

Dorsolateral prefrontal cortex and the implicit association of concepts and attributes

Michael W. L. Chee,^{1,CA} N. Sriram,² Chun Siong Soon^{1,2} and Kok Ming Lee¹

¹Cognitive Neuroscience Laboratory, Singapore General Hospital, 20 College Road, Singapore 169856; ²Department of Social Work and Psychology, National University of Singapore 119260, Singapore

^{CA}Corresponding author

Received 28 September 1999; accepted 3 November 1999

Acknowledgements: This work was supported by the National Medical Research Council of Singapore, Grant NMRC 98/00270. We are grateful to Michael I Posner Ph.D. for helpful critique of the manuscript and to Jia Jia Hoon and Marilyn Chee, who contributed to running the experiments and processing the data.

The Implicit Association Test (IAT) examines the differential association of two object categories (e.g. flower and insect) with attribute categories (e.g. pleasant and unpleasant). When items from congruent categories (e.g. flower + pleasant) share a response key, performance is faster and more accurate than when items from incongruent categories (e.g. insect + pleasant) share a key. Performing incongruent word classification engages

inhibitory processes to overcome the prepotent tendency to map emotionally congruent items to the same response key. Using fMRI on subjects undergoing the IAT, we show that the left dorsolateral prefrontal cortex, and to a lesser extent the anterior cingulate cortex, mediate inhibitory processes where manipulation of word association is required. *NeuroReport* 11:135–140 © 2000 Lippincott Williams & Wilkins.

Key words: Executive function; fMRI; Implicit association; Inhibition; Prefrontal cortex

INTRODUCTION

The need to inhibit habitual or ingrained responses arises in situations where our implicit or habitual tendencies are inappropriate. The capacity to override these habitual or ingrained responses is the function of a supervisory system [1] that includes lateral prefrontal and midline frontal areas. Within the lateral prefrontal cortex, the ventral prefrontal cortex is thought to be critical in mediating inhibitory functions. Imaging studies using the go/no-go task [2,3], an implementation of a Stroop experiment [4] and a verbal, letter-memory inhibition task [5] support this notion.

In the context of the present study, inhibition refers to the process of making a choice contrary to an inherently preferred response. Two well-known tests of frontal lobe function, go/no-go and the Wisconsin Card Sorting Task (WCST) employ newly learned and arbitrary associations between items and responses. In order to generate response selection competition, bias is built up by repeatedly presenting one type of stimulus prior to a shift or change of stimulus type. As such, a novelty or oddball detection effect cannot be excluded as a component process in such experimental designs. In the verbal memory inhibition task used by Jonides *et al.*; [5] a target item was sometimes used later as a source of interference. Stimulus–response map-

ping was not held constant. As a result of incidental memory encoding, interference effects could cumulate in an unpredictable fashion as the task progressed.

In the present study, we sought to characterize the areas activated by inhibitory processing by engaging a task that utilizes items with ingrained (as distinct from arbitrary) associations and one which does not increment working memory load in order to generate interference. We hypothesized that such a task might preferentially recruit dorsolateral prefrontal cortex (DLPFC) on the basis that manipulation of semantic knowledge is required in performing such a task [6].

The IAT [7] examines the differential association of two object categories (e.g. flower and insect) with two attribute categories (e.g. pleasant and unpleasant). In addition to the knowledge of physical and functional characteristics of items contained in the object categories, we also harbor ingrained associations coding emotional valence. These may not be apparent on casual introspection and are thus termed implicit associations [8]. In the IAT, subjects are instructed to make dichotomous choices in two distinct classifications on alternating trials. In the first classification, the subject decides if a word belongs to one or the other category (e.g. a rose is a flower; a beetle is an insect). In the following trial, the subject judges an attribute of a word

(e.g. corpse is unpleasant; heaven is pleasant). The items in the alternating classifications are mutually exclusive. Flower/insect names do not appear in the pleasant/unpleasant task. Previous work has shown that most persons regard flowers as pleasant and insects as unpleasant [7]. When items from congruent categories (e.g. flower + pleasant) share a response key, performance is faster and more accurate than when items from incongruently associated categories (e.g. insect + pleasant) share a key. In the congruent task, categories with similar valences are mapped to the same response key. In the incongruent task, categories with discordant valences are mapped to the same response key. In order to complete the incongruent task successfully, the subject has to remember the counterintuitive assignment of response keys and inhibit the prepotent tendency to assign items with implicitly linked attributes to the same response key. The alternating sequence of the two classifications enhances the inhibitory demands of this task.

In contrast to most other tests exploring inhibition, the change of action schema is predictable in the IAT. The IAT uses the memory of the action taken in the preceding trial to interfere with the current trial and the basis for interference lies in implicit or pre-programmed associations between words and the concepts they allude to. These features enable us to look at the effects of increasing inhibitory processing requirements without changing memory load or incurring novelty effects.

MATERIALS AND METHODS

Implicit associations related to verbal items were evaluated. Eight right-handed undergraduates who gave informed consent participated in the study. In a practice run prior to scanning, subjects began by classifying 7-letter consonant strings (perceptual control) as either upper or lower case using a two-button mouse. Next they classified 20 words as either pleasant or unpleasant. After another block of perceptual controls, subjects classified a further 20 items as flower or insect. This sequence was repeated once. In the congruent task subjects pressed the left mouse in response to 'flower' and 'pleasant' (Fig. 1). Subjects were thus familiarized with the response mappings immediately prior to undergoing fMRI scans. While being scanned, subjects made alternating flower/insect and pleasant/unpleasant decisions every 1.2s. Each block of word classification lasted 30s and was alternated with 20s of upper case/lower case discrimination. Each run comprised

four word classification blocks and four blocks of consonant strings.

After two fMRI runs, another practice was performed for the Incongruent task. This practice was identical in all respects to the congruent task except that the mapping of response keys was switched such that the pleasant words and insect words shared the same response key. Two further fMRI runs were then performed. The order of presentation of the congruent and incongruent tasks was counterbalanced across subjects; four subjects began with the congruent task and the other four subjects with the incongruent task. A total of 200 images per task were obtained per subject.

Imaging and image analysis: Scanning was performed in a 2.0T Bruker Tomikon S200 system (Bruker, Karlsruhe, Germany) fitted with a 30mT/m gradient system. A gradient-echo EPI sequence with the following parameters was used: effective TE 40ms, TR 2000ms, FOV 23 × 23 cm, a 128 × 64 pixel matrix (128 pixels in the A-P direction). Seventeen contiguous oblique axial slices approximately parallel to the AC-PC line 4mm thick (with a 2mm gap) were acquired. High-resolution, T1-weighted anatomical reference images were obtained as a set of 128 contiguous axial slices with a reformatted matrix of 256 × 56 pixels, using a three-dimensional spoiled gradient-echo sequence. Head motion was reduced using a bite bar system (Biomat, Singapore).

Methodology related to image pre-processing has been described in previous work from our laboratory [9]. Functional images were separated into two sets for further analysis: congruent and control; incongruent and control. A direct comparison between incongruent and congruent tasks was not possible because these tasks were not interleaved in our design. Images were shifted 4s to allow for a time lag in the rise of BOLD signal following presentation of word stimuli. Unpaired *t*-tests were then applied to the groups of images. A Z-score threshold of 3.5 was used to generate the activation maps. These were transformed into Talairach space and co-registered with the high resolution anatomical images. Averaging Talairach-transformed statistical and structural images from each subject separately and then re-computing the statistical map resulted in pooled activation maps. Voxels above the specified threshold in the prefrontal, midline frontal and parietal regions were counted. The DLPFC was distinguished from the ventral prefrontal cortex by defining it as

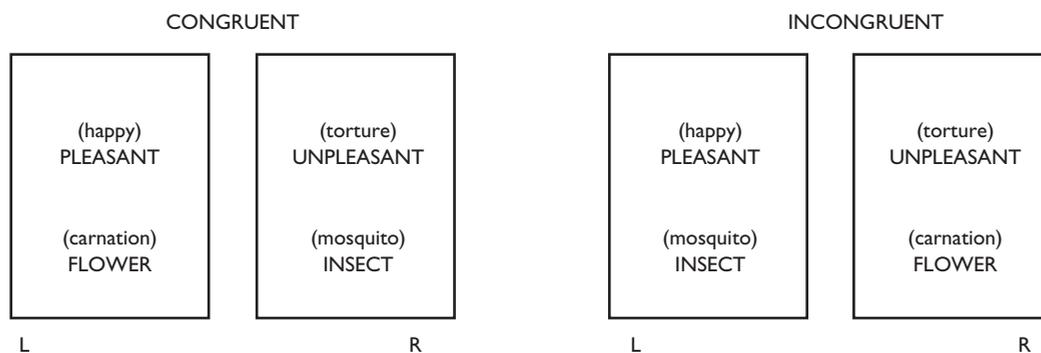


Fig. 1. Congruent and incongruent mapping of object and attribute categories to responses.

the prefrontal cortex above Talairach co-ordinate $z = +28$ mm. This corresponds to the average level of the inferior frontal sulcus. ANOVA was computed on log-transformed values of pixel counts in the prefrontal regions taking into account laterality of activation, location (dorsal *vs* ventral), task (incongruent or congruent) and order of task performance (congruent followed by incongruent or vice versa).

A region of interest in the DLPFC located at $z = +32$ mm was defined in each subject who performed the congruent task first. Pixels surrounding the activation maxima (approximate location $-40, 18, 32$) were included in the region of interest. The region was identical across congruent and incongruent tasks. Signal time courses from individuals and tasks were averaged and the signal intensities normalized to obtain percentage change plots of activations.

RESULTS

Word classification in the incongruent task was associated with increased reaction times $A(F(1,6) = 13.4; p = 0.01)$ as

well as reduced accuracy ($F(1,6) = 10.3; p = 0.021$) compared with the congruent task and control. Repeated presentation of the same stimuli in the incongruent task resulted in some performance improvement, both in terms of reaction time and accuracy (Fig. 2). However, compared to the congruent task, reaction times for the incongruent task were always more delayed.

In the pooled data, word classification of either type compared to case determination resulted in activation of the left prefrontal and superior parietal regions (Fig. 3a,b). Parietal activations were present in an area $(-32, -58, 46)$ corresponding to BA7 occupying the superior end of the inferior parietal lobule and the posterior portion of the superior parietal lobule.

Comparing activations from incongruent and congruent tasks, more extensive activation in the incongruent task was noted in the prefrontal and midline frontal areas as well as the superior parietal areas bilaterally (Fig. 3a, Fig. 4a). Activation maxima in the frontal regions were

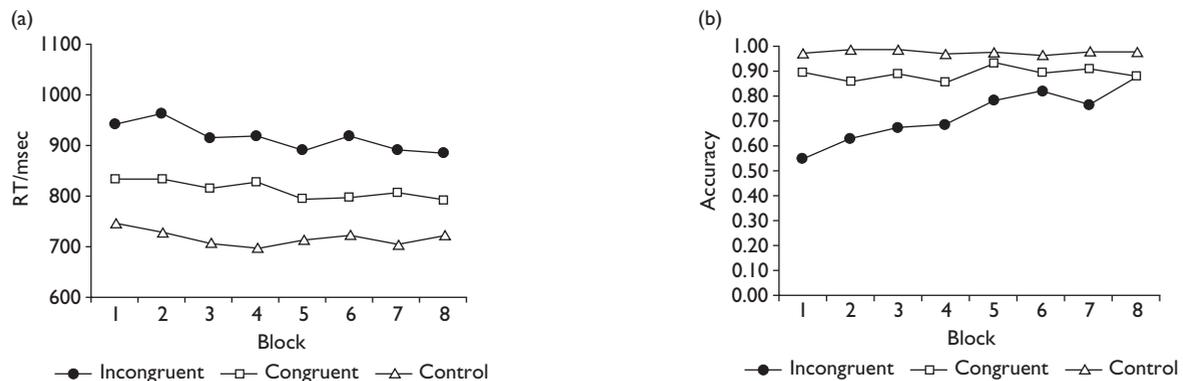


Fig. 2. (a) Mean reaction time data from eight subjects plotted as a function of time. The numbers on the x-axis denote the block number. (b) Accuracy data from the same eight subjects segregated by block.

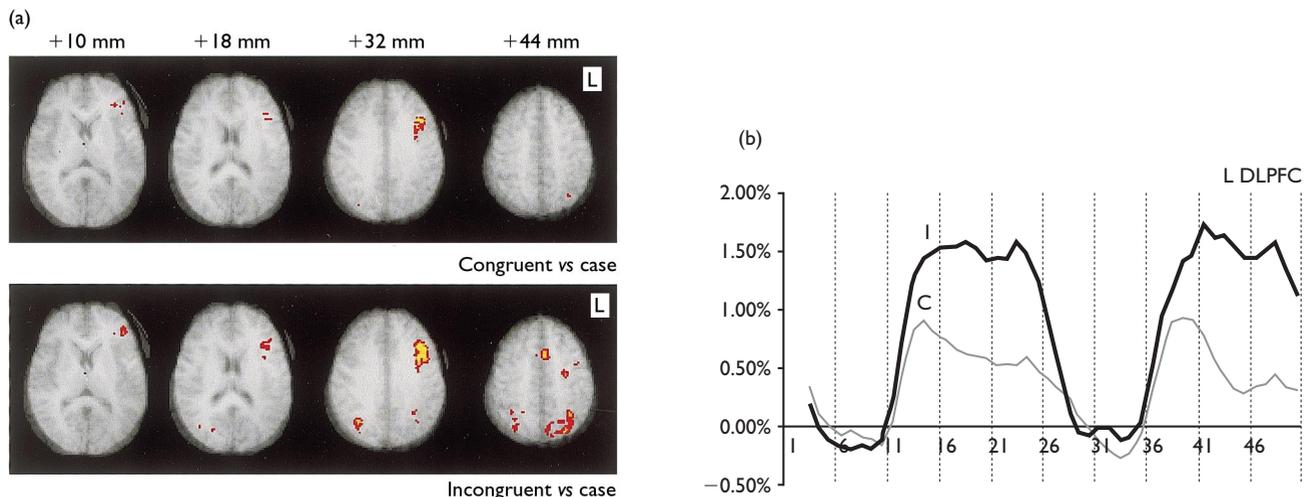


Fig. 3. (a) Axial slices from Talairach transformed, averaged images derived from eight subjects' data showing prefrontal, anterior cingulate and parietal activations. The numbers above the images indicate the distance above the intercommisural plane in millimeters. (b) Averaged time course across task blocks and subjects collapsed into two task-control cycles. Voxel intensities during the incongruent (I; solid line) and congruent (C; shaded line) tasks from a region of interest in the left prefrontal cortex corresponding to $z = +32$ mm are shown. The first 10 images relate to the control task and the next 15 images to the task.

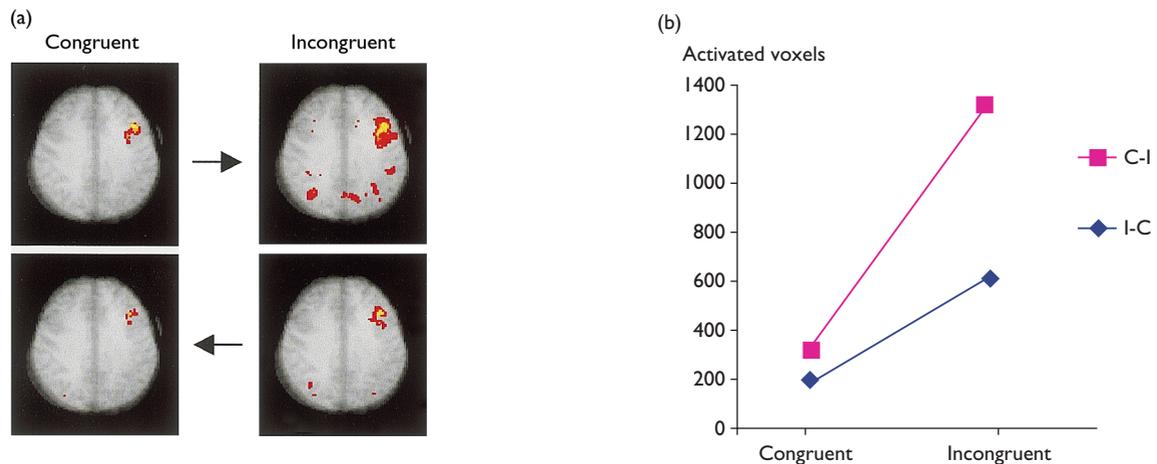


Fig. 4. (a) Axial slices from Talairach transformed, averaged images derived from subjects performing the incongruent task first (I-C), upper panel and the congruent task first (C-I) lower panel. The arrows indicate the order of task performance. (b) Plot showing mean voxel counts taken from a region of interest in the left prefrontal region segregated by order of task performance and task.

Table 1. Location of activation maxima in left prefrontal region (Talairach coordinates)

Subject	t-statistic	Incongruent			t-statistic	Congruent		
		x	y	z		x	y	z
1	7.69	-40	16	34	4.57	-42	16	34
2	11.0	-42	12	28	3.70	-42	8	32
3	8.47	-40	26	32				
	7.87	-40	18	30				
4	10.2	-38	10	36	6.26	-34	8	36
	8.25	-44	18	30	5.04	-36	20	32
5	6.08	-34	28	32	3.53	-30	28	32
6	5.24	-34	20	30	4.94	-34	20	42
	4.09	-34	20	40	4.61	-34	20	30
7	7.44	-36	18	28	5.17	-36	20	28
	6.29	-44	14	40				
8	8.82	-30	24	40	4.53	-30	24	40
	6.90	-28	14	32	3.87	-28	14	32

located in the left ventral prefrontal cortex corresponding to BA 47 (Talairach co-ordinates -40,34,8), the DLPFC (-40,16,32; Table 1) in the middle frontal gyrus corresponding to BA9 and the pre-SMA/anterior cingulate region (-2,10,48). There was an increase in the spatial extent as well as magnitude of activations in the incongruent task *vs* control (Fig. 3a,b). The increase in prefrontal activations was significantly greater in left prefrontal region than the corresponding right prefrontal region (Table 2; $F(1,6)=157$; $p=0.00001$). Modulation of activation across tasks was greater in the DLPFC compared to the ventral prefrontal cortex ($F(1,6)=7.5$; $p=0.03$). This effect involved the left but not right prefrontal activations ($F(1,6)=8.8$; $p=0.021$). It is important to point out that the DLPFC is in fact contiguous with parts of the ventral prefrontal cortex that lie below it. Our intention in making the distinction between ventral and dorsal areas is to merely to highlight a functional difference in the more dorsal parts of the prefrontal cortex.

Across individuals, there was consistent activation in the dorsal prefrontal, posterior ventral prefrontal and parietal regions in each subject. However, the spatial

extent of activation was very variable (Table 2), necessitating a log transform of pixel counts (all counts were incremented by 1 to set zero as the minimum value in the log transformed data) in the prefrontal region to facilitate analysis by ANOVA. Activations in BA 47 and the anterior cingulate were less robust and tended to appear in subjects who underwent the congruent-incongruent version of the test.

When time courses from individual subjects were pooled, we noted that there was no decrement in the magnitude of activations from block to block in the incongruent task. This was in contrast to the improvement in reaction time and accuracy shown in the behavioral data (Fig. 2) and serves to illustrate that imaging and behavioral data can provide independent information about cognitive processes.

The modulation of prefrontal and anterior cingulate activations across congruent and incongruent tasks was greater when the congruent task was performed before the incongruent task ($F(1,6)=20.7$; $p=0.004$; Fig. 4a,b). This may stem from reinforcement of implicit associations when subjects performed the congruent task first.

Table 2. Activated voxels tabulated by region in congruent (C) and incongruent (I) tasks

Subject	L ventral frontal (BA 47,44; -40,34,8; -38,12,20)		R ventral frontal (BA 47,44; -40,34,8; -38,12,20)		L DLPFC (BA 9,44; -40,16,32)		R DLPFC (BA 9,44; -40,16,32)		Anterior cingulate/pre- SMA (BA 32; -2,10,48)	
	C	I	C	I	C	I	C	I	C	I
C-I										
1	0	91	0	0	19	217	0	41	2	8
2	122	159	1	301	33	984	0	350	1	283
3	61	36	1	30	0	344	0	25	0	107
4	45	198	0	41	52	353	0	122	0	124
I-C										
5	23	130	1	29	32	210	0	8	7	17
6	2	13	0	0	32	61	0	0	6	9
7	12	36	0	0	74	168	0	10	0	4
8	2	4	0	0	0	102	0	0	0	0

Data from eight individual subjects segregated according to order of task performance: C-I indicates that the congruent task was performed first.

DISCUSSION

Our results suggest that within the frontal regions, inhibition of implicit word associations involve prefrontal and the anterior cingulate cortex. The prefrontal activations observed were strongly lateralized to the left hemisphere, dorsal to where activations related to inhibitory processing have previously been demonstrated [3-5,10], and were robustly reproduced across subjects.

The organizing principles underlying memory and executive processes in the prefrontal lobe are subjects of considerable interest (see [11] for a review). While much work has been channelled to the study of working memory, relatively less attention has been paid to inhibitory processes despite their involvement in operations considered executive or supervisory.

On the basis of primate studies, it was long held that the ventral prefrontal cortex had an exclusive role in mediating inhibition [12,13]. This view has prevailed despite evidence from patient studies showing impaired performance in the WCST following damage to the DLPFC [14]. Damage to this region as a result of phenylketonuria results in a loss of ability to act counter to a pre-programmed initial tendency [15].

New lesion [16] and computational studies have strengthened the notion that the suppression of inappropriate prepotent responses may be a property of the entire prefrontal cortex [17,18]. Despite these points, functional imaging studies to date have imputed inhibitory function only to the ventral prefrontal cortex [2,3,5,10].

What underlies the dorsal extension of prefrontal activations in the present study? Previous experiments involving mnemonic processing where the DLPFC was activated, relate to situations where there is increasing item load [19,20], where manipulation of the contents of working memory is required [21] and where task monitoring is required [22]. Activation of the left middle and inferior frontal gyrus in the DLPFC occurs where the need to select a semantically appropriate response from competing responses increases [6]. Given that short-term retention of semantic contents is an essential part of context-based control processes mediated by the frontal lobes [18], one

interpretation of the present data is to posit that the DLPFC activation is driven by the need to access semantics and then manipulate that information. In the present experiment, a subject seeing the word 'termite' in the incongruent task realizes that this is an insect. However, pressing the left key is interfered with because the same key maps to pleasant in the previous classification schema.

Suppressing the implicit tendency to map 'termite' to the unpleasant response key constitutes a type of inhibition not studied in prior investigations that uncovered ventral prefrontal activation [2-5].

N-back tasks activate the DLPFC and the prevailing notion is that this is because manipulation of working memory is required [11,21]. As N increases, the task becomes more difficult, arguably because of an increasing load on working memory as well as the increase in competing, task-irrelevant responses. Inhibitory as well as mnemonic processes are involved in N-back tasks; the latter providing the context for the former to act on. Since increasing memory load itself can result in DLPFC activation [20], imputing inhibitory function on the DLPFC requires a task that dissociates memory load from inhibitory requirements. In the IAT, the constancy of stimulus-response mappings and the use of identical items in the congruent and incongruent tasks ensure a constant working memory load and allows us to relate DLPFC activations to inhibitory demands.

The lateralization of prefrontal activation in the present experiment was striking and is opposite to the pattern seen with the go/no-go task. The present findings are inconsistent with the notion that the right inferior frontal area has universal utility in performing inhibitory operations [3,23]. Compared with the dorsoventral organization of working memory, there appear to be few or no organizing principles for lateralization of inhibitory processes. Human case studies suggest that left frontal damage is associated with memory deficits involving verbal items and right frontal damage with non-verbal items [13]. The go/no-go task, which involves attention to action using non-verbal items activates the right prefrontal region when one only considers responses related to the inhibition of the 'go'

response [3]. The suppression of prepotent responses to recently shown letters [5,10] and the suppression of word reading in a color naming Stroop task [4] activate the left ventral prefrontal region. Before considering a media-specific effect (words *vs* non-words), we note that tasks that utilize shapes or objects and ostensibly recruit inhibitory processing [23,24] activate both prefrontal cortices, not preferentially the right. This could reflect a tendency to use verbally based strategies whenever a task lends itself to such a possibility [25]. However, another consideration is that inhibitory processes are different in all of these experiments. For example, the go/no-go and WCST involve detection of a change in the experimental milieu, inhibiting the tendency to use the current action schema and replacing it with a new one. The color word naming Stroop test involves selective attention to word color while suppressing an innate tendency to read words. While a direct comparison of different inhibitory tasks has not been performed, a double dissociation between inhibitory processes involving affect and higher order attentional factors has been demonstrated in a lesion study involving marmosets [16]. This supports the notion that spatially distinct areas within the prefrontal region can mediate different inhibitory processes. As such our findings complement rather than contradict existing findings.

The appearance of anterior cingulate activations in the Incongruent task is consonant with the view that this area mediates conflict resolution and is activated when there is an increased likelihood of making an erroneous response [26]. Compared with the increased prefrontal activations seen in all subjects when the incongruent and congruent tasks were compared, an increase in cingulate activation was less robust. Of interest is the observation that cingulate activation was more prominent when the congruent task was preformed first. A plausible explanation for this is that the reinforcement of an inherent bias enhances expectations of congruent responses and leads to greater cingulate activation [11].

CONCLUSION

The left dorsolateral prefrontal cortex and, to a lesser extent, the anterior cingulate cortex are activated during the inhibition of prepotent responses to implicit word associations.

REFERENCES

1. Norman D and Shallice T. Attention to action: willed and automatic control of behavior. In: Davidson RJ, Schwartz GE and Shapiro D, eds. *Consciousness and Self Regulation*. New York: Plenum Press, 1986: 1–18.
2. Kawashima R, Satoh K, Itoh H *et al*. *Brain Res* **728**, 79–89 (1996).
3. Konishi S, Nakajima K, Uchida I *et al*. *Brain* **122**, 981–991 (1999).
4. Taylor SF, Komblum S, Lauber EJ *et al*. *Neuroimage* **6**, 81–92 (1997).
5. Jonides J, Smith EE, Marshuetz C *et al*. *Proc Natl Acad Sci USA* **95**, 8410–8413 (1998).
6. Thompson-Schill SL, D'Esposito M, Aguiuffe GK and Farah MJ. *Proc Natl Acad Sci USA* **94**, 14792–14797 (1997).
7. Greenwald AG, McGhee DE and Schwartz JL. *J Pers Soc Psychol* **74**, 1464–1480 (1998).
8. Greenwald AG and Banaji MR. *Psych Rev* **102**, 4–27 (1995).
9. Chee MW, Caplan D, Soon CS *et al*. *Neuron* **23**, 127–137 (1999).
10. D'Esposito M, Postle BR, Jonides J and Smith EE. *Proc Natl Acad Sci USA* **96**, 7514–7519 (1999).
11. Smith EE and Jonides J. *Science* **283**, 1657–1661 (1999).
12. Iversen SD and Mishkin M. *Exp Brain Res* **11**, 376–386 (1970).
13. Fuster JM. Human neuropsychology. In: Fuster JM, ed. *The Prefrontal Cortex: Anatomy, Physiology and Neuropsychology of the Frontal Lobe*. Philadelphia: Lippincott-Raven, 1997: 150–184.
14. Milner B. *Arch Neurol* **9**, 90–100 (1963).
15. Diamond A. Evidence for the role of dopamine for prefrontal cortex functions in early life. In: Roberts AC, Robbins TW and Weiskrantz L, eds. *The Prefrontal Cortex: Executive and Cognitive Functions*. Oxford: Oxford University Press, 1998: 144–164.
16. Dias R, Robbins TW and Roberts AC. *J Neurosci* **17**, 9285–9297 (1997).
17. Roberts AC, Robbins TW and Weiskrantz L. Discussions and conclusions. In: Roberts AC, Robbins TW and Weiskrantz L, eds. *The Prefrontal Cortex: Executive and Cognitive Functions*. Oxford: Oxford University Press, 1998: 221–242.
18. Cohen JD, Braver TS and O'Reilly RC. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. In: Roberts AC, Robbins TW and Weiskrantz L, eds. *The Prefrontal Cortex: Executive and Cognitive Functions*. Oxford: Oxford University Press, 1998: 195–220.
19. Rypma B, Prabhakaran V, Desmond JE *et al*. *Neuroimage* **9**, 216–26 (1999).
20. Rypma B and D'Esposito M. *Proc Natl Acad Sci USA* **96**, 6558–6563 (1999).
21. Smith EE, Jonides J, Marshuetz C and Koeppel RA. *Proc Natl Acad Sci USA* **95**, 876–882 (1998).
22. Petrides M, Alivisatos B, Meyer E and Evans AC. *Proc Natl Acad Sci USA* **90**, 878–882 (1993).
23. Nagahama Y, Fukuyama H, Yamanauchi H *et al*. *Brain* **119**, 1667–1675 (1996).
24. Konishi S, Nakajima K, Uchida I *et al*. *Nature Neuroscience* **1**, 80–84 (1998).
25. Kelley W, Buckner RL, Miezin F *et al*. *Neuroimage* **9**, S937 (1999).
26. Carter CS, Braver TS, Barch DM *et al*. *Science* **280**, 747–749 (1998).