Visual Recognition: Evidence for Two Distinctive Mechanisms from a PET Study

Priyantha Herath,^{1*} Shigeo Kinomura,² and Per E. Roland¹

¹ Division of Human Brain Research, Department of Neuroscience, Karolinska Institute, Stockholm, Sweden

² Department of Nuclear Medicine and Radiology, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan

Abstract: In this study, we examined the hypothesis that two distinct sets of cortical areas subserve two dissociable neurophysiological mechanisms of visual recognition. We posited that one such mechanism uses category specific cues extractable from the viewed pattern for the purpose of recognition. The other mechanism matches the pattern to be recognized with a pre-encoded memory representation of the pattern. In order to distinguish the cortical areas active in these two strategies, we measured changes in regional cerebral blood flow (rCBF) with positron emission tomography (PET) and ¹⁵O Butanol as the radiotracer. Ten subjects performed pattern recognition tasks based on three different short-term memory conditions and a condition based on visual categories of the patterns. When subjects used representations of the patterns held in short-term memory for the purpose of recognition, the precune were bilaterally activated. Recognition based on visual categories of the patterns activated the right (R) angular gyrus, left (L) inferior temporal gyrus, and L superior parieto-occipital cortex. These findings demonstrate that the R angular gyrus, the L inferior temporal gyrus, and the L superior parieto-occipital cortex are associated with recognition of patterns based on visual categories, whereas recognition of patterns using memory representations is associated with the activity of the precune. This study is the first to show functional dual dissociation of active cortical fields for different mechanisms of visual pattern recognition. *Hum. Brain Mapping* 12:110–119, 2001. © 2001 Wiley-Liss, Inc.

Key words: cortical fields; dual dissociation; rCBF; inferior temporal gyrus; visual categories; short term memory; and pattern recognition; precuneus

INTRODUCTION

Visual pattern recognition has been the subject of extensive empirical studies. The mechanisms and the functional anatomy of visual pattern recognition has been investigated in brain damaged patients [Farah, 1990; Grüsser and Landis, 1991], experimental lesions

and single-unit recordings in primates [Logothetis and Sheinberg, 1996; Tanaka 1997], optical imaging of intrinsic signals [Wang et al., 1996], and in humans using the noninvasive functional imaging methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) [Roland and Gulyás, 1995; Kanwisher et al., 1996; Bly and Kosslyn, 1997; Tanaka, 1997]. These studies and many others have accumulated a large body of knowledge about visual pattern recognition. However, except for the evidence for the importance of the inferior temporal gyrus in pattern recognition and a possible role of posterior parietal areas [Faillenot et al., 1997], no other

Contract grant sponsor: The VOLVO Foundation of Sweden.

*Correspondence to: Priyantha Herath, Division of Human Brain Research, Dept. of Neuroscience, Karolinska Institute, Stockholm 171 77, Sweden. E-mail: Priyantha.Herath@neuro.ki.se

Received for publication 17 December 1999; accepted 5 October 2000

consistent evidence for dissociated cortical areas engaged in different mechanisms of visual recognition is available yet.

We investigated two possible computational strategies and their neuro-anatomical substrates. Visual recognition can be based on category specific cues extractable from the patterns. For instance, humans can recognize the color blue as being blue. In this case, the semantic label “blue” is used to specify the visual category. It is also possible to recognize patterns simply because they have been seen before. In this instance, the recognition is based on representations of that particular pattern stored in the memory. In the visual category based recognition we expected to see increased cortical activity in areas that act as the interface between visual image representations and their semantic representations or, more simply, the areas that may be involved in processing meanings of viewed objects and patterns. In memory based pattern recognition, we expected increased rCBF in cortical areas that are active in retention and recognition of visual short-term memories. As the search for “blue” or “quadrangle” may be viewed as finding targets among distracters (i.e., other colors and geometric figures in the patterns viewed), we attempted to partition the brain activity of memory driven recognition by using visual recognition tasks with and without different types of distracters embedded within the patterns (i.e., color and long-term memory distracters).

MATERIALS AND METHODS

Subjects

Ten normal healthy male volunteers (age 21–37 year) who had no significant past medical, neurological, or psychiatric histories participated in the study. All subjects gave informed written consent for the scanning experiment. The local ethics committee of the Karolinska hospital, Stockholm, approved the study. Nine subjects were right handed and one was left handed according to a Swedish version of the Edinburgh Handedness Inventory [Oldfield, 1971]. All had normal or corrected to normal visual acuity, and none had visual field defects. Before the scanning, the subjects were naïve about the psychophysical experiment.

Electrophysiological measurements

All subjects had electroencephalogram (EEG) and bilateral electro-oculograms (EOGs) recorded during the scanning procedure. The EEG and EOGs were recorded with a Siemens Mingograph, at 100× ampli-

fication. For EEG recordings, needle electrodes were placed in the scalp at F3, F4, P3, P4, and Cz according to the international 10–20 system. The ground electrode was attached to the skin over the left clavicle. The recordings were analyzed off line for the proportions of different frequency domains. The EOGs were measured with two silver electrodes attached to the skin over the lateral epicanthi of both eyes. With this, we measured the number of saccadic eye movements equal or greater than 2 degrees in amplitude. The measurement for movements less than 2 degrees is imprecise with this method.

Data acquisition

Each subject rested comfortably in the supine position and had their heads fixed to the scanner bed by a stereotaxic helmet [Bergström et al., 1981] that stabilized the head in identical position during the anatomical MR and functional PET scanning procedures. First, a high-resolution anatomical magnetic resonance (MR) image of the brain was obtained on a GE Signa 1.5 T scanner, with a 3D SPGR sequence (TE = 5 ms, TR = 21 ms, flip angle 50 degrees, NEX = 1, FOV = 25.6 cm) yielding a 3D brain volume with 1 mm isotropic voxels. These MR images were used to guide the head positioning in the PET scanner.

The method of measuring rCBF with PET and ^{15}O Butanol has been described in detail [Roland et al., 1993], and only a brief description is given here. The PET measurements were made in a Scandtronix 2048-15B PET camera. This PET camera had an in-plane resolution of 4.5 mm and an interslice distance of 6.5-mm and a maximum axial field of view of 10.5 cm. With this, we were able to acquire a brain volume between $z = -24$ and $z = +61$ in all subjects. As such, our imaging data included the brain activity of almost the whole brain except for a small part of the premotor cortex, the lowermost part of the temporal poles, and the cerebellum. Subjects received 70 mCi of ^{15}O Butanol as a bolus injection at the beginning of each PET scan. The bolus was administered through an intravenous cannula placed in the left cubital fossa before the scanning procedures. Four injections of ^{15}O Butanol corresponding to four tasks were given to each subject. The sinograms were reconstructed using a 4-mm Hanning filter. The reconstructed imaging data were corrected for radioactive attenuation. For each condition, the count rates of frame 1–10 representing 0–50 sec from the beginning of the injection were integrated. The rCBF was calculated by an autoradiographic method [Meyer, 1987].

Stimuli

The background illumination during all scanning sessions was 0.27 cd/m². The subjects were adapted by a 10-min exposure to this level before the PET scanning sessions began. All stimuli (e.g., Fig. 1) that were used in the memory sets and subsequent displays during the scanning sessions have been well characterized previously [Roland and Gulyás, 1995]. The stimuli had an internal color contrast of 0.32. The average luminance, information content in bits and the signal energy were balanced across all items. The spatial frequency within each pattern varied from one to a few cycles per degree [Roland and Gulyás, 1995]. The computer-generated patterns were drawn from a pool of 400 such items. Therefore, the patterns were not repeated across the conditions. However, some patterns were repeated within the sessions (see below). These patterns were projected onto a screen in such a way that they subtended a 15 degree × 20 degree (vertical × horizontal) incidence angle. Subjects viewed the patterns binocularly. An electronic shutter driven via a computer controlled the exposure times. The order of task presentation was randomized according to a balanced schedule across subjects, and they received stimuli for each condition at the same rate of presentation. Thus, during the scanning period the rate of stimuli presentation for each condition was 33 stimuli/50-sec acquisition period. In all conditions, subjects received 11–12 targets and 21–22 distracters within a 50-sec long scanning session. The proportion of target: nontarget pattern was 1:2. Each stimulus was exposed for 1,000 ms and stimuli were separated by 500 ms interstimulus interval.

Activation tasks

Object recognition by visual category specific cues (CatRec)

Geometric patterns [Roland and Gulyás, 1995] (e.g., Fig. 1) were presented at the rate of 1 pattern/sec plus 500 ms intertrial intervals. The task was to “recognize those patterns that had either a blue colored part or a quadrangular part somewhere in it and respond as fast as possible” by pressing a response key with their right thumb to confirm the recognition of the patterns containing the targets. The concept of quadrangles was further defined to the subjects as any figure in the patterns containing four corners or angles. The proportion of target: nontarget pattern was 1:2.

Figure 1.

Geometrical patterns used in the experiment. Note that some patterns contain “blue” or “quadrangular” parts while others do not.

Recognition based on short-term memory (STM)

A set of five geometrical patterns (e.g., Fig. 1) was presented for 8 sec each with 1-sec interval in between the stimuli. These sets of stimuli were called the memory set. Subjects were instructed to retain these patterns in their memory. Immediately after the 45-sec exposure to the target stimuli, the PET scanning started. During the scanning, additional geometrical patterns, randomly intermixed with the items from the memory set were presented at a rate of 1 pattern/sec with 500 ms intervals. The task was to press a response key as quickly as possible when a pattern from the memory set was recognized from the items that appeared on the display. The proportion between old (memory set): a new pattern was $\approx 1:2$ and target stimuli were thus represented twice during the scanning.

Short-term memory with color distracters (STMCOLDISTR)

As in STM, a set of five geometrical patterns was presented for 8 sec each with 1-sec intervals. When PET scanning started following this, the geometrical patterns belonging to the memory set was presented intermixed with patterns that had identical geometry but different colors (ratio of memory set presentations: distracters $\approx 1:2$), target stimuli repeated twice. The task was to press the response key when an item from the memory set was recognized from the items that appeared in the display.

Working memory with long-term memory distracters (STMLTMDISTR)

The subjects learned a set of 15 geometrical patterns while lying in the scanner bed. These were presented as three sets of five patterns each (LTM memory set). Each stimulus was shown for 8 sec each with 1-sec intervals in between. The pattern sets were shown for approximately 45 min, during which time 100 memory sets (of five patterns each) were presented. Following this learning period, five new geometrical patterns (STM memory set), each for a period of 8 sec followed by 1-sec interval were presented. The PET scanning then began. During the scanning, items from the short-term memory set just displayed were presented intermixed with items from the long-term memory set (ratio of STM memory set: LTM memory set $\approx 1:2$). The task was to recognize the patterns that were shown immediately before the scanning session.

Data processing

The functional images were corrected for between scan motion using Woods AIR algorithms [Woods, 1992]. The first PET scans for each subject served as the reference image. The anatomical structures of each subject's high-resolution brain MR image were fitted interactively to the structures of the Human Brain Atlas [Roland et al., 1994] using a number of spatial transformations. The adaptation parameters thus obtained were then used to reformat the functional PET images to standard atlas brain anatomy [Roland et al., 1994] with the Talairach and Tournoux [1988] coordinate system. The reformatted functional data were then filtered with a 5-mm FWHM Gaussian filter that yielded a final FWHM of 6 mm for the images. The anatomically standardized functional images had a voxel size of $2 \times 2 \times 2 \text{ mm}^3$.

The smoothed images were then modeled for statistically significant activations. We used a General Linear Model [Ledberg, 1998], in which the design matrix included subjects (10) and conditions (4) as factors and global counts as a covariate. As such, the rCBF was normalized in the statistical treatment of data. Statistical t-maps were calculated for contrasts of interest between different activation tasks. The t-maps were converted to Z maps. Both the Z threshold and the cluster size for significant activations for the whole brain ($P < 0.05$, omnibus) was determined by 5,000 Monte-Carlo simulations [Ledberg, 1998]. This simulation process provided the following Z threshold for the whole brain and the corresponding cluster size: $Z = 2.48$. Cluster size = 123 voxels with a volume of 984 mm^3 .

Clusters of voxels that were statistically significant are displayed by coregistering them to the standard anatomical MR image of the Human Brain Atlas [Roland et al., 1994].

All responses were analyzed for response times (RT) and sorted as correct hits ([response]|target), misses ([no response]|target) and false alarms ([response]|no target). Performance was calculated as the area under the curve of the receiver-operating characteristic (ROC) [Gescheider, 1997]. The ROC curve is a graphical representation of how the hit rate of an observer changes as a function of changes in the false alarm rate.

RESULTS

Psychophysical performance

Analysis of response times during different task conditions showed that all recognition tasks based on

TABLE I. Electrophysiological and response measurements during the brain activation conditions

	STM	STMCOLDISTR	STMLTMDISTR	CatRec
% α blockade in EEG mean $\pm$ SD	88 $\pm$ 9.8	90 $\pm$ 9.5	86 $\pm$ 15.2	85 $\pm$ 13.2
Eye movements (Hz) mean $\pm$ SD	0.75 $\pm$.10	0.72 $\pm$ 0.16	0.77 $\pm$ 0.13	0.74 $\pm$ 0.11
Response latency/ms	635 $\pm$ 63	621 $\pm$ 54	594 $\pm$ 52	554 $\pm$ 99
Mean $\pm$ SD				
Performance: the area under the ROC curve	0.87 $\pm$ 0.09	0.85 $\pm$ 0.07	0.91 $\pm$ 0.08	0.81 $\pm$ 0.04
Mean $\pm$ SD				

STM had longer response times as compared to the visual category recognition task (One way ANOVA, $P < 0.05$, $F_{(3,36)} = 9.23$). Neither the response times between the three different STM recognition tasks nor the performance of the four activation conditions showed any difference at a statistically significant level (one-way ANOVA, $P > 0.05$) in all cases. For a summary of psychophysical results, see Table I.

Electrophysiological measures

All four activation tasks had about the same α -blockade (measured as % α -blockade within each condition) in all leads in EEG. Analysis of the EOG indicated that the number of saccadic eye movements was not different between tasks (Table I). We were unable to see any differences in the amplitudes of the eye movements across conditions.

PET activations

Contrasted to the visual category recognition task, the short-term memory tasks (STM + STMCOLDISTR + STMLTMDISTR – 3CatRec), [i.e., 1 1 1 –3], activated a large cortical field in the precuneus. This contrast emphasizes the main effect of STM based recognition processing compared to the category based processing. The active cortical field had a volume of 1,358 mm³. The center of mass was in the L precuneus, above the subparietal sulcus (–3.6, –57.6, 23.8). This volume of gray matter extended in +10 to +16, –44 to –60, and +11 to +35 in the x, y, and z planes respectively. Laterally, the active cortical field did not reach as far as the superior parietal lobule. The active cortical field was coextensive, across the midline, with the R precuneus. It had a “dumbbell” shape across the midline with the most of the active mass almost symmetrically distributed in the precuneus on both sides. This produced the only cortical area that is active in the visual short-term memory based recognition tasks (Fig. 2A, Table II).

When the visual category recognition task was contrasted with the visual short-term memory tasks (3CatRec – (STM + STMCOLDISTR + STMLTMDISTR)) [i.e. 3 –1 –1 –1], a set of cortical fields located in the R angular gyrus, L inferior temporal gyrus, and L superior parieto-occipital cortex was activated. (Fig. 3a, Table II). The center of mass of the L inferotemporal cortical field was (41.3, –70, –6.6) and its volume was 1,000 mm³. It lay quite laterally to the fusiform gyrus and confined to the inferior temporal gyrus. As such, it extended between +34 to +51, –62 to –80, and –15 to +39 in the x, y, and z planes. The R angular gyrus cortical field had a volume of 1,056 mm³, with its center of mass at –42.2, –43.9, and –43.4. Although this activity spread across the intraparietal sulcus to the superior parietal lobule, the bulk of the activity and therefore the center of the mass was on the angular gyrus. The largest area of cortical field activity was in the L superior parieto-occipital cortex, with its center of mass at (21.2, –71.9, 43.2) and a gray matter volume of 1,710 mm³. These activations were almost identically replicated when the two STM conditions with distracters were compared to the visual category recognition condition CatRec – (STMCOLDISTR + STMLTMDISTR) (Fig. 3b, Table II). Once again, the cortical activity was confined to the L inferior temporal gyrus (2,176 mm³), two discrete clusters in R angular gyrus (1,224 mm³ + 1360 mm³), and the largest cortical field in L superior parieto occipital cortex (1,496 mm³). The extent of the clusters in x, y, z planes corresponded to what was seen in the contrast for (3CatRec – (STM + STMCOLDISTR + STMLTMDISTR)).

When the recognition task with long-term memory distracters was compared to the visual category based recognition (STMLTMDISTR – CatRec) (i.e., 0 0 1 –1), there was bilateral activity of the precuneus, with a cortical field that had its center of mass at –3.9, –57.1, 32.0 and a volume of 1,864 mm³ (Fig. 2B, Table II). It appeared that most of this mass was right lateralized. However, the active cortical field was coexten-

Figure 2.

A. Activation of the precuneus when the short term memory recognition tasks are contrasted to the visual category based recognition task (sections are displayed at $x = +4$, $y = +21$, and $z = -59$). **B.** Activation of the precuneus when short-term memory with distractors task contrasted to the visual category recognition task (sections are displayed at $x = -4$, $y = +32$, and $z = -51$).

sive across the midline with the left precuneus, resembling the active cluster of $(STM + STMCOLDISTR + STMLTMDISTR) - 3CatRec$. However, this cluster extended more anteriorly so that it almost reached the posterior parts of the cingulate gyrus.

The contrasts between $STMLTMDISTR - STM$, $STMLTMDISTR - STMCOLDISTR$, and $STM - CatRec$

did not show any significant increase in cortical activity.

We performed a volume of interest (VOI) based analysis for comparison of STM recognition tasks with visual category recognition tasks, taking the active precuneal cortical field as the volume of interest. This showed highly significant activity in $[STMCOLDISTR + STMLTMDISTR - STM - CatRec]$, $[STMCOLDISTR + STMLTMDISTR - CatRec]$, $[STMLTMDISTR - CatRec]$, $[STMCOLDISTR - CatRec]$, and in $[STM - CatRec]$ ($P < 0.001$, Bonferroni corrected for multiple comparisons).

DISCUSSION

We have psychophysically demonstrated that visual patterns can be recognized by (a) using visual category specific information extractable from the pattern and (b) matching a perceptual description of the pattern to a representation of that pattern previously stored in memory. Using PET and ^{15}O Butanol, we have then shown that the visual category based recognition activated the R angular gyrus, L superior parieto-occipital cortex, and the L inferior temporal gyrus, while the short-term memory-based recognition strategy activated the precuneus bilaterally.

Category based recognition > STM based recognition
 $\in$

R angular gyrus, L superior
 parieto-occipital cortex and L inferior
 temporal gyrus

STM based recognition > category based recognition
 $\in$ L and R precuneus

Essentially, this is clear evidence that shows the presence of a dual dissociation of activation of cortical fields between visual category cued and short-term memory based pattern recognition. The present study, to our knowledge, is the first to show such a dual dissociation for visual pattern recognition.

We observed no significant differences in the detection of targets in the four activation tasks. Especially, the probability of detecting targets between the two STM recognition tasks with distractors ($STMCOLDISTR$ and $STMLTMDISTR$) was similar. This indicates that the level of difficulty between tasks was well balanced. We also noted that there were no differences in the number of eye movements in the four conditions. Keeping in mind that our stimuli are large field projections that covered approximately 15×20 degrees of the visual field, the subjects had to move their eyes in order to see the whole pattern. While the

TABLE II. Location of the maxima of cortical fields activated in different contrasts between activation tasks

Contrast	Cluster size/mm ³	x	y	z	Mean t value	Anatomical region
STM + STMCOLDIST + STMLTMDIST – 3CatRec	1358	–3.6	–57.6	23.8	2.94	R-precuneus
3CatRec – STM + STMCOLDISTR + STMLTMDISTR	1000	41.3	–70	–6.6	3.09	L-inferior temporal gyrus
	1056	–42.2	–43.9	43.3	3.21	R-angular gyrus
	1710	21.2	–71.9	43.2	3.04	L-superior parieto-occipital cortex
2CatRec – STMCOLDISTR + STMLTMDISTR	1176	43	–69	–7.0	3.19	L-inferior temporal gyrus
	1224	–41.6	–43.1	43.3	3.23	R-angular gyrus
	1360	–29.4	–62.2	46.8	3.09	R-angular gyrus
	1496	20.9	–69.1	43.7	3.12	L-superior parieto-occipital cortex
STMLTMDISTR – CatRec	1864	–3.9	–57.1	32.0	2.96	R-precuneus

subjects were doing this, the amplitudes of the eye movements did not seem to differ. This indicated that the visual search in all tasks was roughly similar. Therefore, it is unlikely that the subjects used different search strategies (i.e., global search and local search) in the visual category task and the short-term memory tasks. The degree of α blockade was similar between the tasks. This suggests that the attentional demands in the four tasks were closely matched. The only behavioral difference that we noticed was that the three STM tasks had significantly longer response times than the CatRec task. This is an important finding because the visual patterns of the memory sets were repeated twice during the three STM tasks while the patterns of the CatRec were not repeated. It could be argued that the STM tasks were confounded by possible visual priming effects. The longer RTs in STM tasks makes this unlikely. The study can be construed to be limited in one aspect. We acquired the images in the 2D mode, and this limited the number of repetitions. Further, we were unable to measure the rCBF in the dorsal most part of the premotor cortex, the tips of the temporal poles, and the posterior lobes of cerebellum in all subjects. Although we cannot totally exclude activations in some subjects in these regions, this would be of no importance for the visual activations found. The search for “blue” or “quadrangle” may be viewed as finding targets among distracters (i.e., other colors and geometric figures in the patterns viewed). This is why the visual category task was matched against distracters in STMCOLDISTR task. Apparently, whether the distracters were present or not did not affect the contrast between visual category recog-

nition task and the short-term memory driven recognition.

In pattern recognition with visual category, subjects had to recognize the patterns by associating viewed features of the objects with semantically definable visual categories such as “blue” or a “quadrangle.” Recognition at this level depends on having previously established categories of “What is blue?” and “What is a quadrangle?” The terms “blue” and “quadrangle” contain semantic reference to the two visual categories. Therefore, we regard CatRec task as a top-down recognition process. For the task in which recognition was to be carried out based on visual (short-term) memories, the patterns themselves were meaningless. The only way to solve the task was to match a pattern to a stored representation. Presumably, the patterns being viewed need to be processed to some extent starting from their elementary cues such as spectral composition, luminance, edges, and so on before this matching process can take place. Thus, we regard the STM tasks as a bottom-up type of recognition. Notwithstanding these formal descriptions of the requirements to solve the tasks, it should be emphasized that we had no objective means to decide which strategies individual subjects might have used to solve the tasks. We have only calculated and localized the statistically significant activations in the group of subjects.

The cells of the monkey inferior temporal gyrus (ITC) have been implicated in visual pattern recognition based on converging experimental evidence from many studies. For instance, lesions in the IT cortex produce severe defects in an animal’s ability to recognize objects [Ungerleider and Mishkin, 1982]. From

Figure 3.

A. Activation of R angular gyrus, L superior parieto-occipital cortex and the L infero temporal cortex in the visual category based recognition tasks (sections are displayed at $z = +45$ and $z = -10$). **B.** Activation of the identical cortical fields as in Fig 3a when category recognition is contrasted to the short term memory tasks with distracters (sections are displayed at $z = +46$ and $z = -10$). Note: the ratio of target to distracters in both these tasks was 1:2.

single cell recordings it has been shown that cells in the IT cortex show specific activity to complex visual stimuli [Bruce et al., 1981; Young and Yamane, 1992; Miyashita and Cheng, 1988; Tanaka et al., 1990; Fujita et al., 1992; Miyashita 1993]. This suggests that the primate ITC has the physiological machinery needed to recognize patterns.

Many reports from neuroimaging literature implicate human IT cortex and occipito-temporal cortex in pattern and object recognition tasks, albeit the functional homologues between monkeys and humans in these areas are not clear [Roland et al., 1990; Roland and Gulyás, 1995; Haxby, 1991; Kanwisher, 1996; Price et al., 1996; Gauthier et al., 1997; Faillenot et al., 1997; Kraut et al., 1997; Ishai et al., 1999]. The activity of the

IT cortex observed in the present study conforms to the findings from experiments on primates. Specifically, however, in the paradigm we employed, the computational requirement extended beyond simple hierarchical processing of primary visual cues that leads to a complex representation of features in the IT cortex. It was expected that the subjects would rely on the visual categorical cues of the target for the recognition. Clearly, before visual categories can be employed for the purpose of recognition, features of the target patterns have to be processed in the usual manner, viz., using both dorsal and ventral streams of visual processing pathways. However, since this presumably happens in all four task conditions, we do not expect to see cortical activity specific to simple visual form and color processing in the contrast CatRec - (STM + STMCOLDISTR + STMLTMDISTR). Rather, we expected to see cortical activity that represents usage of visual categories. The fact that IT cortex is active in this situation suggests that cells in the human IT cortex are capable of computations that extend beyond complex representations of the viewed patterns. In fact, there is recent evidence for differences of processing in the identification of objects from different semantic classes [Perani et al., 1995; Schacter et al., 1995; Damasio et al., 1996; Martin et al., 1996]. In these studies, the ITC has been implicated as an important cortical area processing semantic category specific information. These studies also explored the relationship between visual identification and access to semantic properties. In particular, Damasio et al. [1996] have shown that the activity in the IT cortex in object recognition contains an intermediate step between prelexical processing and the lexicon.

The angular gyrus and superior parieto-occipital cortex are less well characterized in their contribution to visual information processing. However, there is experimental evidence for presence of effective cross talk between the dorsal (i.e., occipito-parietal) and ventral visual information processing streams [Young, 1992, 1993a,b]. Additionally, there is evidence from human functional imaging studies to show coactivations of dorsal and ventral visual processing pathways in processing of object shape [MacIntosh, 1994; Kraut 1997]. The superior parieto-occipital cortex and the angular gyrus are known to be active in visual stimulation with the geometrical patterns almost identical to the present ones during recall or retrieval tasks of geometric patterns [Roland et al., 1990; Roland and Gulyás, 1995].

In the tasks where subjects were expected to recognize objects using short-term memory representations, precuneus was the only specific area of cortical activ-

ity. As was described earlier, the volume of activity was large and extended across the midline so that both precuneal areas were active in tasks that relied on short-term memory representation of the patterns. The instructions for the task specified that the only strategy to be used to solve the problem was to remember the patterns. The patterns themselves had no semantic content. The presence of a visual recognition system reliant on different types of stored representations has already been proposed [Young, 1994; Ullmann, 1995; Kanwisher et al., 1996; Logothetis and Sheinberg, 1996]. Out of the cortical areas that are implicated in retrieval of visual memories, ITC is probably one of the most important, given its consistent activity across many different visual memory retrieval tasks [Miyashita and Cheng, 1988; Miller, 1991; Fujita, 1992; Miyashita, 1993; Miller, 1994; Young, 1994; Moscovitch, 1995; Logothetis and Sheinberg, 1996; Tanaka, 1996].

In the recent years, however, in addition to the traditional interest in the IT cortex, other cortical areas have emerged as important structures in the visual memory processes. One such area is the precuneus. There are a number of reports that showed increased rCBF in precuneal areas in recall of geometrical patterns and familiar surroundings [Roland et al., 1987, 1990; Roland and Gulyás, 1995]. Further, use of visual imagery during a word recall task (i.e., memory-related imagery) was shown to be associated with significant activation of the precuneus [Fletcher et al., 1995; Halsband et al., 1998]. It has been concluded that the precuneus is a key area of the neural substrate of visual imagery used in conscious recall. However, the precuneus is also active during other imagery tasks and visual tasks without any memory components [Mellet et al., 1995, 1996; Gulyás and Roland, 1991, 1994]. In the present study, there was increased rCBF in the precuneus when memories were used to recognize patterns. The most parsimonious conclusion from the paradigm that we employed is that the neurons in the precuneal areas are involved in some form of visual memory processing, particularly important in matching of targets to templates. The cue invariant nature of the recognition process is perhaps due to this possibility to match targets to many templates or to select from a group of visual categories. The significantly longer response times for all three STM recognition tasks suggest that the matching of processed complex features of a pattern to a memory template requires more processing steps than searching for a match of semantic category for the same. This looks counterintuitive because one assumes that it would be easier to match a target to

templates than to analyze and sort through semantic content of a target.

CONCLUSION

We conclude that the cortical areas involved in visual recognition comprise at least two different sets of extrastriate cortical visual areas. These visual areas are involved in different modes of visual recognition. One such mode of visual recognition can be the use of short-term memories. The other mode appears to be the usage of visual categories as cues for the purpose of recognition. The results of this study add to a rapidly growing repertoire of the neurophysiology of visual recognition.

REFERENCES

- Bergström M, Boets J, Eriksson L, Greitz T, Widen L (1981): Head fixation device for reproducible positron alignment in transmission CT and positron emission tomography. *J Comput Assist Tomogr* 5:130–141.
- Bly BM, Kosslyn SM (1997): Functional anatomy of object recognition in humans: evidence from positron emission tomography and functional magnetic resonance imaging. *Curr Opin Neurol* 10:5–9.
- Bruce C, Desimone R, Gross CG (1981): Visual properties of neurons in a polysensory area in superior temporal sulcus in the macaque. *J Neurophysiol* 46:369–384.
- Damasio H, Grabowski TJ, Tranel D, Hicwa RD, Damasio AR (1996): A neural basis for lexical retrieval. *Nature* 380:499–505.
- Faillenot I, Toni I, Decety J, Gregoire MC, Jeannord M (1997): Visual pathways for object oriented action and object recognition. Functional anatomy with PET. *Cereb Cortex* 7:77–85.
- Farah MJ (1990): Visual agnosia: disorders of object recognition and what they tell us about normal vision. Cambridge, MA: MIT Press.
- Fletcher PC, Frith CD, Baker SC, Shallice T, Frackowiak RS, Dolan RJ (1995): The mind's eye precuneus activation in memory related imagery. *Neuroimage* 2:195–200.
- Fujita I, Tanaka K, Ito M, Cheng K (1992): Columns for visual features of objects in monkey inferior temporal gyrus. *Nature* 360:343–346.
- Gauthier I, Anderson AW, Tarr MJ, Skudlarski P, Gore JC (1997): Levels of categorizations in visual objects studied with functional MRI. *Curr Biol* 7:645–651.
- Gescheider GA (1997): *Psychophysics: the fundamentals* (3rd ed.). Hillsdale, NJ: Erlbaum.
- Grüsser OJ, Landis T. (eds) (1991): *Visual agnosias and other disturbances of visual perception and cognition*. London: Macmillan.
- Gulyás B, Roland PE (1991): Cortical fields participating in form and color discrimination in the human brain. *Neuroreport* 2:585–588.
- Gulyás B, Roland PE (1994): Processing and analysis of form, color and binocular disparity in the human brain: functional anatomy by positron emission tomography. *Eur J Neurosci* 6:1811–1828.
- Halsband U, Krause BJ, Schmidt D, Herzog H, Tellmann L, Muller-Gartner HW (1998): Encoding and retrieval in declarative learn-

- ing: a positron emission tomography study. *Behav Brain Res* 97:69–78.
- Haxby JV, Grady CL, Horowitz B, Ungerleider LG, Mishkin M, Carsson RE, Herscovitch P, Schapiro MB, Rapoport SI (1991): Dissociation of object and spatial visual processing pathways in human extra striate cortex. *PNAS* 88:1621–1625.
- Ishai A, Ungerleider LG, Martin A, Schouten JL, Haxby JV (1999): Distributed representation of objects in the human ventral visual pathway. *PNAS* 96:9379–9384.
- Kanwisher N, Chun MM, McDermott J, Ledden PJ (1996): Functional imaging of human visual recognition. *Cogn Brain Res* 5:55–67.
- Kraut M, Hart J Jr, Soher BJ, Gordon B (1997): Object shape processing in the visual system evaluation using functional MRI. *Neurology* 48:1410–1420.
- Ledberg A, Åkerman S, Roland PE (1998): Estimation of the probabilities of 3D clusters in functional brain images. *Neuroimage* 8:113–128.
- Logothetis NK, Sheinberg DL (1996): Visual object recognition. *Ann Rev Neurosci* 19:577–621.
- Martin A, Wiggs CL, Ungerleider LG, Haxby JV (1996): Neural correlates of category specific knowledge. *Nature* 379:649–652.
- McIntosh AR, Grady CL, Ungerleider LG, Haxby JV, Rapoport SI, Horowitz B (1994): Network analysis of cortical visual pathways mapped with PET. *J Neurosci* 14: 655–666.
- Mellet E, Tzourio N, Denis M, Mazoyer B (1995): A positron emission tomography study of visual and mental spatial exploration. *J Cogn Neurosci* 7:433–445.
- Mellet E, Tzourio N, Crivello F, Joliot M, Denis M, Mazoyer B (1996): Functional anatomy of spatial mental imagery generated from verbal instructions. *J Neurosci* 16:6504–6512.
- Meyer E (1987): Simultaneous correction for tracer arrival delay and dispersion in CBF: measurement by $H_2^{15}O$ autoradiographic method and dynamic PET. *J Nucl Med* 30:1069–1078.
- Miller EK, Lin Li, Desimone R (1991): A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* 254:1337–1379.
- Miller K, Desimone R (1994): Parallel neuronal mechanisms for short term memory. *Science* 263:520–522.
- Miyashita Y, Cheng HS (1988): Neuronal correlate of pictorial short term memory in the primate temporal cortex. *Nature* 331:68–70.
- Miyashita Y (1993): Inferior temporal cortex: where visual perception meets memory. *Ann Rev Neurosci* 16:245–263.
- Moscovitch M, Kapur S, Köhler S, Houle S (1995): Distinct neural correlates of visual long term memory for spatial location and object identity: a positron emission tomography study in humans *PNAS* 92:3721–3725.
- Oldfield RC (1971): The assessment and analysis of handedness. The Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Perani D, Kappa SF, Bettinardi V, Bressi S, Gorno-Tempini M, Matarrese M, Fazio F (1995): Different neural systems for the recognition of animals and man made tools. *Neuroreport* 6:1637–1641.
- Price CJ, Moore CJ, Humphreys GW, Frackowiak RS, Friston KJ (1996): The neural regions sustaining object recognition and naming. *Proc R Soc Lond B Biol Sci* 223:1501–1507.
- Roland PE, Eriksson L, Widen L, Stone-Elander S (1987): Does mental activity change the oxidative metabolism in the brain? *J Neurosci* 7:2373–2389.
- Roland PE, Gulyás B, Seitz RJ, Stone-Elander S (1990): Functional anatomy of storage, recall and recognition of a visual pattern in man. *Neuroreport* 1:53–56.
- Roland PE, Levin B, Kawashima R, Åkerman S (1993): Three-dimensional analysis of clustered voxels on ^{15}O Butanol brain activation images. *Hum Brain Mapp* 1:3–19.
- Roland PE, Graufelds CJ, Wählin J, Ingelman L, Andersson M, Ledberg A, Pedersen J, Åkerman S, Dabringhus A, Zilles K (1994): Human brain atlas: for high resolution functional and anatomical mapping. *Hum Brain Mapp* 1:173–184.
- Roland PE, Gulyás B (1995): Visual memory, visual imagery, and visual recognition of large field patterns by the human brain: functional anatomy by positron emission tomography. *Cereb Cortex* 1:79–93.
- Schacter DL, Reiman E, Uecker A, Porster MR, Yun LS, Cooper LA (1995): Brain regions associated with retrieval of structurally coherent visual information. *Nature* 376:587–590.
- Talairach J, Tournoux P (1988): Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers.
- Tanaka K, Saito H, Fukuda Y, Morioka M (1990): Integration of form, texture and color information in the inferotemporal cortex of the macaque. In: Iwai E, Misitkin M, editors. *Vision memory and the temporal lobe*. New York: Elsevier. p 101–109.
- Tanaka K (1996): Neuronal mechanism of object vision. *Ann Rev Neurosci* 19:109–139.
- Tanaka K (1997): Mechanisms of visual object recognition: monkey and human studies. *Curr Op Neurobiol* 7:523–529.
- Ullmann S (1995): The visual analysis of shape and form. In: Gazzaniga MS, editor. *The cognitive neurosciences*. Cambridge, MA: MIT Press. p 339–350.
- Ungerleider LG, Mishkin M (1982): Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW, editors. *Analysis of visual behavior*. Cambridge, MA: MIT Press. p 549–586.
- Wang G, Tanaka K, Tanifuji M (1996): Optical imaging of functional organization in the monkey inferior temporal gyrus. *Science* 272:1665–1668.
- Woods RP, Cherry SR, Mazziotta JC (1992): Rapid automated algorithm for aligning and reslicing PET images. *J Comput Assist Tomogr* 16:620–633.
- Young MP (1992): Optimization analysis of the organization of cortico-cortical connections between areas of monkey cerebral cortex. *Soc Neurosci Abstr* 18:313.9.
- Young MP (1993a): Visual cortex: module for pattern recognition. *Curr Biol* 3:44–46.
- Young MP (1993b): The organization of neural systems in primate cerebral cortex. *Proc R Soc Biol Sci* 252:13–18.
- Young MP (1994): Open questions about the neural mechanisms of visual pattern recognition. In: Gazzaniga M, editor. *The cognitive neurosciences*. Cambridge, MA: MIT Press. p 463–474.
- Young MP, Yamane S (1992): An analysis at the population level of the processing of faces in the IT cortex. In: Ono T, Squire LR, Petrett DI, Fukuda M, editors. *Brain mechanisms of perception and memory: from neuron to behavior*. Oxford: Oxford University Press. p 47–71.