateral frontal cortex characteristically adhere to previously valid behavior, which has been attributed to impairment of the inhibition of perseverative interference from a previously acquired set. The lateral frontal involvement in set shifting has been supported by a number of neuroimaging studies demonstrating prominent activations in the lateral prefrontal cortex, especially in the inferior prefrontal cortex (Dove et al., 2000; Monchi et al., 2001; Konishi et al., 2002a; Nakahara et al., 2002).

Activation associated with set shifting has been investigated by assessing signal increase at the time of dimensional changes in set-shifting paradigms; however, dimensional changes might well

involve multiple processes other than inhibitory ones (Owen et al., 1991; Delis et al., 1992). In particular, an important distinction regarding component processes related to set shifting has been proposed between shifting from previous dimensions and shifting to new dimensions (Owen et al., 1993), both of which may inevitably be recruited at the time of dimensional changes. To isolate neural correlates of the inhibitory mechanism, a modified card-sorting task was devised in the present functional magnetic resonance imaging (fMRI) study by introducing “dual-match” stimuli in card-sorting trials (see Fig. 1*a*). The dual-match stimuli allowed us to temporally separate inhibitory processes away from dimensional changes during which multiple processes related to set shifting are recruited (see Fig. 1*b*). The dimensional change was signaled to subjects by visual presentation of a subsequent dimension (Konishi et al., 1999b, 2002a). When the dual-match stimulus was presented immediately after the dimensional change, the subjects reconfigured a new task set but were not required to inhibit a previous set acquired before the dimensional change, because new correct matching in the dual-match trials was compatible with previously correct matching. In a “single-match” trial presented immediately after a series of the dual-match trials (an “inhibition trial”), however, the subjects were now exposed to the previous set and were required to inhibit it to choose a correct answer. The present study directly analyzed the inhibition trials in which inhibition was required away from the dimensional changes (experiment 1). One important feature of the inhibition trials would be that the subjects were less aware of the dimensional changes and therefore of the exposure to a previous set at the inhibition trials, which might result in changes in the inhibitory mechanisms. Experiment 2 was conducted to control this possible effect after the subjects had been informed and made aware of the exposure.

Received April 9, 2003; revised July 7, 2003; accepted July 7, 2003.

This work was supported by a grant-in-aid for Specially Promoted Research (14002005) to Y.M. and a grant (14780598) to S.K. from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Correspondence should be addressed to either of the following: Dr. Seiki Konishi, Department of Physiology, The University of Tokyo School of Medicine, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan, E-mail: konishi@m.u-tokyo.ac.jp; or Prof. Yasushi Miyashita, Department of Physiology, The University of Tokyo School of Medicine, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan, E-mail: yasushi_miyashita@m.u-tokyo.ac.jp.

Copyright © 2003 Society for Neuroscience 0270-6474/03/237776-07\$15.00/0

Materials and Methods

Subjects and fMRI procedures. Informed consent was obtained from 36 healthy right-handed subjects (18 males, 18 females; age, 20–31 years). They were scanned by fMRI using experimental procedures approved by the institutional review board of the University of Tokyo School of Medicine. Thirty-six subjects underwent fMRI scanning without being informed of the exposure to the interference from previous task set (experiment 1). Sixteen of these subjects were additionally scanned performing the task after having been made aware of the exposure (experiment 2).

Scanning was conducted using a 1.5 T fMRI system. Scout images were first collected to align the field of view centered on the subject's brain. Then T2-weighted spin-echo images were obtained for anatomical reference [repetition time (TR) = 5.5 sec; echo time (TE) = 30 msec; 75 slices, slice thickness = 2 mm; in-plane resolution = 2×2 mm]. For functional imaging, gradient echo echo-planar sequences were used (TR = 4 sec; TE = 50 msec; flip angle = 90°). Each functional run consisted of 68 whole-brain acquisitions (28×4 mm slices; in-plane resolution 4 mm). The first four functional images in each run were excluded from the analysis to take into account the equilibrium of longitudinal magnetization.

Behavioral procedures. Visual stimuli were presented to subjects by projecting the stimuli onto a screen. Subjects viewed the screen through prism glasses. A magnet-compatible button press on the basis of a fiber-optic switch was used to record subjects' performances. The tasks used in this study were derived from the WCST (Grant and Berg, 1948) modified in our previous studies (Konishi et al., 1999b, 2002a). In each WCST trial, a five-card stimulus was presented until subjects responded to one of four reference card stimuli at the corner of the screen by matching the attribute of a central card on the basis of the dimension of color, form, or number. A four-channel button was pressed using the right thumb for the choice of one of the four reference card stimuli. A feedback stimulus (right, O; wrong, X) was then presented. After eight or more successive correct trials, the currently relevant dimension was changed to one of the others, and subjects were instructed regarding the subsequent dimension by visual presentation of the word "color," "form," or "number."

There were two forms of the five-card stimuli: single match and dual match (Fig. 1*a*). The single-match stimulus was used in most of the card-sorting trials, and each one of its four reference card stimuli was matched to its central card stimulus on the basis of only one of the three dimensions (Fig. 1*a*, left). On the other hand, the dual-match stimulus, which was presented immediately after dimensional changes, contained a reference card stimulus that was matched to the central card stimulus on the basis of two of the dimensions (Fig. 1*a*, right). When the dimension was changed from color to form, as in Figure 1*b*, subjects were required to reconfigure a new task set but not to inhibit a previous set in the dual-match trials, because matching to the dual-match stimuli on the basis of a previous dimension (color matching) still led to correct matching on the basis of a current dimension (form matching). After successive dual-match trials, the single-match trials followed, and a previous set that is known to last long after task switching (Allport et al., 1994; Meiran et al., 2000) was now required to be inhibited (i.e., inhibition was required independently of dimensional changes) because matching to the single-match stimuli on the basis of a previous dimension no longer led to correct matching. Thus the present design temporally isolated the inhibition trials (the first single-match trial after a series of the dual-match trials) and avoided the contamination of various processes that may occur at the time of dimensional changes (Owen et al., 1991; Delis et al., 1992).

The number of dual-match trials presented after dimensional changes was randomly selected from two and four, and therefore the single-match trials were presented at either the third trials (~ 4 sec) or the fifth trials (~ 8 sec) after the dimensional changes. No dual-match stimuli were presented after the inhibition trials. As a positive control condition, after one-third of the dimensional changes, no dual-match trials were presented (i.e., single-match trials were presented instead) after dimensional changes ("original" dimensional changes). The task used a self-paced design, and the instruction stimuli were presented for 0.5 sec, with each stimulus separated by a blank image for 0.25 sec (therefore, the time

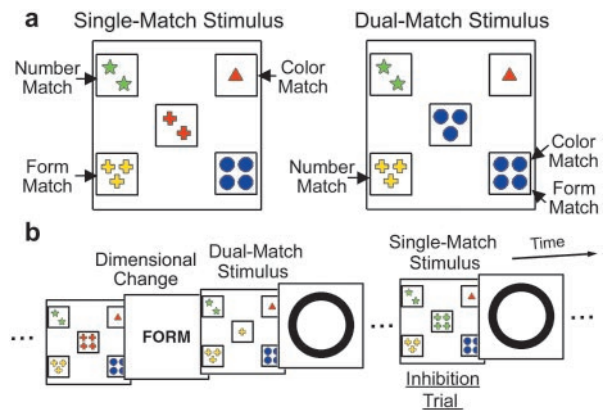


Figure 1. A modified WCST used in the present study. *a*, Single-match (left) and dual-match (right) stimuli. *b*, Presented stimuli (card, feedback, and instruction) shown in temporal order. In this figure the original dimension is color, and the dimension is changed into form. In the inhibition trials after the dual-match trials, subjects were exposed to perseverative interference from a previously acquired set.

between response and presentation of the next trial was 1.0 sec). Eight experimental runs were collected for each subject that contained 106.3 ± 10.4 (mean \pm SD) dimensional changes in total depending on subjects' performance.

A post-scanning interview was conducted in experiment 1 to examine to what degree the subjects were unaware of the exposure to the previously acquired set. The dual-match nature of the card stimuli and the exposure to the previous set at the inhibition trials were explained to the subjects, and they were asked whether they noticed them. Only those subjects who answered strongly in the negative were regarded as having been totally unaware of them. Those subjects who reported awareness of any aspect of them were regarded as having been partially unaware of them, although they reported that they performed the task indifferently to them. It should be noted that this indifferent performance of the task contrasts with the situation in experiment 2, in which subjects were made aware of them and a distinct pattern of activation was obtained (see Results). The procedure might underestimate the number of subjects who performed the task totally unaware of them but would certify the form of inhibitory processes recruited in that situation. Sixteen of the 36 subjects were additionally scanned in a control experiment (experiment 2) after they had been informed and made aware of the exposure. More specifically, the dual-match nature of the card stimuli and the exposure to the previous set at the inhibition trials were explained to the subjects, and they were instructed to try not to make an error in the critical inhibition trials.

Data analysis. Data were analyzed using SPM99 (<http://www.fil.ion.ucl.ac.uk/spm/>). Functional images were realigned, slice timing was corrected, normalized to the default template with interpolation to a $2 \times 2 \times 2$ mm space, and spatially smoothed (full width, half maximum = 6 mm). Then event timing was coded into a general linear model (GLM) (Friston et al., 1994; Worsley and Friston, 1995). The activation of interest in the present study, the first single-match trials after dimensional changes in which the inhibition was involved, were coded into a GLM using the canonical hemodynamic response function in SPM99, time-locked to the onset of stimulus presentation (Fig. 1*b*); however, it is possible that residual activation derived from the dimensional changes, although separated by ~ 4 and 8 sec (1 trial, ~ 2 sec), contaminated the activation related to the inhibition. To cancel out this possible effect, the third or fifth trials after the dimensional changes were defined as control trials when the inhibition trials were presented at the fifth or third trials after the same dimensional changes, respectively. These temporally equivalent control trials, together with the dimensional changes and error trials, were also coded into a GLM. Images of parameter estimates for signal response magnitudes in these events were then analyzed for group analysis using a random-effect model. Peak coordinate locations in activation maps were generated using a threshold of 19 or more contiguous

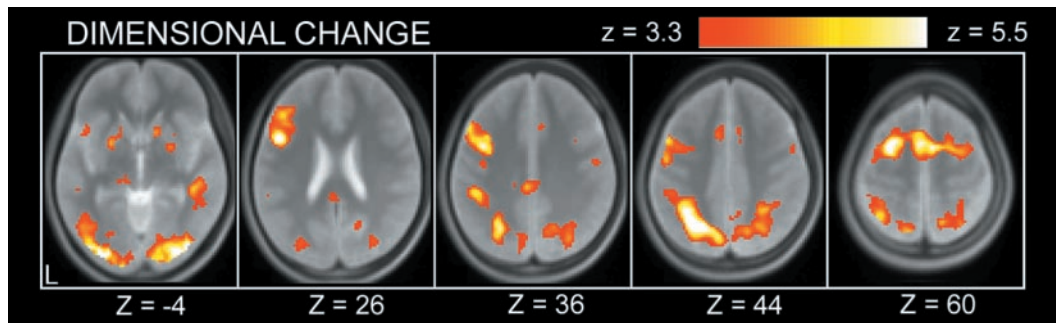


Figure 2. Statistical activation maps for signal increase and decrease during original dimensional changes in experiment 1. The color scale in the maps reflects statistical significance, using the threshold of $z > 3.3, p < 0.001$ (uncorrected) for a display purpose. Activation maps are displayed as transverse sections and overlaid on the anatomic image averaged across subjects. The transverse section level is indicated by the Z coordinates of Talairach space at the bottom.

significant voxels above $p < 0.001$ ($z > 3.3$) (each voxel, $2 \times 2 \times 2 \text{ mm}^3$), calculated using an empirical analysis of control data set (Buckner et al., 1998; Konishi et al., 2001). For critical activations in the superior and inferior prefrontal cortex highlighted in Results, $p < 0.05$ corrected was cleared.

Results

Mean correct performance in experiment 1 was 98.1% in the inhibition trials (presented at the third and fifth trials after the dimensional changes) and 99.7% in the temporally equivalent control trials (presented at the fifth and third trials after the same dimensional changes, respectively; see Materials and Methods). This performance difference was significant ($1.6 \pm 2.0\%$; mean \pm SD; $t_{(35)} = 4.9$; $p < 0.001$). Mean reaction time in correct trials was 1074.8 msec in the inhibition trials and 1008.2 msec in the temporally equivalent control trials; this increase in mean reaction time was also significant (66.6 ± 53.6 msec; mean \pm SD; $t_{(35)} = 7.5$; $p < .001$). A post-scanning interview confirmed that 17 of the 36 subjects were totally unaware of the exposure to interference from the previous set at the inhibition trials, whereas the remaining 19 subjects reported general indifference to it. When the reaction time analysis was restricted to these 17 subjects, the performance difference ($1.5 \pm 2.4\%$; mean \pm SD; $t_{(16)} = 2.6$; $p < 0.05$) and the reaction time increase (51.6 ± 34.4 msec; mean \pm SD; $t_{(16)} = 6.2$; $p < .001$) were similarly significant. These behavioral results suggest that inhibitory processes were recruited at the inhibition trials, even when the subjects were unaware of the exposure.

The image data set from a pool of the 36 subjects was analyzed by a GLM implemented in SPM99 and was applied to a random effect model. A positive control condition was included in the dimensional changes after which no dual-match trials were presented. As shown in Figure 2, the dimensional changes elicited prominent activations in a set of mostly left-lateralized regions, including the posterior inferior frontal sulcus regions, consistent with our previous study using a similar modification of the WCST (Konishi et al., 2002a). More central to the activation of interest in the present study, at the inhibition trials, prominent activation was revealed in the left middle frontal gyrus near the superior frontal sulcus (Fig. 3, top). It is possible that this activation is contaminated by residual activation derived from the dimensional changes two or four trials apart. This possibility was canceled by calculating activation in the temporally equivalent control trials (Fig. 3, middle) and comparing it with that in the inhibition trials. Although some of the activations were weakened, the superior prefrontal activation remained as significant (Fig. 3, bottom), which is consistent with the observation that this

region was not significantly activated at the dimensional changes ($p > 0.05$) (Fig. 2). The complete list of the activation results is shown in Table 1. Note that the anatomical label and Brodmann area, including those for the superior prefrontal activation, should be considered approximate.

The superior prefrontal activation at $(-32, 32, 44)$ (Table 1) was investigated further by dividing the inhibition trials from two independent sources: the third and fifth trials after the dimensional changes. The percentage signal in the inhibition trials was 0.139% ($t_{(35)} = 5.7$; $p < 0.001$; correction unnecessary) in the third trials and 0.105% ($t_{(35)} = 3.8$; $p < 0.001$) in the fifth trials. Moreover, the superior prefrontal activation was similarly significant in the 17 (of the 36) subjects who were totally unaware of the exposure to previous set (0.132%, $t_{(16)} = 3.5$, $p < 0.005$ in the third trials; 0.095%, $t_{(16)} = 3.6$, $p < 0.005$ in the fifth trials), confirming the reliability of the superior prefrontal activation in these two data sets, even in the totally unaware group only.

In experiment 2, 16 of the 36 subjects were additionally scanned as a positive control after they had been informed and made aware of the exposure. Mean correct performance was 97.9% in the inhibition trials and 99.8% in the temporally equivalent control trials, and the performance difference was significant ($1.9 \pm 2.3\%$; mean \pm SD; $t_{(15)} = 3.3$; $p < 0.005$). Mean reaction time in correct trials was 1370.7 msec in the inhibition trials and 1090.7 msec in the temporally equivalent control trials; this mean reaction time increase was also significant (280.0 ± 160.6 msec; mean \pm SD; $t_{(15)} = 7.0$; $p < 0.001$). The transient activation related to the inhibition trials was revealed in a set of brain regions that included the posterior inferior frontal sulcus region (Fig. 4a, Table 2) and appears relatively similar to the activation pattern during the original dimensional changes shown in Figure 2, where similar inhibition was involved. The superior prefrontal region that was prominently activated at the inhibition trials in experiment 1, on the other hand, was not significantly activated ($p > 0.05$).

The contrast of central interest, “inhibition minus control,” was directly compared between experiments 1 and 2, and the differential activation is presented in Figure 4b. Only the superior prefrontal region was activated significantly more in experiment 1 than in experiment 2 (i.e., the superior prefrontal peak at Table 1 was included in the significant activation shown in Fig. 4b), whereas the opposite differential activation was observed in the posterior inferior frontal sulcus region, as well as many of the regions presented in Figure 4a (i.e., the inferior prefrontal peak at Table 2 was included in the significant activation shown in Fig. 4b).

The signal difference between experiments 1 and 2 was further

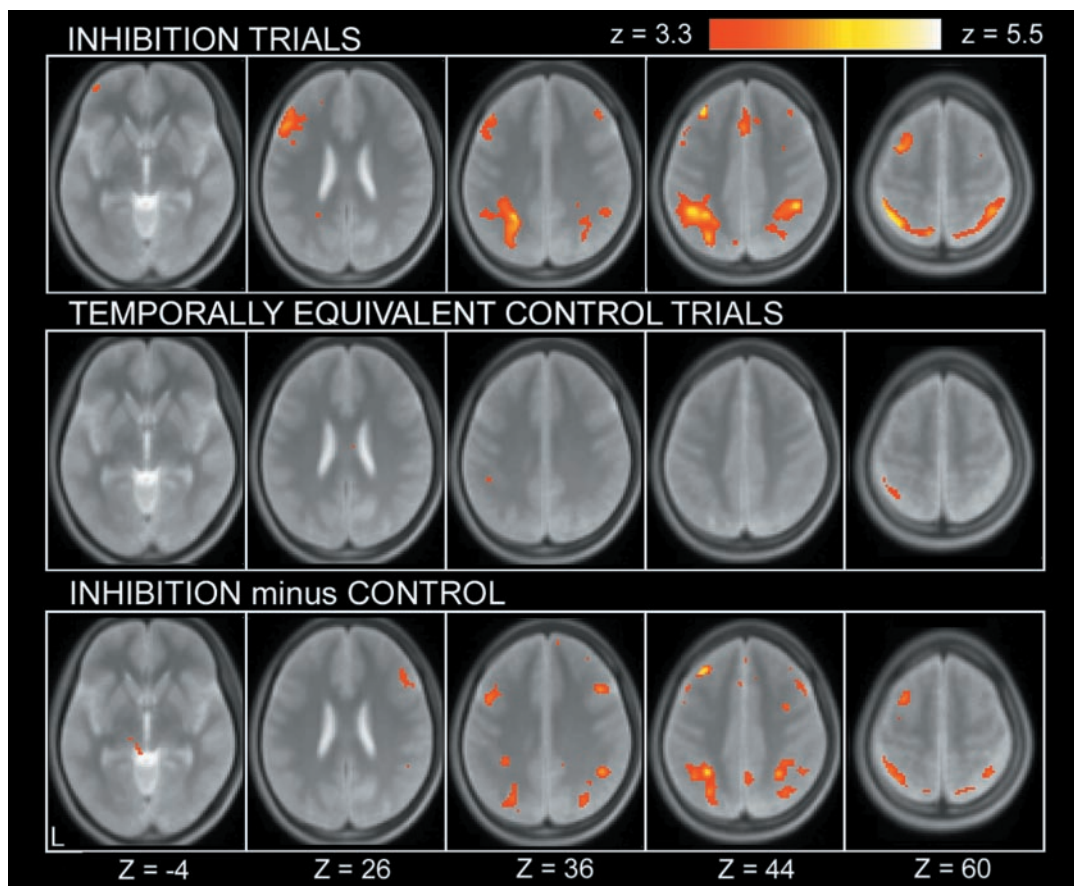


Figure 3. Statistical activation maps for signal increase during inhibition trials (top), temporally equivalent control trials (middle), and the difference (bottom) in experiment 1. Format is similar to Figure 2.

Table 1. Brain regions showing signal increase in inhibition minus control in experiment 1

	Coordinates			z value	BA/Area
	x	y	z		
Lateral frontal cortex	-32	32	44	5.14*	8/9
	-42	16	32	4.97	9
	44	18	38	4.65	9
	-30	6	60	4.26	6
	38	34	42	3.87	8/9
	-42	36	12	3.55	46
Medial frontal cortex	-2	18	48	3.94	6/8
Parietal cortex	-28	-52	42	5.04*	40/7
	30	-64	54	4.66	7
	34	-44	42	4.58	40/7
	-28	-68	44	4.36	7
	26	-56	68	4.20	7
Occipital cortex	4	-56	46	4.11	7
	26	-50	0	4.31	19/18
	30	-76	34	4.08	19
Others	-10	8	8	4.29	Caudate

Coordinates are listed in the Talairach space (Talairach and Tournoux, 1988) with negative values of x on the left. BA is the Brodmann area near the coordinates and is approximate. Significant activation was detected using a threshold of ≥ 19 contiguous significant voxels above $p < 0.001$. * $p < 0.05$ corrected for whole-brain multiple comparison.

tested for each of the inferior and superior regions using independent data sets. For the inferior prefrontal region, coordinates were represented on the basis of our previous studies of inhibitory control (Konishi et al., 2001, their Table 2) at (-44, 16.5, 23), and the coordinates were then applied to the present data set. The difference between experiments 2 and 1 was significant ($t_{(50)} =$

2.7; $p < 0.01$). As for the superior prefrontal region, because of the lack of previous knowledge of its precise coordinates, a replication approach was used, wherein the data set of experiment 1 was divided into two: a hypothesis generating and a hypothesis testing (Buckner et al., 1996). The first data set from experiment 1 ($n = 18$) generated a peak at (-30, 36, 42) ($z = 3.8$), and the coordinates were then applied to the second hypothesis testing data set from experiment 1 ($n = 18$) and the whole data set of experiment 2. The difference between experiments 1 and 2 was significant ($t_{(32)} = 2.5$; $p < 0.05$). It should be noted that these analyses did not require whole-brain multiple comparisons because the two foci had already been shown to be significant at $p < 0.05$ (corrected) (Tables 1, 2).

A final analysis tested the possible reaction time effect to the differential activation pattern (Christoff et al., 2001). The reaction time increase for each subject for both experiments 1 and 2 was coded into a GLM, together with the differential effect, and examined in a group analysis using a random effect model. The significant signal difference was still present in both the superior ($t_{(49)} = 3.4$; $p < 0.005$, correction unnecessary) and inferior ($t_{(49)} = 2.3$; $p < 0.05$) prefrontal regions, whereas the reaction time effect was not significant in either of the regions ($p > 0.05$), confirming that the differential pattern does not simply reflect the reaction time difference.

Discussion

In the present study, the WCST was modified in a manner that required subjects to inhibit a previously acquired set temporally segregated from dimensional changes. Behavioral analysis re-

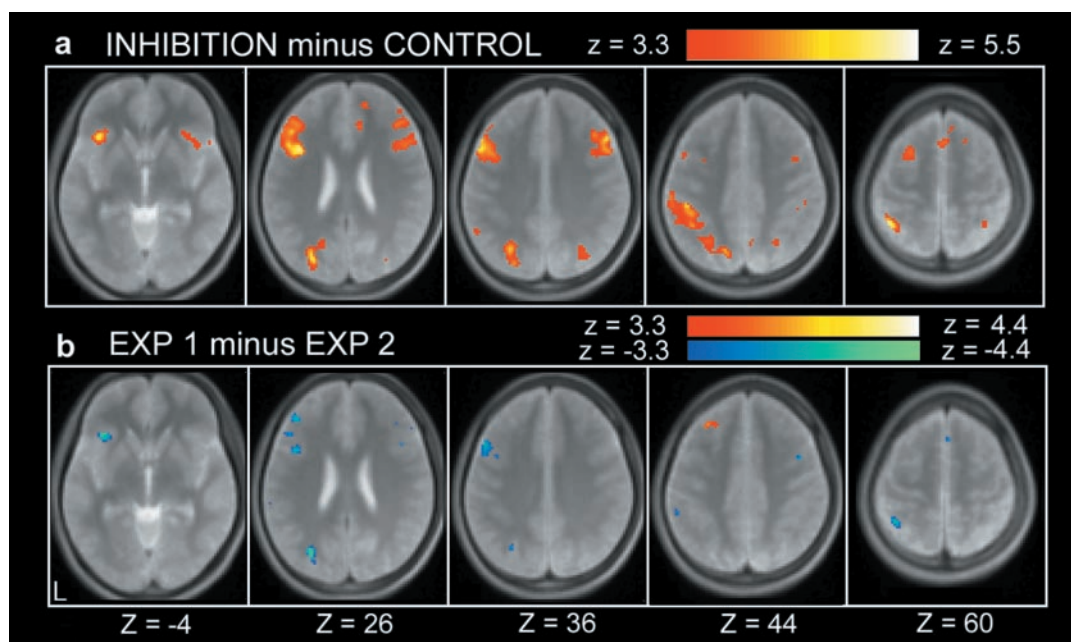


Figure 4. *a*, Statistical activation maps for signal increase and decrease during inhibition minus control trials in experiment 2. *b*, Direct comparison of the contrast inhibition minus control between experiments 1 and 2.

Table 2. Brain regions showing signal increase in inhibition minus control in experiment 2

	Coordinates			z value	BA/Area
	x	y	z		
Lateral frontal cortex	-38	14	26	5.27*	45/44
	-38	24	-4	5.10	47/12
	50	24	36	4.76	9
	38	4	48	4.49	6
	20	16	62	4.27	6
	34	18	-12	4.20	47/12
	-22	6	62	4.20	6
Medial frontal cortex	26	2	64	4.00	6
	4	18	58	4.42	6
	14	52	28	4.03	9
Parietal cortex	12	32	26	3.97	32
	-40	-52	60	5.21*	7/5
	-26	-64	50	5.18*	7
	-44	-34	44	4.67	40
	6	-68	48	4.67	7
	-54	-54	38	4.54	40
Occipital cortex	30	-56	54	4.32	40/7
	48	-36	54	3.87	40
	-26	-80	28	5.09	19
	36	-74	34	4.18	19
	-20	-62	18	3.55	18/17

Coordinates are listed in the Talairach space (Talairach and Tournoux, 1988) with negative values of x on the left. BA is the Brodmann area near the coordinates and is approximate. Significant activation was detected using a threshold of ≥ 19 contiguous significant voxels above $p < 0.001$. * $p < 0.05$ corrected for whole-brain multiple comparison.

vealed a robust reaction time increase at the inhibition trials. Transient activation elicited at the inhibition trials was observed in the left middle frontal gyrus near the superior frontal sulcus. A distinct pattern of activation was revealed on the inhibition trials, including the left posterior inferior frontal sulcus, after the subjects had been informed and made aware of the exposure to a previous set. These results suggest dissociable mechanisms within the prefrontal cortex that implement two forms of inhibition.

Activation in the superior part of the prefrontal cortex in general has rarely been observed in neuroimaging studies using a

wide range of cognitive tasks (Duncan and Owen, 2000). The superior prefrontal activation during inhibition reported in this study is located superior to the standard middle prefrontal activation during working memory tasks (Petrides et al., 1993; McCarthy et al., 1994; D'Esposito et al., 1995; Cohen et al., 1997; Courtney et al., 1997). The superior prefrontal activation is also different from the activation in the posterior part of the superior frontal sulcus during spatial working memory tasks (Courtney et al., 1998; Rowe et al., 2000). The previous studies using set-shifting paradigms also appear insignificant in the superior prefrontal activation (Berman et al., 1995; Nagahama et al., 1996; Barcelo et al., 1997; Konishi et al., 1998, 2002a; Dove et al., 2000; MacDonald et al., 2000; Rogers et al., 2000; Monchi et al., 2001; Dreher et al., 2002; Nakahara et al., 2002; Rushworth et al., 2002), consistent with the negative results in the original dimensional changes (see Materials and Methods) shown in Figure 2. These negative observations suggest the unique status of the superior prefrontal activation associated with a cognitive domain that has rarely been investigated. The left hemisphere dominance of the superior prefrontal activation, however, is common to left frontal dominance reported in the previous neuropsychological literature (Milner, 1971; Rogers et al., 1998) and to left-dominant prefrontal activation during set shifting (Konishi et al., 2002a), suggesting general left hemisphere dominance for inhibitory control of this type.

The superior prefrontal activation presented here contrasts with the activation in the posterior inferior frontal sulcus. This region conforms to the one reported in our previous studies of set shifting (Konishi et al., 2001, their Table 2) and is activated during the performance of cognitive tasks requiring executive control processes such as the Stroop task (Taylor et al., 1997; Chee et al., 2000), go/no-go tasks (Konishi et al., 1999a; Bunge et al., 2002; Durston et al., 2002), controlled retrieval tasks (Thompson-Schill et al., 1997; Dobbins et al., 2002; Gold and Buckner, 2002; Konishi et al., 2002b), and set-shifting tasks (Berman et al., 1995; Nagahama et al., 1996; Barcelo et al., 1997; Konishi et al., 1998, 2002a; Dove et al., 2000; MacDonald et al.,

2000; Rogers et al., 2000; Monchi et al., 2001; Dreher et al., 2002; Nakahara et al., 2002; Rushworth et al., 2002; Houde and Tzourio-Mazoyer, 2003), consistent with the executive roles expected for inhibition of cognitive set. One potential confound regarding this inferior prefrontal activation relates to the instruction given to the subjects in experiment 2 (see Materials and Methods) that may introduce unwanted perceptual or cognitive processes, but it is to be noted that the results in experiment 2 are consistent with many of prominent activations, including the posterior inferior frontal sulcus region, that have been observed during similar inhibitory processes cited above.

The set-shifting paradigm using the dual-match stimuli in the present study was designed to isolate the inhibitory processes that were originally recruited at the time of dimensional changes; however, quite importantly, the superior prefrontal activation elicited at the inhibition trials was not observed during dimensional changes in the original condition during which inhibition of the previous set was required (Fig. 2). The negative result suggests that although the superior prefrontal activation may indeed be related to inhibition of cognitive set, the form of underlying inhibitory processes was qualitatively changed when temporally segregated from dimensional changes. One obvious difference between these two events is that subjects were less aware of the exposure to the previous set at the inhibition trials than at the dimensional changes. To control the level of such awareness at the inhibition trials, experiment 2 was conducted, and clear double dissociation in the superior and inferior prefrontal activations was revealed. One straightforward explanation for the dissociation results is that the inferior prefrontal activation reflects an explicit form of inhibitory processes, whereas the superior prefrontal activation reflects an implicit form of inhibitory processes that might be similar to those recruited during spontaneous shifting (Rylander, 1939; Halstead, 1940; Goldstein, 1944), during which inhibition is required without subjects being aware of dimensional changes. Another possibility would be that the superior prefrontal region is activated only when no other alternative inhibitory mechanisms (i.e., inferior prefrontal ones) were involved that are to be activated during normative dimensional changes. Although definite interpretation of the superior prefrontal activation is not fully elucidated by the results of the present study, the double dissociation reported in the present study suggests the existence of another form of inhibitory mechanisms in the superior prefrontal cortex that had previously remained unrevealed.

The differential activation was detected by comparing experiments 1 and 2 scanned in the same subjects. It is unlikely that the differential activation can be explained only by learning effects, in light of previous literatures that have demonstrated that such effects on the WCST were indistinguishable between naive and practiced subjects (Berman et al., 1995). Moreover, the differential activation pattern during the inhibition trials across experiments 1 and 2 (Fig. 4b) was consistent with that observed within the data set of experiment 1: the inhibition trials (Fig. 3) and the original dimensional changes that involved inhibition similar to that in experiment 2 (Fig. 2). It is also unlikely that the superior prefrontal activation simply reflects increased involvement attributable to less facilitation in the inhibition trials than in the dual-match trials, because the task structure used in experiment 1 is common to that in experiment 2 in which no superior prefrontal activation was observed.

The dissociable neural substrates for the two forms of inhibition might be relevant to the debate regarding the precise critical lesion foci for inhibitory functions within the lateral prefrontal

cortex. As has been pointed out (Milner, 1964), there appears to be some variability of critical lesion foci among previous neuropsychological studies in terms of superior versus inferior prefrontal regions (Milner, 1963, 1964; Mishkin, 1964; Butter, 1969; Iversen and Mishkin, 1970; Passingham, 1972; Dias et al., 1996; Aron et al., 2003). One possibility might be related to the fact that, unlike most neuroimaging studies, neuropsychological studies use naive subjects and require only a small number of repetitions. Under such behavioral conditions, the subjects' strategy (the form of inhibitory processes to be recruited) would not be completely established, which might result in variability of recruited processes among subjects and even among trials in the same subjects. Although open to questions as to how these two forms of the function interact with each other in such naive subjects, the present study suggests the inhibitory mechanisms implemented in the superior prefrontal cortex, alternative to the inferior prefrontal ones, that contribute to our flexible behavior.

References

- Allport DA, Styles EA, Hsieh S (1994) Shifting intentional set: exploring the dynamic control of tasks. In: Attention and performance XV (Umiltà C, Moscovitch M, eds), pp 421–452. Cambridge, MA: MIT.
- Aron AR, Fletcher PC, Bullmore ET, Sahakian BJ, Robbins TW (2003) Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat Neurosci* 6:115–116.
- Barcelo F, Sanz M, Molina V, Rubia FJ (1997) The Wisconsin card sorting test and the assessment of frontal function: a validation study with event-related potentials. *Neuropsychologia* 35:399–408.
- Berman KF, Ostrem JL, Randolph C, Gold J, Goldberg TE, Coppola R, Carson RE, Herscovitch P, Weinberger DR (1995) Physiological activation of a cortical network during performance of the Wisconsin card sorting test: a positron emission tomography study. *Neuropsychologia* 33:1027–1046.
- Buckner RL, Raichle ME, Miezin FM, Petersen SE (1996) Functional anatomical studies of memory retrieval for auditory words and visual pictures. *J Neurosci* 16:6219–6235.
- Buckner RL, Goodman J, Burock M, Rotte M, Koutstaal W, Schacter DL, Rosen BR, Dale AM (1998) Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron* 20:285–296.
- Bunge SA, Dudukovic NM, Thomason ME, Vaidya CJ, Gabrieli JDE (2002) Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron* 33:301–311.
- Butter CM (1969) Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in *Macaca mulatta*. *Physiol Behav* 4:163–171.
- Chee MWL, Sriram N, Soon CS, Lee KM (2000) Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *NeuroReport* 11:135–140.
- Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, Gabrieli JDE (2001) Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage* 14:1136–1149.
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE (1997) Temporal dynamics of brain activation during a working memory task. *Nature* 386:604–608.
- Courtney SM, Ungerleider LG, Keil K, Haxby JV (1997) Transient and sustained activity in a distributed neural system for human working memory. *Nature* 386:608–611.
- Courtney SM, Petit L, Maisog JM, Ungerleider LG, Haxby JV (1998) An area specialized for spatial working memory in human frontal cortex. *Science* 279:1347–1351.
- Delis DC, Squire RL, Bihle A, Massman P (1992) Componential analysis of problem-solving ability: performance of patients with frontal lobe damage and amnesic patients on a new sorting test. *Neuropsychologia* 30:683–697.
- D'Esposito M, Detre JA, Alsop DC, Shin RK, Atlas A, Grossman M (1995) The neural basis of the central executive system of working memory. *Nature* 378:279–281.
- Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380:69–72.
- Dobbins IG, Foley H, Schacter DL, Wagner AD (2002) Executive control

- during episodic retrieval: multiple prefrontal processes subserve source memory. *Neuron* 35:989–996.
- Dove A, Pollmann S, Schubert T, Wiggins CJ, von Cramon DY (2000) Prefrontal cortex activation in task switching: an event-related fMRI study. *Cognit Brain Res* 9:103–109.
- Dreher JC, Koechlin E, Ali SO, Grafman G (2002) The roles of timing and task order during task switching. *NeuroImage* 17:95–109.
- Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci* 23:475–483.
- Durston S, Thomas KM, Worden MS, Yang Y, Casey BJ (2002) Effect of preceding context on inhibition: an event-related fMRI study. *NeuroImage* 16:449–453.
- Friston KJ, Jezzard P, Turner R (1994) Analysis of functional MRI time-series. *Hum Brain Mapp* 1:153–171.
- Gold BT, Buckner RL (2002) Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. *Neuron* 35:803–812.
- Goldstein K (1944) The mental changes due to frontal lobe damage. *J Psychol* 17:187–208.
- Grant DA, Berg EA (1948) A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *J Exp Psychol* 38:404–411.
- Halstead WC (1940) Preliminary analysis of grouping behavior in patients with cerebral injury by the method of equivalent and non-equivalent stimuli. *Am J Physiol* 96:1263–1294.
- Houde O, Tzourio-Mazoyer N (2003) Neural foundations of logical and mathematical cognition. *Nat Rev Neurosci* 4:507–514.
- Iversen SD, Mishkin M (1970) Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res* 11:376–386.
- Konishi S, Nakajima K, Uchida I, Kameyama S, Nakahara K, Sekihara K, Miyashita Y (1998) Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nat Neurosci* 1:80–84.
- Konishi S, Nakajima K, Uchida I, Kikyo H, Kameyama S, Miyashita Y (1999a) Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. *Brain* 122:981–991.
- Konishi S, Kawazu M, Uchida I, Kikyo H, Asakura I, Miyashita Y (1999b) Contribution of working memory to transient activation in human inferior prefrontal cortex during performance of the Wisconsin card sorting test. *Cereb Cortex* 9:745–753.
- Konishi S, Donaldson DI, Buckner RL (2001) Transient activation during block transition. *NeuroImage* 13:364–374.
- Konishi S, Hayashi T, Uchida I, Kikyo H, Takahashi E, Miyashita Y (2002a) Hemispheric asymmetry in human lateral prefrontal cortex during cognitive set shifting. *Proc Natl Acad Sci USA* 99:7803–7808.
- Konishi S, Uchida I, Okuaki T, Machida T, Shirouzu I, Miyashita Y (2002b) Neural correlates of recency judgment. *J Neurosci* 22:9549–9555.
- MacDonald III AW, Cohen JD, Stenger VA, Carter CS (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288:1835–1838.
- McCarthy G, Blamire AM, Puce A, Nobre AC, Bloch G, Hyder F, Goldman-Rakic P, Shulman RG (1994) Functional magnetic resonance imaging of the prefrontal cortex activation during a spatial working memory task. *Proc Natl Acad Sci USA* 91:8690–8694.
- Meiran N, Chorev Z, Sapir A (2000) Component processes in task switching. *Cognit Psychol* 41:211–253.
- Milner B (1963) Effects of different brain lesions on card sorting. *Arch Neurol* 9:90–100.
- Milner B (1964) Some effects of frontal lobectomy in man. In: *The frontal granular cortex and behavior* (Warren JM, Akert K, eds), pp 313–334. New York: McGraw-Hill.
- Milner B (1971) Interhemispheric differences in the localization of psychological processes in man. *Br Med Bull* 27:272–277.
- Mishkin M (1964) Perseveration of central sets after frontal lesions in monkeys. In: *The frontal granular cortex and behavior* (Warren JM, Akert K, eds), pp 219–241. New York: McGraw-Hill.
- Monchi O, Petrides M, Petre V, Worsley K, Dagher A (2001) Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J Neurosci* 21:7733–7741.
- Nagahama Y, Fukuyama H, Yamaguchi H, Matuszaki S, Konishi J, Shibasaki H, Kimura J (1996) Cerebral activation during performance of a card sorting test. *Brain* 119:1667–1675.
- Nakahara K, Hayashi T, Konishi S, Miyashita Y (2002) Functional MRI of macaque monkeys performing a cognitive set-shifting task. *Science* 295:1532–1536.
- Owen AM, Roberts AC, Polkey CE, Sahakian BJ, Robbins TW (1991) Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampotomy in man. *Neuropsychologia* 29:993–1006.
- Owen AM, Roberts AC, Hodges JR, Summers BA, Polkey CE, Robbins TW (1993) Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain* 116:1159–1175.
- Passingham RE (1972) Non-reversal shifts after selective prefrontal ablations in monkeys (*Macaca mulatta*). *Neuropsychologia* 10:41–46.
- Petrides M, Alivisatos B, Meyer E, Evans AC (1993) Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc Natl Acad Sci USA* 90:878–882.
- Rogers RD, Sahakian BJ, Hodges JR, Polkey CE, Kennard C, Robbins TW (1998) Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain* 121:815–842.
- Rogers RD, Andrews TC, Grasby PM, Brooks DJ, Robbins TW (2000) Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J Cognit Neurosci* 12:142–162.
- Rowe JB, Toni I, Josephs O, Frackowiak R, Passingham RE (2000) The prefrontal cortex: response selection or maintenance within working memory? *Science* 288:1656–1660.
- Rushworth MFS, Passingham RE, Nobre AC (2002) Components of switching intentional set. *J Cognit Neurosci* 14:1139–1150.
- Rylander G (1939) Personality changes after operations on the frontal lobes. Oxford: Oxford UP.
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme.
- Taylor SF, Kornblum S, Lauber EJ, Minoshima S, Koeppe RA (1997) Isolation of specific interference processing in the Stroop task: PET activation studies. *NeuroImage* 6:81–92.
- 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 USA* 94:14792–14797.
- Worsley KJ, Friston KJ (1995) Analysis of fMRI time-series revisited—again. *NeuroImage* 2:173–181.