

# Overcoming Confounds of Stimulus Blocking: An Event-Related fMRI Design of Semantic Processing

L. K. Pilgrim,\* J. Fadili,\* P. Fletcher,† and L. K. Tyler\*‡

\*Department of Experimental Psychology, †Department of Psychiatry, and ‡Wolfson Brain Imaging Center, University of Cambridge, England

Received September 26, 2001

**The way in which meaning is represented and processed in the brain is a key issue in cognitive neuroscience, which can be usefully addressed by functional imaging techniques. In contrast to previous imaging studies of semantic knowledge, which have primarily used blocked designs, in this study we use an event-related fMRI (erfMRI) design, which has the advantage of enabling events to be presented pseudorandomly, thus reducing strategic processes and enabling more direct comparison with psychological behavioral studies. We used a semantic categorization task in which events were words representing either artifact or natural kinds concepts. Significant areas of activation for semantic processing included inferior frontal lobe bilaterally (BA 47) and left temporal regions, both inferior (BA 36 and 20) and middle (BA 21). These are areas that have been identified in previous neuroimaging studies of semantic knowledge. However, there were no significant differences between artifact and natural kinds concepts. These results are consistent with our previous imaging studies using blocked designs and suggest that conceptual knowledge is represented in a unitary, distributed neural system undifferentiated by domain of knowledge. These findings demonstrate that event-related designs can generate activations that are similar to those seen in blocked designs investigating semantics and, moreover, offer a greater capacity for interpretation free from the confounds of block effects.** © 2002 Elsevier

Science (USA)

## INTRODUCTION

Recently, neuroimaging techniques have been used to investigate how semantic knowledge is represented and processed (e.g., Perani *et al.*, 1995; Martin *et al.*, 1996; Cappa *et al.*, 1998; Tyler *et al.*, 2001; Devlin *et al.*, 2001). Currently there are three main types of model concerned with the organisation of semantic knowledge: Feature-based models which argue that the main organizing principle of semantic memory is type of

semantic property, such as perceptual & functional features (e.g., Warrington and McCarthy, 1983, 1987); domain-specific models, which suggest that there are dedicated knowledge systems sub-served by distinct neural mechanisms (e.g., Caramazza and Shelton, 1998); and more recently, distributed models, which argue that conceptual knowledge is represented in a unified neural system, which is not differentiated by category or domain and where category structure emerges out of the structure and content of individual concepts (e.g., Tyler *et al.*, 2000; Tyler and Moss, 2001).

Although neuroimaging studies have reliably identified brain regions involved in semantic processing, there has been little consistent evidence of domain-specific or category-specific semantic activation at a neural level (see Price and Friston, 2001; Martin and Chao, 2001, for reviews). While a number of studies have shown increased activation for tools in the left posterior middle temporal lobe (e.g., Chao *et al.*, 1999; Martin *et al.*, 1996; Perani *et al.*, 1999; Moore and Price, 1999), many of these effects were only significant at a very low threshold or when uncorrected for multiple comparisons. In addition, not all studies report differential activation in this area (Devlin *et al.*, 2001; Perani *et al.*, 1995). In a previous series of three studies using fMRI and PET, we used lexical decision and semantic categorization tasks to investigate the neural correlates of semantic knowledge (Devlin *et al.*, 2001). Our items were carefully matched in terms of various nuisance variables, and activations were corrected for multiple comparisons. Although we found a robust semantic network in the left hemisphere consistent with previous studies, there were no differences as a function of domain or category. The only detectable differences were at an uncorrected level and these tended to be inconsistent across the three studies. It is plausible that lack of consistency across neuroimaging studies of category-specificity may be attributable to the different, often lenient, thresholds chosen for significance level, as well as inadequate matching of stimuli across conditions.

Moreover, PET and fMRI studies investigating category-specificity have typically used blocked designs. A number of items of a single type (e.g., animals) are presented in a block, which is followed by another block of a different stimulus type (e.g., tools), and so on. These blocks may last from 20 seconds to several minutes. Blocked designs are powerful in their ability to detect activation: "responses to events of the same type are summed, concentrating power at the block alternation frequency" (Josephs and Henson, 1999). However, they have important consequences in both neural and cognitive terms.

At a cognitive level, blocked presentation of items may induce strategies (see Strayer and Kramer, 1994 for comparison of blocked and randomized trials on performance of a memory search task) and therefore may not reflect normal automatic semantic processing (den Heyer *et al.*, 1983; Keefe and Neely, 1990). Blocking items may also affect the way people evaluate stimuli, for instance, whether they use semantic or perceptual/contextual judgements (Johnson *et al.*, 1997). In addition, the psychological literature relating to semantic satiation suggests that extended repetition of a category word can result in loss or decrement of meaning of items within that category (see for example, Smith and Klein, 1990; Balota and Black, 1997; Kounios *et al.*, 2000).

At a neural level we also know that repetition can lead to decreased activation in brain regions that were activated during the initial processing of a stimulus (Schacter and Buckner, 1998; Wiggs and Martin, 1998; Wagner *et al.*, 2000). This may be a particular problem in blocked designs where items from one category are repeated within a single block. This habituation effect may lead to insufficient activation to allow detection of differences between semantic conditions, even if they do exist. In addition, habituation effects may not be consistent over different types of stimuli or different brain regions. Henson, Shallice and Dolan (2000) found an interaction between stimulus familiarity and repetition effects in the right fusiform gyrus. For repeated meaningful stimuli there was reduced activation (i.e., habituation), whereas for repeated meaningless stimuli there was increased activation. In addition, Henson (2001) reported habituation effects for repeated words in contrast to enhanced activation for repeated non-words in the left fusiform gyrus.

Given these problems we cannot be confident that blocked imaging designs provide reliable data on the neural basis of category specificity. However, if we acquire the same results without blocking items, we can be more confident that our previous findings were not confounded by factors which are not germane to semantic processing.

In event-related designs items are presented in a pseudo-randomised order rather than in blocks. These types of designs have several advantages over blocked

designs and have consequences both in imaging and cognitive terms. Although they have lower detection power than blocked designs, event-related designs are better able to estimate the shape of the hemodynamic response (Dale, 1999) allowing timing information to be included (which may be particularly important in language studies where we know that various brain regions involved in language processing show transient or phasic responses to stimuli, see Price and Friston, 1997). Pseudorandom presentation also decreases context or strategic effects (i.e. predictability of the next item decreases). In addition, if we can obtain the same results with an event-related design as with a blocked design we can rule-out habituation as a reason for obtaining no differences between two semantic conditions. Furthermore, pseudorandom presentation of stimuli is comparable to existing behavioural and ERP designs. Thus, using an event-related, pseudo-randomised design can better model both the hemodynamic response and the cognitive processes under investigation. Given the advantages of pseudo-randomised, event-related fMRI designs, particularly that they rule out the potential confounds produced by blocking stimuli, we used such a design to investigate semantic processing. Although previous studies (e.g., Wagner *et al.*, 1998) have used event-related designs to investigate such processes as memory encoding for example, these designs have not yet been used specifically to examine semantic processing of different categories of knowledge, which was the aim of the present study.

Our main goal was to determine whether an event-related design is sensitive enough to detect areas commonly associated with semantic processing, and our secondary aim was to ask whether there are cortical areas selectively responsive to items from either living or nonliving domains. In the discussion section we also compare the results from the present study using an *erfMRI* design with our previous imaging studies which all used blocked designs. A direct statistical comparison was not undertaken as any differences between the two designs could be due to either the psychological effects of stimulus blocking or to the differences in power between the two designs and would therefore be uninterpretable. Thus, our focus was on the convergence of patterns of activation obtained in blocked and event-related designs.

## METHODS

### *Participants*

Fifteen right-handed (assessed using the Edinburgh Handedness Inventory), native British English speakers were tested. Each gave informed consent after the experimental methodology was explained. One participant was removed from the analysis because of exces-

**TABLE 1**  
Example Stimuli

Task	Cue 1	Cue 2	Cue 3	Target	Response
Semantic	string	rope	cord	THREAD	Same
	pig	goat	cow	SHEEP	Same
	mug	saucer	cup	SANDALS	Different
	apricot	melon	lemon	RABBIT	Different
Baseline	ddd	dddd	dddd	DDDD	Same
	gggggggg	gggggg	gggggg	LLLLLL	Different

sive movement (exceeding 5 mm and 4° in terms of translation and rotation corrections). Data from 14 volunteers aged 18–29 years (mean 23, 8 female) was analyzed.

### *Experimental Materials and Design*

Before scanning began participants were given written instructions on the nature of the categorisation task. Once inside the scanner they were given 36 practice items (none of which were in the main testing sessions) to judge whether they were able to perform the task adequately. Responses were made via an in-house designed, two-button response box.

Stimuli were presented visually via computer and projected to a mirror directly above the participant's head, at eye-level. Effective viewing distance was approximately 18 inches. Items were presented in Arial 22 font, in black on a white background. Items in the two test conditions consisted of either natural kinds (e.g., animals, birds, fruit/vegetables) or artifacts (e.g., tools, vehicles, clothing). The baseline consisted of single letter strings (e.g., kkkk, ddddddd, rrr). There were 96 items in each of the three conditions. Artifacts and natural kinds were matched on word familiarity and frequency. Artifacts had a mean familiarity of 503 (standard deviation (SD) 47.1) and a mean frequency of 16 (SD 10.2). Natural kind trials had a mean familiarity of 496 (SD 32.9) and a mean frequency of 15 (SD 14.2). The items in the letter task were matched on letter length with those in the semantic task. Stimuli ranged from three to nine letters in length with a mean of 5.4 (SD 0.9) letters.

For the semantic conditions, an event consisted of three cue words followed by a fourth target word. Participants were required to press a response button to indicate whether the target word was in the same or different category as the three cue words. Different responses were taken from within the same domain (e.g., artifacts) but from different categories within that domain (e.g., tools vs vehicles). Cue words were presented in lower case and target words were presented in upper case. In the Baseline task cues and targets were letter strings (cues in lower case and targets in upper case). Participants indicated whether

the target letter string was composed of the same or different letters as the cues. Thus the baseline task shared the same stimulus and response characteristics as the semantic categorisation task but had no lexical or semantic component. See Table 1 for examples.

Each of the four words or letter strings was presented for 200 ms with a between-item delay of 400 ms. The target item was followed by a delay of 3 s. Each event (quadruple of words/letter strings) lasted for 5 s. The same timing parameters were used for both the semantic and baseline tasks. Items were presented in four sessions each lasting approximately 6 min, with 1–2 min rest between each session. There was some repetition of individual words, although no event (i.e., quadruple of words) contained the same sequence of four words. Independent *t* tests comparing number of word repetitions within each session showed no significant differences between natural kinds and artifacts. In session 1 there were an average of 2.81 (SD 1.47) repetitions of artifact words and 2.39 (SD 1.17) repetitions of natural kinds words ( $t_{72} = 1.331$ ,  $P < 0.187$ ). In session 2 there were 3.11 (SD 1.20) repetitions of artifact words and 2.65 (SD 0.9) repetitions of natural kinds words ( $t_{43} = 1.392$ ,  $P < 0.171$ ). In session 3 there were 2.75 (SD 1.07) repetitions for artifact words and 2.33 (SD 0.8) repetitions of natural kinds words ( $t_{36} = 1.325$ ,  $P < 0.194$ ). In session 4 there were 2.76 (SD 1.30) repetitions of artifact words and 2.23 (SD 1.28) repetitions of natural kinds words ( $t_{54} = 1.540$ ,  $P < 0.130$ ).

### *fMRI Procedure and Analysis*

Scans were carried out at the Wolfson Brain Imaging Center, Cambridge, England, on a 3 Tesla Bruker Medspec Avance S300 system (Bruker Medical, Ettlingen, Germany). A gradient-echo EPI sequence was used for image collection (TR = 3000 ms, TE = 30 ms, flip angle 90°, FOV 25 × 25 cm, 21 oblique slices, 4mm thick (1-mm gap between slices, 128 × 128 in-plane resolution, 126 repetitions) with head coils, 200 kHz bandwidth and spin echo guided reconstruction. The event duration (5 seconds) was not an integer multiple of the TR (3 s) and therefore data was sampled at various points along the HRF ensuring an effective sampling

rate (Josephs *et al.*, 1997). All participants wore protective earplugs and ear-defenders.

Data was analyzed using Statistical Parametric Mapping software (SPM99, Wellcome Institute of Cognitive Neurology, [www.fil.ion.ucl.ac.uk](http://www.fil.ion.ucl.ac.uk)), implemented in Matlab (Mathworks Inc. Sherborn, MA). The first six scans of each time series were discarded to allow for T1 equilibrium before the test trials started.

### *Spatial Preprocessing*

To correct for different acquisition times (relative to the onset of each event), signal in all slices was shifted relative to the acquisition of the first slice using a sinc interpolation in time. Our previous research showed that the loss of BOLD signal near air-tissue interfaces at high magnetic field strengths (Devlin *et al.*, 2000) may lead to errors in normalization (i.e., signal may be dragged down from other areas during normalization), therefore we masked areas of susceptibility prior to normalisation. Masking was done by hand using MRIcro (MRIcro, Chris Rorden, [chris.rorden@nottingham.ac.uk](mailto:chris.rorden@nottingham.ac.uk)), and areas affected by susceptibility in each affected slice were filled. This mask was saved as a region of interest (ROI) and then used during normalization (the masked areas were then not taken into account during normalization). After normalizing to the EPI template in SPM99 (i.e., transforming the images onto the MNI [Montreal Neurological Institute] mean brain) the data was spatially smoothed using an 8-mm smoothing kernel.

### *Model Specification and Parameter Estimation*

The data was modelled using the general linear model (GLM) (Friston *et al.*, 1995). Four sessions and three variables were entered into the model (artifacts, natural kinds, and baseline). In order to partial-out any effect of reaction time, participant's reaction times for each event (with errors replaced by the mean for that item type) were entered as a parametric modulator (with linear expansion). Trials were entered as events and modelled using a canonical hemodynamic response function (HRF) with temporal derivatives. The model was then estimated in the following way. In order to be comparable with our previous blocked design studies global scaling was applied. Data was high-pass filtered (minimum cut-off point 62 s) to remove low frequency drifts in signal. The data was temporally smoothed with an HRF-shaped filter to attenuate high frequency components of the data.

A first level, within-subjects analysis (i.e., a fixed-effects model) using the GLM was performed on the functional data from each participant individually. Five contrasts per participant were computed:

1. Artifacts + natural kinds (semantics) – letter detection (baseline)

2. Artifacts – natural kinds
3. Natural kinds – artifacts
4. Artifacts – baseline
5. Natural kinds – baseline

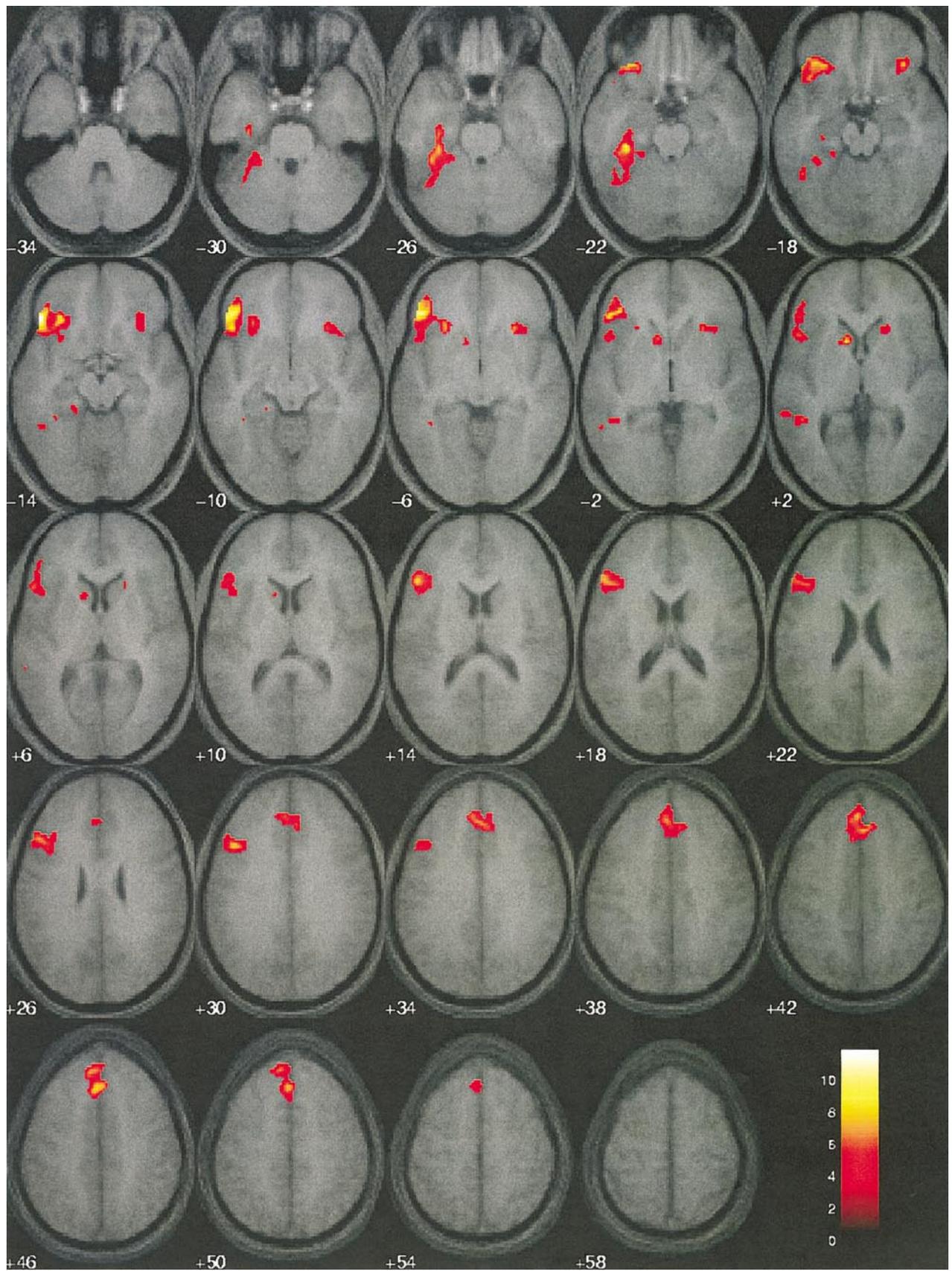
Each of the resulting contrast images was taken through to a second-level, between-subjects group analysis (i.e., a random-effects model). The results from this random-effects group analysis are presented here. In order to address the two main questions outlined in the introduction, three one-sample *t* tests were performed on the data. To specify areas which were activated by both semantic categories compared to the baseline task, the contrast artifacts + natural kinds – baseline was inclusively masked with artifacts – baseline and natural kinds – baseline. The inclusive masking ensured that areas of activation were greater than baseline for both the semantic categories, rather than by only one of the categories.

To address whether there were any areas specific to artifacts, the contrast artifacts – natural kinds was inclusively masked with artifacts – baseline. The masking ensured that areas were activated significantly more for artifacts than for either natural kinds or baseline conditions and also ensured that any differences were not merely the result of deactivation of natural kinds relative to baseline. To assess whether there were any areas modulated by natural kinds the contrast, natural kinds – artifacts was inclusively masked with natural kinds – baseline. Again the inclusive masking ensured that activations in natural kinds activated significantly more than both artifacts and baseline conditions and were also not the result of deactivation of artifacts relative to baseline.

## RESULTS

### *Behavioural Results*

In the semantic categorisation task participants' mean reaction time for Artifacts was 716 ms (SD 107.3), which was not significantly different from the mean RT of 699 ms (SD 96.2) for Natural Kinds. Error rates for Artifacts were 6.1% (SD 11.5) and 9.2% (SD 14.7) for Natural Kinds. Collapsed over the two semantic domains the mean reaction time was 708 ms (SD 102) with a mean error rate of 7.6% (SD 13.3). Mean reaction time for the letter categorisation task was 572 ms (SD 50.9) with a mean error rate of 2.4% (SD 4.7). A univariate analysis of variance comparing the two semantic domains (artifacts and natural kinds) showed no main effect of domain ( $F(2,176) = 1.378$ ,  $P < 0.242$ ), i.e., there was no significant difference in reaction times between artifacts and natural kinds. Similarly, error rates did not differ as a function of domain ( $F(2,188) = 2.665$ ,  $P < 0.104$ ). As expected reaction times to baseline items were significantly faster than those to semantic items. A univariate analysis of vari-



**FIG. 1.** The areas activated in the semantic-baseline (masked) contrast. All activations reached corrected significance  $P < 0.05$  at a cluster level

**TABLE 2**  
Semantic Activations

Description	Extent	SPM $\{t\}$	x	y	z
Activations reaching corrected significance $P < 0.05$ at a cluster level					
Frontal lobes					
L. Inferior Frontal (BA 47)	2085	11.86	-50	30	-14
		10.40	-50	34	-6
		9.64	-44	36	-10
R. Inferior Frontal (BA 47)	297	7.35	26	22	-4
		7.26	36	30	-18
		5.22	22	22	2
L/R. Medial Frontal (BA 8)	994	7.91	-2	24	48
		7.41	4	28	44
		7.02	-2	38	42
Temporal Lobe					
L. Inferior Temporal (BA 36, 20)	723	8.88	-36	-30	-22
		6.72	-40	-40	-26
		6.67	-48	-50	-18
L. Middle Temporal (BA 21)	93	6.65	-48	-44	0
		5.42	-62	-40	4
		5.08	-40	-46	-4
Caudate	89	8.98	-12	14	2

Note. The table shows significant clusters, after correcting for multiple comparisons, and the peak voxels within each cluster.

ance collapsing over domain showed a significant main effect of the semantic compared to baseline condition ( $F(2,282) = 81.311$ ,  $P < 0.001$ ). Error rates were also significantly higher for the semantic compared to baseline items ( $F(2,288) = 8.940$ ,  $P < 0.001$ ).

Reaction times were entered into the model as a parametric regressor and therefore we performed an explicit analysis of this regressor using the same contrasts as we used for the test items (i.e., artifacts-natural kinds and natural kinds-baseline). In line with the behavioral data, there were no significant domain differences and therefore this analysis will not be discussed further.

### Imaging Results

#### Common Semantic Activations

As noted above, activations common to both artifacts and natural kinds were computed by contrasting artifacts and natural kinds with the letter detection task (i.e., artifacts + natural kinds - baseline). This contrast was then inclusively masked<sup>1</sup> by each of the semantic domains minus letter detection (i.e., artifacts - baseline + natural kinds - baseline). Activations were thresholded at SPM  $\{t\} = 3.85$   $P < 0.001$  uncorrected and clusters are reported as significant if they reach  $P < 0.05$  corrected. This masked contrast produced several large clusters of activation in the left hemisphere (LH) which extended across the frontal and

temporal gyri. In addition there was a single smaller cluster in the right hemisphere (RH) in the inferior frontal gyrus and small cluster in the left caudate (see Fig. 1 and Table 2).

The largest cluster, of over 2000 voxels, was in the left frontal lobe with peaks in inferior frontal gyrus (GFi) BA 47, extending across Broca's area (BA 45) the pars triangularis and into the adjacent pars opercularis (BA 44), then anteriorly (BA 11) and dorsally into the middle frontal gyrus (GFm) (BA 46). There was also activation in a similar, but smaller region to the left GFi activation in RH extending across BA 47.

There was also a large cluster of nearly 1000 voxels in the medial aspect of the superior frontal lobe (GFd), extending across both LH and RH, with peaks in BA 8 and BA 9 bilaterally. In addition there was a smaller cluster (89 voxels) in the left caudate.

In the LH temporal lobe there were two significant clusters of activation. The first was along the ventral aspect of the inferior temporal lobe, where there was a large cluster of activation (over 700 voxels) with peaks in the parahippocampal gyrus (GH) and fusiform gyrus (GF) (BA 36). This cluster extended anteriorly into the uncus (U) (BA 20) and posteriorly into the cerebellum. In addition there was also a smaller (93 voxels) cluster in the middle temporal gyrus (GTm) (BA 21).

#### Domain Activations

There was no significant activation, either at a cluster or voxel level, for Natural Kinds relative to Artifacts or Artifacts relative to Natural Kinds, at a corrected level of significance. Although some studies have reported selective activation for tools in the pos-

<sup>1</sup> All masking was inclusive and used a threshold of  $P < 0.05$ , uncorrected. Thus, an effect had to be present in each masking contrast at least with an uncorrected probability of 0.05

terior middle temporal gyrus we did not find any significant activation in this area, either at a corrected or uncorrected level. It may be argued that there is a possibility of type II errors with correction for multiple comparison. Therefore, we lowered the threshold to a more lenient  $P < 0.010$  uncorrected (SPM  $\{t\} = 2.65$ ). A single cluster reached significance for artifacts relative to natural kinds. This cluster was in GF (BA 37) and significant at  $P < 0.022$  (corrected) although it was not significant at a voxel level (peak at  $z = -44$ ,  $y = -62$ ,  $z = -16$ , SPM  $\{t\} = 5.61$ ,  $P < 0.94$  corrected). There was no significant activation for natural kinds relative to artifacts.

In addition, it may be that lack of significant activation in the domain-specific contrasts is due to signal drop-out which affects areas near air-tissue interfaces, particularly at high magnetic field strengths (see Devlin *et al.*, 2000, for a discussion of the susceptibility issue). Areas affected include regions of the temporal lobe near the sinuses; these are areas which are potentially of interest in studies of domain-specific knowledge. In order to reduce our search volume and increase our ability to detect activation in these areas we produced a mask image derived from our previous imaging data (Devlin *et al.*, 2000) to define a ROI. This was a PET study using the same task and items as in the present study and produced significant activations (at a corrected level) in left inferior temporal gyrus and left anteriomedial temporal pole. We masked the SPM  $\{t\}$  map of our fMRI data with this ROI and did a small volume correction calculation (Worsley *et al.*, 1996). Even with this drastically reduced search volume there were still no significant activations modulated by either artifacts or natural kinds.

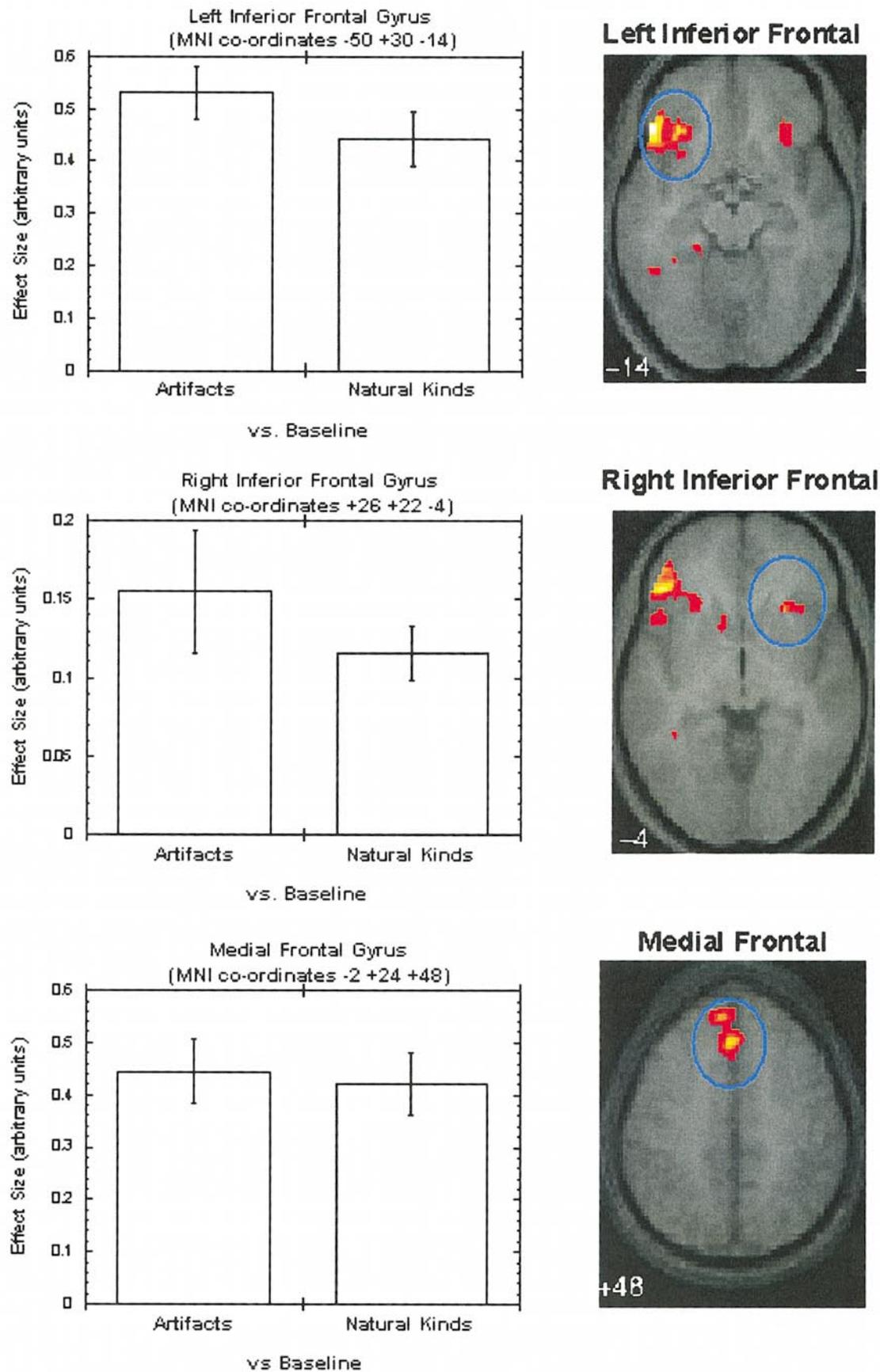
The observed lack of difference between artifacts and natural kinds carries the risk of type II error. While we accept that this possibility can never fully be excluded, we explored it further by looking more closely at each of the regions identified by the semantics-baseline contrast. Identifying each of the peaks we extracted parameter estimates for each subject and carried out paired  $t$  tests between conditions on each (i.e., comparing artifacts to natural kinds relative to baseline). In doing so, we reduced our threshold and avoided the requirement to carry out a conservative correction for multiple comparisons. Both of these measures reduce the risk of type II error. The peaks were in the left inferior frontal lobes, left caudate, left inferior temporal lobe, left middle temporal lobe and left medial frontal lobe (we did not perform a  $t$ -test on the peak in the right inferior frontal gyrus as it did not contain any datapoints at this voxel for all the subjects). Only one of these comparisons showed any difference which approached significance. This was in the left middle temporal lobe ( $P < 0.054$ ,  $t_{13} = 2.16$ ). As a further illustration of the lack of measurable differences in any region, except arguably the left middle temporal lobe, we have

plotted the parameter estimates for all conditions across each of these regions (see Fig. 2). In addition, we performed a small volume correction on the artifacts-natural kinds and natural kinds-artifacts contrasts. We specified a sphere of 10mm diameter centered at each of the peaks in the six significant clusters and used this to specify a ROI. There were no domain differences in any of these severely restricted areas, except in the artifacts-natural kinds contrast where there was a single small cluster of 6 voxels which reached significance (peak at  $z = -30$ ,  $y = -30$ ,  $z = -28$ , SPM  $\{t\} = 4.87$ ,  $P < 0.034$  corrected).

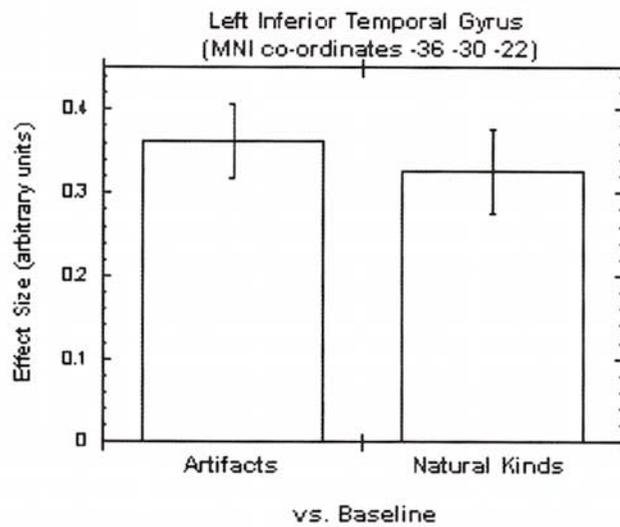
## DISCUSSION

Using an event-related fMRI design we found extensive frontal and temporal activations, primarily in the left hemisphere. A large cluster in the left inferior frontal lobe and a smaller cluster in a similar right frontal region are areas typically activated in semantic processing tasks (e.g., Devlin *et al.*, 2000; Thompson-Schill *et al.*, 1999; Mummery *et al.*, 1999; Vandenberg *et al.*, 1996; Tyler *et al.*, 2001). In addition we found activation in the medial aspect of the superior frontal lobe (GFd), in both left and right hemispheres. Similarly located activations have been found by both Devlin *et al.* (2001) and Cappa *et al.* (1998) using semantic categorization and semantic decision tasks, respectively. A smaller cluster was also activated in the left caudate. Both Mummery *et al.* (1998) and Price *et al.* (1997) found similar activation in the caudate for semantic judgements compared to syllable judgements. There were also two significant clusters of activation in the left temporal lobe. The first extended across the ventral aspect of the inferior temporal lobe and into the cerebellum. There was also a smaller cluster in the middle temporal gyrus. Similar temporal lobe areas have also been reported by Perani *et al.* (1999) with a lexical decision compared to letter detection task and Vandenberg *et al.* (1996) in a matching to sample task.

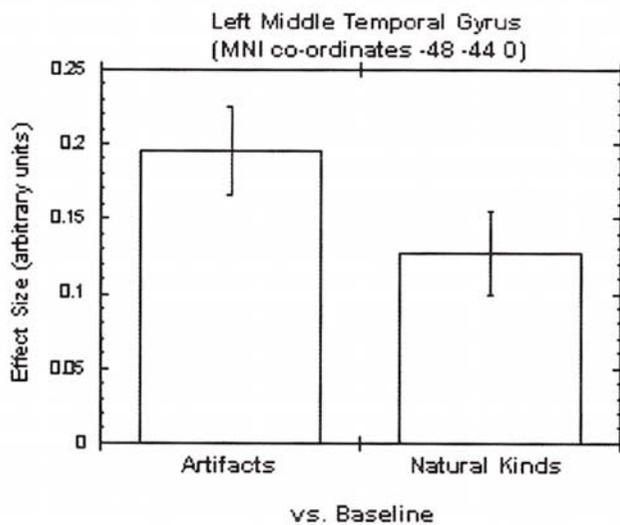
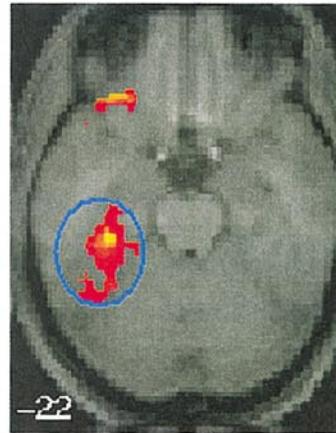
We found no significant activation, either at a cluster or voxel level, for natural kinds relative to artifacts at a corrected level. This replicates our previous findings (Devlin *et al.*, 2001) using the same task & items in both PET and blocked design fMRI. This result is also consistent with studies reviewed by Price and Friston (2001) and Devlin *et al.* (2001) in which no neural region was selectively activated for natural kinds. In addition, there was no significant activation for artifacts relative to natural kinds at a corrected level. Although some studies have reported selective activation for tools in the left posterior middle temporal gyrus (e.g., Chao *et al.*, 1999; Martin *et al.*, 1996; Moore and Price, 1999; Perani *et al.*, 1999) we did not find robust significant activation in this area.



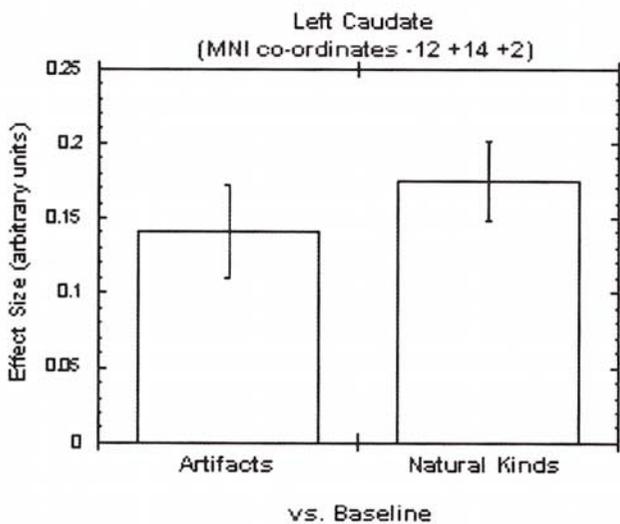
**FIG. 2.** Plots of parameter estimates at the peaks of each significant cluster.



## Left Inferior Temporal



## Left Middle Temporal



## Left Caudate

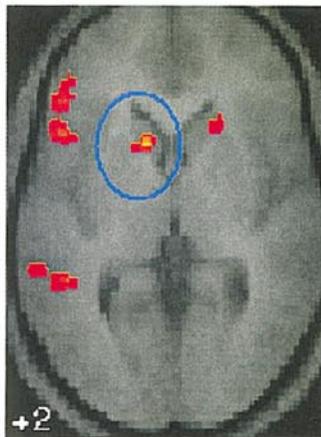


FIG. 2—Continued

We also compared the patterns of activation for all 14 participants tested in this event-related study and compared them with data we acquired from 8 participants in a blocked design study using the same task and materials (Devlin *et al.*, 2001). We found significant overlap in brain areas activated in these two studies. There was extensive left hemisphere activation in frontal and temporal areas for both designs. Our event-related study found several large clusters of activation in the frontal lobes one centring around the GF<sub>i</sub> (peaks in BA 47) and extending into the GF<sub>m</sub> (BA 9), another in the GF<sub>d</sub> and a smaller GF<sub>i</sub> activation in the RH. In the blocked design there was also a very large cluster in LH GF<sub>i</sub> and GF<sub>d</sub> and a homologous GF<sub>i</sub> activation in RH. There was some LH GF<sub>s</sub>/GF<sub>m</sub> activation in the blocked design that was not present in the event-related design. Both designs also showed temporal lobe activation. In the event-related study there was a large cluster of activation extending across the GT<sub>i</sub> and GT<sub>m</sub>. The cluster of activation in the blocked design was smaller and tended to be more superior with peaks in the GT<sub>s</sub> and GT<sub>m</sub> (−57 −37 0). It should be noted that the event-related activations are the result of a random effects level analysis, which is more stringent than the fixed effects analysis undertaken on the blocked design. A comparison of the patterns of temporal lobe activation produced by fixed effects analysis on the event-related data also shows GT<sub>m</sub>/GT<sub>s</sub> activation, with a peak activation which was almost identical (−59 −35 2).

Furthermore, neither study found significant domain-specific activation at a corrected level. Only with very lenient (uncorrected) thresholds did any of the comparisons show differential activation and these were not consistent across the two designs. This replication of findings across blocked and event-related designs suggests that lack of differences across domains are not merely the result of confounds such as habituation or of changes in the way people evaluate stimuli when confronted with items from a single domain being repeatedly presented.

To conclude, in an event-related fMRI study we found significant activation of an extended semantic network that encompassed many of the regions reported in previous imaging studies. These results show that event-related designs can identify areas involved in semantic processing. Given that event-related designs are both sensitive to semantic variables and overcome the confounds associated with blocking stimuli (i.e., psychological factors such as strategic effects), we anticipate their increased use in future fMRI studies of semantic processing. With respect to the neural representation of conceptual knowledge, these results are most compatible with claims for a conceptual system which is undifferentiated by domain (Tyler *et al.*, 2000; Tyler and Moss, 2001).

## ACKNOWLEDGMENTS

We thank members of the WBIC for their help in collecting data for this study. We also thank our volunteers for their participation. This work was supported by an MRC program grant and a McDonnell-Pew grant to L. K. Tyler and an MRC studentship to L. K. Pilgrim.

## REFERENCES

- Balota, D. A., and Black, S. 1997. Semantic satiation in healthy young and older adults. *Mem. Cognit.* **25**(2): 190–202.
- Cappa, S. F., Perani, D., Schnur, T., Tettamanti, M., and Fazio, F. 1998. The effects of semantic category and knowledge type on lexical-semantic access: A PET study. *Neuroimage* **8**(4): 350–359.
- Caramazza, A., and Shelton, J. R. 1998. Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *J. Cogn. Neurosci.* **10**(1): 1–34.
- Chao, L. L., Haxby, J. V., and Martin, A. 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat. Neurosci.* **2**(10): 913–919.
- Dale, A. M. 1999. Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* **8**(2–3): 109–114.
- den heyer, K., Briand, K., and Dannenbring, G. L. 1983. Strategic factors in a lexical-decision task—Evidence for automatic and attention-driven processes. *Mem. Cognit.* **11**(4): 374–381.
- Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., Matthews, P. M., and Tyler, L. K. 2000. Susceptibility-induced loss of signal: Comparing PET and fMRI on a semantic task. *Neuroimage* **11**(6 Pt 1): 589–600.
- Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Moss, H. E., Fadili, J., and Tyler, L. K. 2001. Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. Submitted.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-B., Frith, C. D., and Frackowiak, R. S. J. 1995. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2**: 189–210.
- Henson, R., Shallice, T., and Dolan, R. 2000. Neuroimaging evidence for dissociable forms of repetition priming. *Science* **287**(5456): 1269–1272.
- Henson, R. N. 2001. Repetition effects for words and nonwords as indexed by event-related fMRI: A preliminary study. *Scand. J. Psychol.* **42**(3): 179–186.
- Johnson, M. K., Nolde, S. F., Mather, M., Kounios, J., Schacter, D. L., and Curran, T. 1997. The similarity of brain activity associated with true and false recognition memory depends on test format. *Psychol. Sci.* **8**(3): 250–257.
- Josephs, O., and Henson, R. N. A. 1999. Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Philos. Trans. R. Soc. London Series B Biol. Sci.* **354**(1387): 1215–1228.
- Josephs, O., Turner, R., and Friston, K. 1997. Event-related fMRI. *Hum. Brain Mapp.* **5**(4): 243–248.
- Keefe, D. E., and Neely, J. H. 1990. Semantic priming in the pronunciation task—The role of prospective prime-generated expectancies. *Mem. Cognit.* **18**(3): 289–298.
- Kounios, J., Kotz, S. A., and Holcomb, P. J. 2000. On the locus of the semantic satiation effect: evidence from event-related brain potentials. *Mem. Cognit.* **28**(8): 1366–1377.
- Martin, A., and Chao, L. L. 2001. Semantic memory and the brain: structure and processes. *Curr. Opin. Neurobiol.* **11**(2): 194–201.
- Martin, A., Wiggs, C. L., Ungerleider, L. G., and Haxby, J. V. 1996. Neural correlates of category-specific knowledge. *Nature* **379**(6566): 649–652.

- Moore, C. J., and Price, C. J. 1999. A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* **122**(Pt5): 943–962.
- Mummary, C. J., Patterson, K., Hodges, J. R., and Price, C. J. 1998. Functional neuroanatomy of the semantic system: Divisible by what? *J. Cogn. Neurosci.* **10**(6): 766–777.
- Mummary, C. J., Shallice, T., and Price, C. J. 1999. Dual-process model in semantic priming: A functional imaging perspective. *Neuroimage* **9**(5): 516–525.
- Perani, D., Cappa, S. F., Bettinardi, V., Bressi, S., Gorno-Tempini, M., Matarrese, M., and Fazio, F. 1995. Different neural systems for the recognition of animals and man-made tools. *Neuroreport* **6**(12): 1637–1641.
- Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., and Fazio, F. 1999. The neural correlates of verb and noun processing: A PET study. *Brain* **122**(Pt 12): 2337–2344.
- Perani, D., Schnur, T., Tettamanti, C., Gorno-Tempini, M., Cappa, S. F., and Fazio, F. 1999. Word and picture matching: A PET study of semantic category effects. *Neuropsychologia* **37**(3): 293–306.
- Price, C. J., Moore, C. J., Humphreys, G. W. and Wise, R. J. S. 1997. Segregating semantic from phonological processes during reading. *J. Cogn. Neurosci.* **9**(6): 727–733.
- Price, C. J., and Friston, K. J. 2001. Functional imaging studies of category specificity. In *Category Specificity* (G. H. E. Forde, Ed.)
- Price, C. J., and Friston, K. J. 1997. Cognitive conjunction: A new approach to brain activation experiments. *Neuroimage* **5**(4 Pt 1): 261–270.
- Schacter, D. L., and Buckner, R. L. 1998. Priming and the brain. *Neuron* **20**(2): 185–195.
- Smith, L., and Klein, R. 1990. Evidence for semantic satiation—Repeating a category slows subsequent semantic processing. *J. Exp. Psychol. Learn. Mem. Cogn.* **16**(5): 852–861.
- Strayer, D. L., and Kramer, A. F. 1994. Strategies and Automaticity .2. Dynamic Aspects of Strategy Adjustment. *J. Exp. Psychol. Learn. Mem. Cogn.* **20**(2): 342–365.
- Thompson-Schill, S. L., Aguirre, G. K., D'Esposito, M., and Farah, M. J. 1999. A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia* **37**(6): 671–676.
- Tyler, L. K., and Moss, H. E. 2001. Towards a distributed account of conceptual knowledge. *Trends Cogn. Sci.* **5**(6): 244–252.
- Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., and Levy, J. P. 2000. Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain Lang.* **75**(2): 195–231.
- Tyler, L. K., Russell, R., Fadili, J., and Moss, H. E. 2001. The neural representation of nouns and verbs: PET studies. *Brain* **124**(Pt 8): 1619–1634.
- Tyler, L. K., Dick, E., Tavares, P., Pilgirm, L. K., Bright, P., Moss, H. E., and Fletcher, P. C. 2001. Is there neural specialisation for pictures of living and non-living things? Submitted.
- Vandenberghe, R., Price, C., Wise, R., Josephs, O., and Frackowiak, R. S. 1996. Functional anatomy of a common semantic system for words and pictures. *Nature* **383**(6597): 254–256.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., and Buckner, R. L. 1998. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* **281**(5380): 1188–1191.
- Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., and Buckner, R. L. 2000. Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex* **10**(12): 1176–1184.
- Warrington, E. K., and McCarthy, R. 1983. Category specific access dysphasia. *Brain* **106**(Pt 4): 859–878.
- Warrington, E. K., and McCarthy, R. A. 1987. Categories of knowledge: Further fractionations and an attempted intergration. *Brain* **110**: 1273–1296.
- Wiggs, C. L., and Martin, A. 1998. Properties and mechanisms of perceptual priming. *Curr. Opin. Neurobiol.* **8**(2): 227–233.
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., and Evans, A. C. 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum. Brain Mapp.* **4**(1): 58–73.