# A Locus in Human Extrastriate Cortex for Visual Shape Analysis

Nancy Kanwisher
Harvard University

Roger P. Woods, Marco Iacoboni, and John C. Mazziotta
University of California, Los Angeles

## Abstract

■ Positron emission tomography (PET) was used to locate an area in human extrastriate cortex that subserves a specific component process of visual object recognition. Regional blood flow increased in a bilateral extrastriate area on the inferolateral surface of the brain near the border between the occipital and temporal lobes (and a smaller area in the right fusiform gyrus) when subjects viewed line drawings of 3-dimensional objects compared to viewing scrambled drawings with no clear shape interpretation. Responses were seen for both novel and familiar objects, implicating this area in the bottom-up (i.e., memory-independent) analysis of visual shape. ■

## INTRODUCTION

Although neurophysiological studies over the last 25 years have provided a richly detailed map of the 30 or more different visual areas of the macaque brain (Felleman & Van Essen, 1991), the mapping of human visual cortex has only just begun. In the last few years functional brain imaging studies using PET and fMRI have discovered areas in human visual cortex that contain topographic representations of space (Sereno et al., 1995), as well as areas that respond selectively to particular dimensions of the visual world such as motion or color (Corbetta, Meizen, Dobmeyer, Shulman, & Petersen, 1990; Watson et al.,1993; Zeki, Watson, Lueck, Friston, Kennard, & Frackowiak, 1991). Sereno et al. (1995) have used fMRI to topographically map the borders of visual areas, and have recently proposed loci in the human cortex for the homologues of the first five visual areas in the macaque (V1-V5). In this paper we extend the mapping of human visual cortex to an extrastriate area involved in a higher level of visual information processing: object recognition.

Visual object recognition is a good candidate for a functional brain localization study because it has all the hallmarks of a modular process (Fodor, 1983)—that is, a process that uses dedicated brain hardware to carry out domain-specific computations.[1] First, humans are experts at visual recognition, able to recognize objects and scenes at rates of 8 or more items per second (Potter, 1976). Indeed, it is virtually impossible *not* to recognize a foveally-presented familiar object (Smith & Magee,

1980; Glaser & Dungelhoff, 1984). Such efficiency and automaticity are characteristic of modular processes (Fodor, 1983). Second, abilities critical for survival (such as rapid and accurate object recognition) are the ones most likely to be refined by natural selection through the construction of innate and highly specialized brain modules (Dawkins, 1986; Pinker, 1994). Third, visual object recognition is a computationally complex task that is not likely to be handled well by general-purpose cognitive processes but instead may require its own special-purpose computational machinery invoking constraints and rules specific to object recognition (Marr, 1982). Thus theoretical and behavioral considerations alone suggest that human cortex may contain area/s specialized for visual object recognition. Finally, as will be discussed below, a great deal of evidence from neurophysiology and neuropsychology supports the conclusion that visual object recognition is subserved by one or more dedicated brain modules.

What computations would take place in a brain area specialized for object recognition? Most theorists agree that visual recognition consists of (at least) three main component processes: (1) extraction of simple visual features from the retinal image of the object, (2) construction of abstract representation of the object's shape, and (3) matching that shape description to a particular stored representation in memory. In the present study we used PET to investigate the brain locus of the second component process, visual shape analysis. Although few past studies have attempted to distinguish between these three components of object recognition, a great

deal of previous work in neurophysiology, neuropsychology, and functional brain imaging has provided a number of candidate brain areas that appear to be involved in some aspect of visual recognition. We review this work next.

## The Brain Locus of Visual Recognition

Neurophysiology and lesion studies in monkeys have shown that extrastriate cortex is segregated into a dorsal pathway specialized for processing information about spatial location and action, and a ventral pathway specialized for visual recognition (Ungerleider & Mishkin, 1982; Goodale & Milner, 1992). Similar findings have been reported from human patients with focal brain damage: parietal patients have impaired visual attention and spatial localization abilities, whereas patients with ventral occipital and temporal damage can exhibit a variety of deficits in the visual recognition (Farah, 1990) of words (alexia), objects (agnosia), faces (prosopagnosia), and colors (achromatopsia).

Recent PET studies also support a division between the "what" and "where" pathways in human visual cortex. Haxby et al. (1994) compared the areas activated when subjects indicated which of two faces matched a third face with the areas activated when subjects carried out an analogous location-matching task on the same stimuli. They found a large ventral area in the occipitotemporal region that was more strongly activated in the face-matching task than the location-matching task.

A second question is whether face recognition uses the same brain machinery as object recognition. Double-dissociations in brain-damaged patients suggests that it may not: some patients are impaired at face but not object recognition, and others have the reverse pattern of deficit (Farah, 1990, 1991). Sergent, Ohta, and MacDonald (1992) reported a PET study designed to ask whether there are distinct brain regions specialized for face and object recognition. They found that face recognition primarily activated a ventro-medial region in the right hemisphere, whereas object recognition primarily activated an occipitotemporal region of the left hemisphere. While these findings are suggestive, the particular subtractions used in this study leave the results open to multiple interpretations. For example, to find areas involved in visual object recognition, Sergent et al. (1992) compared the brain areas that were activated when subjects categorized photographs of objects as living or nonliving to those that were active when subjects judged the orientation of sine-wave gratings. This comparison cannot determine whether the areas activated reflect (1) extraction of any visual features (other than those included in the gratings), (2) matching to stored visual representations, (3) covert object naming, or (4) extraction of the meaning of the objects recognized. Many other studies that have used functional brain imaging to investigate visual recognition use subtractions that are

similarly broad and that not surprisingly activate large or multiple regions of occipital and temporal cortex (see Kanwisher, Chun, McDermott, & Ledden, in press, for a review and critique).

## Goals and Design of the Present Study

The goal of the present study was to search for brain loci of component processes that should be entailed in virtually any model of object recognition. The design of this study was based on several considerations. First, because people can recognize objects quickly we attempted to increase our signal by keeping the visual recognition system active with a rapid stimulus presentation rate of two items per second (but see also Price, Wise, Watson, Patterson, Howard, & Frackowiak, 1994). This rapid rate is also likely to reduce subjects' tendency to covertly name each object presented. Second, because visual object recognition is automatic in the sense that it is virtually impossible to avoid recognizing a familiar object presented clearly at the fovea (Smith & Magee, 1980; Glaser & Dungelhoff, 1984), we reasoned that it would not be necessary to have subjects carry out an overt task in order to invoke object recognition processes. The present study simply asked subjects to passively view the stimuli as they were presented, allowing us to control the mental processes that were carried out in each condition simply by varying the stimuli (see also Petersen, Fox, Snyder, & Raichle, 1990). Third, we wanted to localize visual shape analysis in particular, unconfounded from lower-level feature extraction processes on the one hand and higher-level memory-dependent processes on the other.

To accomplish these goals we constructed three sets of line drawing stimuli with 140 items in each (see Fig. 1). The first ("familiar") set contained drawings of familiar common objects, adapted from the Snodgrass & Vanderwart (1980) set. The second ("novel") set contained drawings that were designed to resemble the familiar-object drawings in every possible way (e.g., complexity, depth, part structure, approximate number of curved versus straight lines) except for familiarity: items in this set do not match any known object. (Though some items may *suggest* a particular known object to some subjects, few or none would be judged to actually *depict* that object.) The third ("scrambled") set was constructed from the familiar-object set in such a way as to exactly match the two sets (summed over all 140 items in each set) for average luminance at each pixel, total line contour length, average line curvature, local orientation, and many other low-level visual features. For each scan, subjects viewed a sequence of all 140 objects from one of the sets, presented on a computer monitor at a rate of 2 items/second.

The logic of this design (see Fig. 1), which is analogous to an earlier study of word recognition (Petersen, Fox, Snyder, & Raichle, 1990),[2] is as follows. Because visual
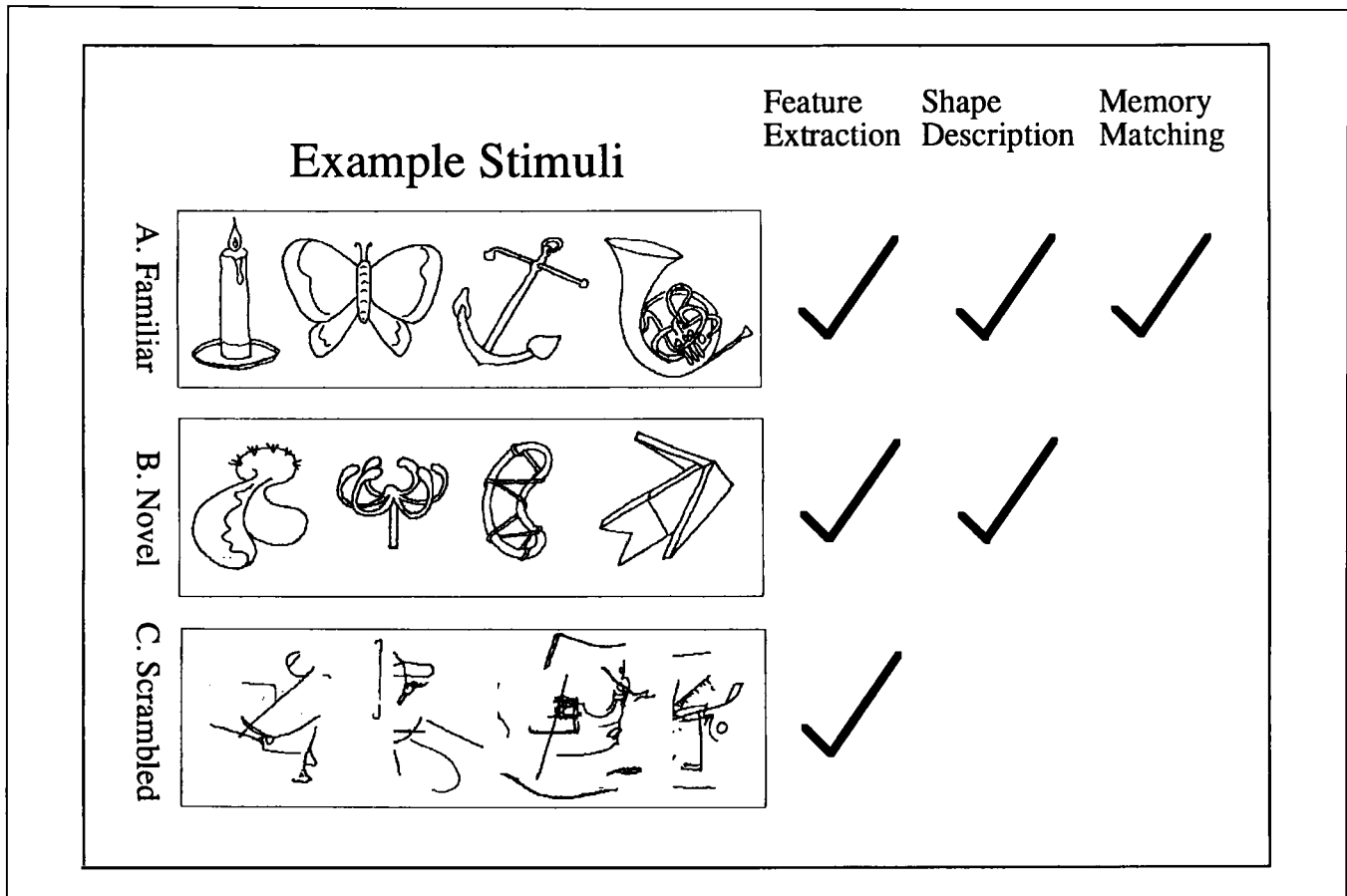
**Figure 1.** Sample stimuli from the three conditions and the component processes likely to be engaged by passive viewing of each.

object recognition proceeds automatically (i.e., independent of task) for clearly-presented visual stimuli under normal conditions, passive viewing of the familiar objects should cause subjects to carry out all three component processes of object recognition: (1) extraction of the visual features from the image, (2) construction of a representation of the object's shape, and (3) matching that shape description to a stored representation of that object in memory. For the novel objects, however, only the first two processes should be successfully completed, because these objects do not match any object stored in memory. Finally, in the scrambled condition, only low-level feature extraction processes should be fully engaged, because the scrambled objects do not have clear interpretations in terms of 3-dimensional shape and do not match any known objects in memory.

Thus to find any brain areas specifically involved in the bottom-up construction of shape descriptions, while avoiding confounds from either feature extraction or memory-matching, we simply compared blood flow in the novel-object condition to that in the scrambled-object condition. To exclude the possibility that the differences seen in this comparison are due to the novelty of the shapes per se, it is important to know whether they also occur when subjects view familiar objects (compared to scrambled objects). Finally, any brain areas involved in the successful matching of visual descriptions to memory and/or accessing the meanings (or names) of familiar objects should be revealed in the comparison of the familiar to the novel conditions.

Each subject underwent 12 scans, four in each of the three conditions (with order counterbalanced across subjects). Two pilot subjects were run first in order to define an a priori region and specific hypotheses. These prior hypotheses were then the basis for one-tailed a priori comparisons within the same regions in six new subjects in the main study.

## RESULTS

Data from two pilot subjects revealed bilateral inferior occipitotemporal areas where relCBF was higher in the novel and familiar conditions than the scrambled condition. This area (outlined in red in Fig. 2) includes two contiguous regions, one in the right hemisphere and the other in the left hemisphere, where the average of the counts in the novel and familiar conditions was greater than in the scrambled condition in the pilot data. No clear differences were found between familiar and novel conditions. The bilateral areas identified in the pilot study provided a priori regions, within which we hypothesized that both (1) the novel condition and (2) the
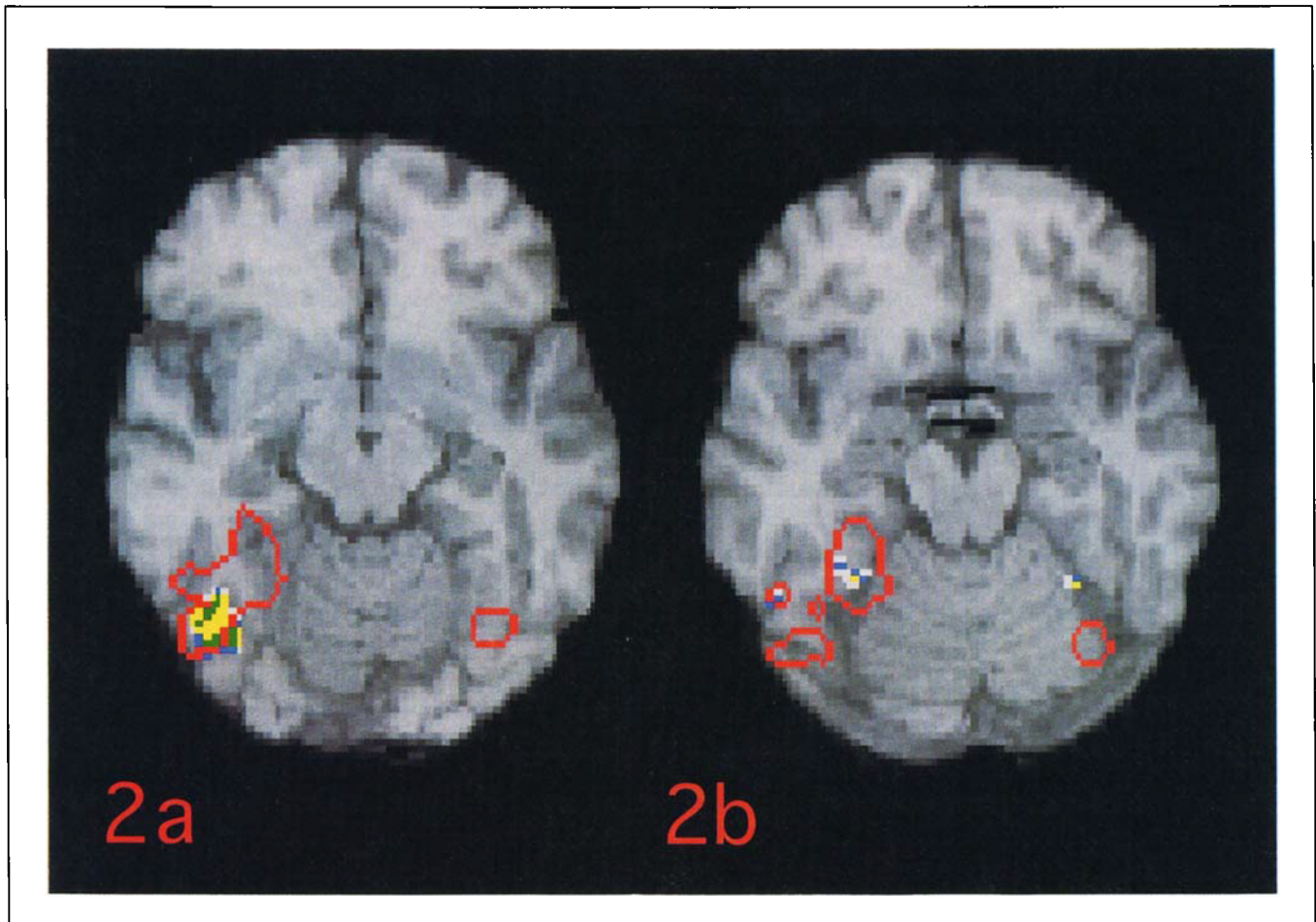
**Figure 2.** (A) The gray-level images show two different transverse sections taken from the anatomical MRI of one of the subjects in our main study. The red outlines indicate the borders of the a priori area derived from the two pilot subjects. The statistical results of the data from the 6 subjects in the main study are shown in the other colors. For the analysis of the 6 subjects in the main study, we corrected for multiple comparisons by computing critical $t$-statistic thresholds based on the size and shape of the prior area and the image resolution (Worsley et al., 1995); the resulting corrected $p$-levels within the red-outlined region are $p < 0.05$ (shown in white), $p < 0.025$ (shown in green), $p < 0.01$ (shown in blue), and $p < 0.001$(shown in yellow), all one-tailed. Figure 2a shows the voxels that reached significance in the comparison of the novel and scrambled conditions; Figure 2b shows the results of the comparison of familiar and scrambled conditions for a slice 10.4 mm more inferior. While the significant regions shown in these two slices are both in the right hemisphere (shown here in conventional radiological orientation on the left), numerous voxels in both hemispheres reached significance in other slices not shown here. Note that because of our conservative correction for multiple spatial comparisons even a single colored pixel within the red-outlined prior region is a significant result.

familiar condition would result in higher relCBF than the scrambled condition. These two prior hypotheses were the basis for one-tailed a priori comparisons within the same regions in six new subjects in the main study.

The data from the six subjects in the main experiment were analyzed in a 3-way (subject × condition × replication number) ANOVA with a planned comparison of means using linear contrasts (Hays, 1988); 3-way interactions were assumed to be due to random error but all other interactions were modeled. Critical $t$-statistic thresholds (30 degrees of freedom) were computed based on the image resolution and the size and shape of the a priori search region (Worsley, Marrett, Neelin, & Evans, 1995), which had a total volume of 13.8 cubic centimeters: $t > 4.29$ ($p < 0.05$, one-tailed) shown in white in Figure 2; $t > 4.59$ ($p < 0.025$, one-tailed) shown

in green in Figure 2; $t > 4.97$ ($p < 0.01$, one-tailed) shown in blue in Figure 2; $t > 5.91$ ($p < 0.001$, one-tailed) shown in yellow in Figure 2. This analysis revealed many voxels within our prior region where relCBF was significantly higher in the novel than the scrambled condition, and many in which relCBF was significantly higher in the familiar than scrambled condition.

The voxels that showed significantly higher relCBF in the novel condition as compared to the scrambled condition are adjacent to, but not always identical to the voxels that showed higher relCBF in the known condition as compared to the scrambled condition. There are two possible interpretations: (1) Anatomically distinct areas responded to novel stimuli and to known stimuli in these regions; (2) Anatomically identical areas responded to both novel and known stimuli with statistical

fluctuations or quantitative (as opposed to qualitative) differences in response magnitudes accounting for the observed differences in precise location. Examination of simple percentage change images supports the latter interpretation. The areas that meet the criteria for statistical significance in the two comparisons are at the maxima of much broader, extensively overlapping regions of increased blood flow. Unequivocal proof of the hypothesis that anatomically distinct sites responded would require demonstration that the novel and known tasks produced significantly different blood flow patterns. Yet in a post hoc comparison of our familiar and novel conditions, no voxels showed a significant difference between the two tasks. Our original intent in using both novel and known stimuli was to allow us to distinguish areas involved in shape analysis from areas not involved in shape analysis. For example, an area responding only to the novelty of the novel objects should not be identified when comparing the known objects to the scrambled objects. Likewise, an area responding only to the nameability of the known objects should not be identified when comparing the novel objects to the scrambled objects. It is unlikely that two independent processes, one responding to the novel stimuli, the other responding to the known stimuli, and each distinct from shape analysis, would independently produce significant blood flow responses in almost identical anatomic locations. The most parsimonious and reasonable interpretation of our findings is that both sets of relCBF increases reflect a single process carried out on both sets of stimuli: shape analysis.

Figure 3 shows a 3-dimensional rendering of the three main areas where novel and familiar stimuli resulted in greater relCBF than the scrambled stimuli. The two more posterior areas are on the inferolateral surface of the brain near the border between the occipital and temporal lobes at the junction of Brodmann areas 19 and 37, with the left-hemisphere activation in a more superior and medial location. The positions of these two areas suggest that they may be right and left hemisphere homologues of one another. The anterior right-hemisphere area is in the fusiform gyrus (Brodmann area 37); its homologous region in the left hemisphere was not included in our a priori region because we chose to define only a single a priori region on each side and the pilot data used to define the a priori regions showed less extensive blood flow changes on the left.

In order to compare the locus of the area activated in the present study with those of other reported studies, the average of the 6 subjects' MRI scans was mapped onto a Talairach space (Talairach & Tournoux, 1988) MRI template (using the 12-parameter affine registration algorithm of Woods, Mazziotta, and Cherry (1993b), which allowed us to derive the following coordinates for the three areas: 43R, −61, −16 and 35L, −60, −10, and 35R, −42, −16.

The repetition of stimulus materials across four replications of each condition did result in the learning of many of the originally novel items, as determined by an informal recognition test conducted after the scan for novel items in the experiment versus new ones the subject had not yet seen. However, this learning of the novel shapes is unlikely to explain the observed similarity in results for the novel and familiar conditions because this similarity was observed in the first scan in each condition (before the novel items were learned), and there was no trend of an interaction of condition by replication number.



**Figure 3.** These images show in red the three contiguous regions that (1) had some overlap with our a priori region of interest, and (2) in which the maximum t-value for the linear contrast of the average novel and familiar conditions to the scrambled condition was significant. These data are superimposed on a 3D rendering of the anatomical MRI from the same subject used in Figure 2. For stereo viewing the two leftmost images are crossed and the right two uncrossed.
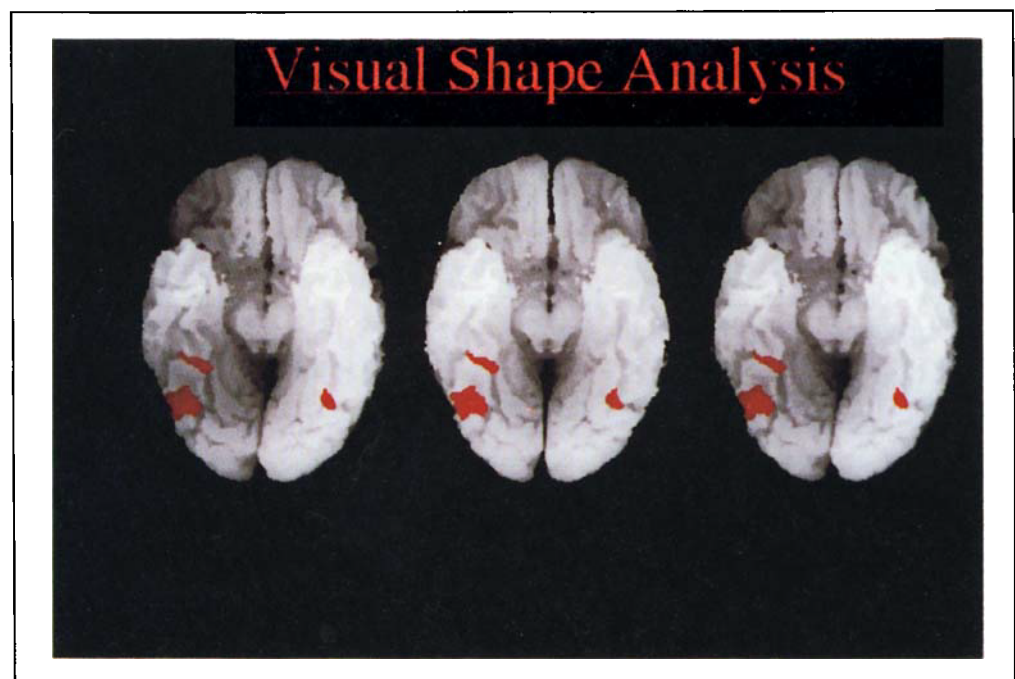
In order to assess any variability across subjects, we looked for voxels that showed a significant interaction of subject by stimulus type, using our predefined anatomic search region to set the statistical threshold. There were no voxels anywhere in the brain that showed a significant interaction at the $p < 0.05$ level. Given this lack of evidence for statistically significant differences between subjects, the different patterns observed in individual subjects' data are unlikely to be meaningful and we do not show them here. However it should be noted that the lack of a significant interaction should not be taken as strong evidence for consistency across subjects because of the low statistical power available in this test. An illustration of the noisiness of the individual-subject data is provided in Table 1, which shows the results for a single voxel across all four replications of each of the three conditions across the six experimental subjects.

In order to make sure that we were not missing an area of greater activation outside our a priori region of interest, we performed a 3-way ANOVA to look for task effects anywhere in the brain using a whole-brain correction for multiple comparisons. The only region that exceeded the $p < 0.05$ threshold was the large right hemisphere activation site already identified as an area of blood flow increase and included in our a priori anatomic search region.

## DISCUSSION

This study has found a bilateral extrastriate area on the inferolateral surface of the brain near the occipitotemporal junction that shows significantly increased relCBF in response to stimuli supporting clear 3-dimensional shape interpretations compared to stimuli containing similar low-level features but no obvious shape interpretation. Because this area responded as strongly to novel as to familiar objects, the response is unlikely to reflect processes associated with memory-matching, naming, or accessing semantic information. We therefore propose that this area is involved in the bottom-up construction of shape descriptions from simple visual features.

The Talairach coordinates computed for our shape area suggest that it is inferior to the area proposed by Zeki et al. (1991) as the human homologue of macaque area V5 and is lateral and inferior to Zeki's proposed area V4. This locus for visual shape analysis is consistent with observations that deficits in object recognition result from bilateral damage to ventral occipitotemporal cortex (Damasio, Tranel, & Damasio, 1989; DeRenzi, Scotti, & Spennler, 1969; Feinberg, Schindler, Ochoa, Kwan, & Farah, 1994). Further, while our more anterior area in the right fusiform gyrus overlaps with areas previously reported to be involved in visual face recognition (Haxby

**Table 1.** An illustration of the intersubjective variability in this experiment: This table shows the data for a single voxel in the main study for four replications × three conditions × six subjects. This voxel had the highest $F$ score of any voxel in the brain on a 3-way task ANOVA, is located in the large region of the right hemisphere indicated in Figure 3, and shows significant differences both between the known and scrambled conditions and between the novel and scrambled conditions.

| | Sub 1 | Sub 2 | Sub 3 | Sub 4 | Sub 5 | Sub 6 |
|---|---|---|---|---|---|---|
| Familiar | | | | | | |
| Rep. 1 | 1.084152 | 1.177580 | 1.183852 | 1.157948 | 1.179909 | 1.143153 |
| Rep. 2 | 1.072584 | 1.128712 | 1.180769 | 1.204823 | 1.147381 | 1.174065 |
| Rep. 3 | 1.099302 | 1.138578 | 1.116677 | 1.118239 | 1.151800 | 1.078001 |
| Rep. 4 | 1.106459 | 1.140955 | 1.209975 | 1.178523 | 1.201816 | 1.169382 |
| Novel | | | | | | |
| Rep. 1 | 1.125693 | 1.217879 | 1.179030 | 1.211345 | 1.183381 | 1.162080 |
| Rep. 2 | 1.043933 | 1.256304 | 1.191232 | 1.193708 | 1.182323 | 1.15098 |
| Rep. 3 | 1.080516 | 1.206103 | 1.139267 | 1.153772 | 1.218811 | 1.099855 |
| Rep. 4 | 1.091433 | 1.151493 | 1.213094 | 1.166995 | 1.153778 | 1.103725 |
| Scrambled | | | | | | |
| Rep. 1 | 1.074013 | 1.101352 | 1.118209 | 1.134501 | 1.108966 | 1.190776 |
| Rep. 2 | 1.040778 | 1.135098 | 1.077233 | 1.151949 | 1.144295 | 1.094437 |
| Rep. 3 | 1.050964 | 1.140219 | 1.111236 | 1.130115 | 1.104968 | 1.078938 |
| Rep. 4 | 1.042348 | 1.057080 | 1.092597 | 1.149366 | 1.143617 | 1.130068 |

et al., 1994; Allison, McCarthy, Nobre, Puce, & Belger, 1994), the bilateral posterior area is centered near locations implicated in object (Kosslyn et al., 1994) but not face recognition (Haxby et al., 1994; Sergent et al., 1992). Further, the areas found in this study (Fig. 3) do not overlap with areas previously implicated in visual word recognition (Petersen et al., 1990). While these results will need to be replicated within subjects, they suggest that the bilateral posterior area found in the present study is specifically involved in the perceptual analysis of visual objects and is distinct from areas involved in visual face and word recognition.

Four recent functional imaging papers have addressed the brain locus of visual object recognition or related processes. In one using PET, Kohler, Kapur, Moscovitch, Winocur, and Houle (1995) showed subjects pairs of two sequentially-presented displays each containing three objects and asked them to judge in one condition whether the three locations were the same, and in another condition whether the three objects were the same. Areas that were significantly more active in the identity task than the location task included the inferior temporal cortex in the region of the fusiform gyrus (Brodmann areas 19 and 37) in the left hemisphere, extending posteriorly into the lingual gyrus (Brodmann areas 18 and 17), and in the ventral occipital cortex of the right hemisphere in the region of the fusiform gyrus. While this region of activation may overlap somewhat with the shape area reported in the present study, it is generally larger, more posterior, and more left-lateralized. This difference is not surprising given that Kohler et al.'s identity matching task could be done on the basis not only of shape per se but also lower-level visual features, names, or meanings.

In another recent PET study, which focused more specifically on visual shape extraction (Schacter et al., 1995), subjects viewed line drawings of novel objects that were either physically possible or impossible. Compared to passive viewing of the same stimuli, an "object decision" task (deciding whether the objects were possible or not) activated areas in the inferior temporal and inferior fusiform gyri—but only for the physically possible objects. (The possible and impossible objects were segregated into different scans, so that the correct response was the same for all 20 items in each scan.) In a second comparison, the effect of a prior exposure to a particular stimulus (i.e., priming) was assessed by subtracting the activation resulting when subjects performed object decisions on previously- unstudied objects from the activation resulting from object decisions on previously-studied objects.[3] This activation, which was also observed only for possible, not impossible, objects was in the inferior temporal and fusiform gyri of the left hemisphere. These data are taken as evidence that the inferior temporal and fusiform regions are "selectively involved in computing global representations of structurally coherent three-dimensional objects" (Schacter et al., 1995, p. 590).

Although it would seem likely that the areas activated in the Schacter et al. study might be the same as those found in the present study, as both are apparently involved in visual shape analysis, it should be noted that the Tailarach coordinates for the Schacter et al. area are anterior to ours. One important difference between the Schacter et al. study and ours is their use of an overt task rather than passive viewing of the stimuli. If shape processing is indeed automatic, as we have argued above, the same computations should be carried out no matter what the subject is attempting to do and task manipulations should be ineffective. On the other hand, the distinction between Schacter et al.'s possible and impossible objects is very subtle and one's subjective impression when carrying out the object decision task is that it requires effort and careful scrutiny of the stimuli. Thus it is possible that while some aspects of shape analysis proceed automatically and in a task-independent fashion, other more subtle aspects of shape analysis occur only when required by the task. If the computations underlying the possible/impossible decision are of the nonautomatic kind, then we might find no difference between Schacter et al.'s possible and impossible objects in a passive-viewing situation. If so, then a reasonable conclusion would be that the present study has revealed an area involved in the automatic aspects of shape analysis, whereas the Schacter et al. study has revealed an area involved in a more effortful and optional component of shape analysis.

The paper most relevant to the present work is an fMRI study reported recently by Malach et al. (1995) that independently arrived at a design very similar to ours. Malach et al. report a new extrastriate area ("LO," for lateral occipital complex) at the lateral-posterior aspect of the occipital lobe just posterior to area MT and argue that this area is involved in an intermediate stage of visual object recognition. This claim is based on the fact that area LO responded more strongly to photographs of familiar objects, famous faces, and unfamiliar 3-dimensional objects, compared to texture fields. The similarity of the design and the results of this study with fMRI and ours with PET is reassuring and suggests that the area identified in both studies may be the same. However, this conclusion may be premature. First, Malach et al.'s LO is posterior to MT, whereas our area appears to be if anything anterior to MT. Second, although no feature control is perfect, we suspect that our scrambled and familiar objects are more closely matched than Malach et al.'s objects and textures are in terms of the component visual features (e.g., edges and contours) likely to be extracted by early visual processing. If so, this would provide an explanation of the more posterior locus of LO relative to the area described here: LO would include not only areas involved in visual shape analysis

but also areas involved in feature extraction. Further work will be necessary to understand exactly why and how the two areas differ.

Finally, Martin, Wiggs, Ungerleider, and Haxby (1996) very recently published a study that, like ours, used PET to compare the activations resulting when subjects viewed line drawings of familiar or novel objects versus visual noise patterns. They found an area active in the novel versus noise comparison that was in the inferior occipital/fusiform region, near the region reported in the present study but more posterior. The more posterior locus of their activation can probably be attributed to the fact that their noise pattern stimuli were so feature-less that their novel versus noise subtraction would be expected to identify areas involved not only in visual shape analysis (as in our study) but also areas involved in many aspects of early visual feature extraction as well. More surprising is the fact that Martin et al. found some areas that were more active when subjects viewed familiar objects than novel ones, whereas we found none. This result can be explained, however, by noting that the subjects in Martin et al.'s study were asked to covertly name the familiar objects but simply stare at the novel ones. Thus the areas active in Martin et al.'s familiar versus novel comparison most likely reflect not only visual recognition per se but also accessing semantic information about the pictures (a necessary prerequisite to naming), covert naming, and even visual attention (because subjects performed a task in the familiar condition but not in the novel condition). Thus, each of the two subtractions that Martin et al. report that are closest to our own include many different component mental processes, so these subtractions are not very useful for isolating specific component processes of visual recognition.

Recent experiments using fMRI have replicated the basic result reported here (Kanwisher, Chun, McDermott, & Ledden, in press). Although the effect failed to reach significance in many of the individual fMRI subjects tested (as was true for the individual subjects in the present PET study), four of the subjects revealed clear and significant effects in individual scans in a bilateral occipitotemporal region that appears to be the same as the region reported in the present study. A similar region has also been observed to be activated in fMRI comparisons of passive viewing of "thresholded" photographs of objects (in which all intermediate grey levels below a certain brightness have been assigned to black and all brighter regions assigned to white) with "scrambled" versions of the thresholded photographs (Kanwisher et al., in press), also consistent with the present claim that this area is involved in the extraction of object shape.

Future work with fMRI should enable us to determine the exact computations that are being carried out in the extrastriate area described here. While we have argued that it is involved in the analysis of visual shape, it is not yet clear whether it is more specifically involved in

figure-ground segregation, part decomposition, depth interpretation, or some other component of shape processing. The answers to these questions will be of interest not only to neuroscientists but also to cognitive psychologists. Indeed, the most exciting promise of functional brain imaging is the possibility that it will enable us to exploit the modular structure of the brain to discover the functional components of the mind.

## METHODS

*Stimuli.* The pictures in the novel-object set were drawn by hand by one of the authors (NK) to resemble the familiar objects in every possible way except for familiarity. The familiar objects, which were taken from the Snodgrass and Vanderwart (1980) set, were then redrawn by the same person to match the line width and handwriting used in constructing the novel set. These pictures were then scanned into PICT format. The scrambled objects used in the main study were created by first digitally dividing each of the 140 familiar object stimuli into five line-component subsets (that re-created the entire familiar object without overlap when superimposed), and then superimposing 5 subset images from 5 different objects (sampling without replacement) to make each scrambled image. Thus, the scrambled and familiar stimulus sets were perfectly matched for average luminance at each pixel, total line contour length, and most simple visual features (although object structure inevitably co-varies to some extent with spatial frequency and with certain visual features like T-junctions). The scrambled stimuli used in the pilot study were created with a different scrambling procedure but looked generally similar to those used in the main study.

*Subjects.* Subjects were 5 men and 3 women (one man and one woman in the pilot study) ranging in age from 19 to 35 years. Informed consent was obtained from all subjects in accordance with protocols approved by the UCLA Human Subjects Protection Committee. Each subject underwent a total of 12 blood flow measurements (four replications for each of the three conditions).

*Equipment and Procedure.* Images were acquired using a Siemens/CTI 831 scanner with an axial field of view of 101.25 mm. A customized foam headholder was used to reduce head motion. For each relCBF measurement, subjects were injected with a 10- mCi (370MBq) bolus of $H_2^{15}O$ in 7 cc of normal saline. Using the partial switch technique described by Cherry, Woods, Doshi, Banerjee, and Mazziotta (1995) to improve signal-to-noise, counts in each study were collected in a single 120-second time frame beginning at the time of injection (10 seconds after subjects began viewing the stimulus sequence) and spanning a further 60 seconds of stimulus viewing plus an additional 60 seconds of viewing a fixation cross.

The scanning procedure and 3D image reconstruction

have been described elsewhere (Cherry, Woods, Hoffman, & Mazziotta, 1993). Original data sets consisted of 15 planes with an interplane distance of 6.75 mm. The images were smoothed using a Gaussian filter to a final isotropic resolution of 10 mm. Head movements were corrected using the algorithm described by Woods, Cherry, and Mazziotta (1992). Global normalization was applied (Mazziotta, et al., 1985). In separate sessions, each subject returned for magnetic resonance imaging (MRI) of the brain performed on a GE Signa scanner using a 3D spoiled GRASS sequence. MRI-PET registration was performed as described in Woods, Mazziotta, and Cherry (1993a). Intersubject stereotaxis based on the MRI data was performed using a 12-parameter affine registration model, as described by Woods, Mazziotta, and Cherry (1993b). After mapping all data into a common space, global normalization was applied (Mazziotta et al., 1985), and the a priori hypotheses were tested using an analysis of variance model.

## Acknowledgments

## Notes

1. Employing dedicated brain hardware is only one of several criteria for modularity discussed by Fodor. Others include the requirements that modules (1) run automatically when supplied with appropriate input, (2) have only restricted access to representations outside the module ("informational encapsulation"), and (3) are largely innately determined.

2. The study of word recognition by Petersen et al. (1990) compared the responses to passive viewing of familiar words (analogous to our familiar object pictures), orthographically regular nonwords (analogous to our novel objects), orthographically irregular nonwords, and false-fonts (analogous to our scrambled objects).

3. The fact that priming produced higher rather than lower activation for studied (compared to unstudied) stimuli differs from earlier priming results (Squire et al., 1992). This difference might reflect the fact that previous studies have presented stimuli well above threshold, whereas the stimuli in the Schacter study were displayed for only 50 ms apiece. In other studies priming might manifest itself via easier processing of primed than unprimed objects, producing lower activation for primed stimuli. However in the Schacter study it is possible that 50 ms

would be sufficient to produce a robust perceptual representation only for primed, not unprimed, objects, producing greater activation for the primed stimuli. The very low performance on object decision for unprimed objects is consistent with this interpretation.

## References

Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex, 5*, 544-554.

Cherry, S. R., Woods, R. P., Hoffman, E. J., & Mazziotta, J. C. (1993). Improved detection of focal cerebral blood flow changes using three-dimensional positron emission tomography. *Journal of Cerebral Blood Flow Metabolism, 13*, 630-638.

Cherry, S. R., Woods, R. P., Doshi, N. K., Banerjee, P. K., & Mazziotta, J. C. (1995). *Journal of Nuclear Medicine, 36*, 307-314.

Corbetta, M., Meizen, F. M., Dobmeyer, S., Shulman, G. L., Petersen, S. E. (1990). Attentional modulation of neuronal processing of shape, color, and velocity in humans. *Science, 248*, 1556-1559.

Damasio, A., Tranel, D., & Damasio, H. (1989). In F. Boller, & J. Grafman (Eds.), *Handbook of Neurology, Vol. 2.* Amsterdam: Elsevier Science Publishers.

Dawkins, R. (1986). *The blind watchmaker.* New York: Norton.

DeRenzi, E., Scotti, G., & Spennler, H. J. (1969). Perceptual and associative disorders of visual recognition: Relationship to the side of the cerebral lesion. *Neurology, 19*, 634-642.

Farah, M. (1990). *Visual agnosia.* Cambridge, MA: MIT Press.

Farah, M. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology, 8*, 1-19.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1-47.

Feinberg, T. E., Schindler, R. J., Ochoa, E., Kwan, P. C., & Farah, M. J. (1994). Associative visual agnosia and alexia without prosopagnosia. *Cortex, 30*, 395-412.

Fodor, J. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Glaser, W. R., & Dungelhoff, F. I. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception & Performance, 10*, 640-654.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience, 15*, 20-25.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. *Journal of Neuroscience, 12*, 6336-6353.

Hays, W. L. (1988). *Statistics.* New York: Holt, Rinehart, & Winston.

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (in press). Functional Imaging of Human Visual Recognition. *Proceedings of the ninth Toyota conference on brain and mind.* Elsevier.

Kohler, S., Kapur, S., Moscovitch, M., Winocur, G., & Houle, S. (1995). Dissociation of pathways for object and spatial vision: A PET study on humans. *NeuroReport, 6,* (14) 1865-1868.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F.,

Rauch, S. L., & Anderson, A. K. (1994). Identifying objects seen from different viewpoints: A PET investigation, *Brain*, *117*, 1055-1071.

Malach, R., Reppas, J. B., Benson, R. B., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. H. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences USA*, *92*, 8135-8138.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649-652.

Mazziotta, J. C., Huang, S. C., Phelps, M. E., Carson, R. E., MacDonald, N. S., & Mahoney, K. (1985). A noninvasive positron computed tomography technique using oxygen-15-labeled water for the evaluation of neurobehavioral task batteries. *Journal of Cerebral Blood Flow Metabolism*, *5*, 70.

Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, *249*, 1041-1044.

Pinker, S. (1994). *The language instinct*. New York: William Morrow.

Price, C. J., Wise, R. J. S., Watson, J. D. G., Patterson, K., Howard, D., & Frackowiak, R. S. J. (1994). Brain activity during reading: The effects of exposure duration and task. *Brain*, *117*, 1255-1269.

Potter, M. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 509-522.

Schacter, D. L., Reiman, E., Uecker, A., Polster, M. R., Yun, L. S., & Cooper, L. A. (1995). Brain regions associated with retrieval of structurally coherent visual information. *Nature*, *376*, 587-590.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J., Brady, T. J., Rosen, B. R., & Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889-893.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain*, *115*, 15-36.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General*, *109*, 373-392.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174-215.

Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., & Raichle, M. E. (1992). Activation of the hippocampus in normal humans: A functional anatomical study of memory. *Proceedings of the National Academy of Sciences USA*, *89*, 1837-1841.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, R. J. W. Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA: MIT Press.

Watson, J. D. G., Myers, R., Frackowiak, R. S. L., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., & Zeki, S. (1993). Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, *3*, 79-94.

Woods, R. P., Cherry, S. R., & Mazziotta, J. C. (1992). Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography*, *16*, 620-633.

Woods, R. P., Mazziotta, J. C., & Cherry, S. R. (1993a). MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, *17*, 536-546.

Woods, R. P., Mazziotta, J. C., & Cherry, S. R. (1993b). Automated image registration. In K. Uemura, N. A. Lassen, T. Jones, I. Kanno (Eds.), *Quantitation of brain function: Tracer kinetics and image analysis in brain PET* (pp. 391-398). Amsterdam: Excerpta Medica.

Worsley, K. J., Marrett, S., Neelin, P., & Evans, A. C. (1995). A unified statistical approach for determining significant signals in location and scale space images of cerebral activation. *NeuroImage*, *2*, S71.

Zeki, S., Watson, J. D. G., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, S. J. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience*, *11*, 641-649.