



High level object recognition without an anterior inferior temporal lobe

IRVING BIEDERMAN,*¶ PETER C. GERHARDSTEIN,† ERIC E. COOPER‡
and CHARLES A. NELSON§

*Department of Psychology, University of Southern California, Hedco Neuroscience Building, Los Angeles, CA 90089-2520, U.S.A.;

†Department of Psychology, Rutgers University, New Brunswick, NJ 08903, U.S.A.; ‡Department of Psychology, Iowa State University, 112 Lagomarcino Hall, Ames, IA 50011, U.S.A.; and §Institute of Child Development, University of Minnesota, 51 East River Road, Minneapolis, MN 55455, U.S.A.

(Received 16 April 1995; accepted 26 June 1996)

Abstract—Seven individuals with unilateral anterior inferior temporal (AIT) lobectomies performed two types of shape recognition tasks with line drawings of 3D objects briefly presented in either the left or the right visual field. In one task, subjects named familiar objects in a name priming paradigm. In the other task, subjects judged whether two objects, presented sequentially with an intervening mask, were the same or different in shape, disregarding differences in orientation of up to 60° in depth. They could not use names or basic level concepts to do the matching as the stimuli were either nonsense objects or, if familiar objects, were of same name–different-shaped exemplars on different trials. The disadvantage of presenting an image to the lobectomized hemisphere was negligible in both tasks. Two non-exclusive possibilities are suggested by this result: (a) Object recognition is completed posterior to AIT, likely at the temporal–occipital boundary, with no deleterious retrograde effects on object recognition from the AIT section, or (b) Callosal transfer of object information prior to AIT is completely efficient. These results, along with results of single unit recording and lesion experiments in the monkey, PET and MRI imaging in humans, and a plausibility argument based on the pattern of callosal connections suggest both are correct. Rather than mediating real-time object recognition, AIT may code representations for visual episodes and scenes. © 1997 Elsevier Science Ltd. All rights reserved.

Key Words: anterior inferior temporal cortex; object recognition; extrastriate visual systems; temporal lobectomy; higher level vision.

Introduction

Temporal lobectomy is a procedure in which, minimally, the anterior half of one temporal lobe is removed, typically as a treatment for epilepsy. Although the superior (speech) region is spared if the surgery is performed in the dominant hemisphere, the anterior inferior temporal cortex (AIT), an area shown to be exclusively concerned with vision in the monkey [1], is always removed. In the present investigation, pictures of objects were briefly presented to temporal lobectomized individuals for identification in either the left or the right visual field. Our major interest was in whether performance would be impaired when the images were presented to the lobectomized hemisphere compared to the intact hemisphere.

Whether an impairment is evidenced under these con-

ditions depends on two factors: the participation of AIT in recognition, and whether the information for recognition is confined to the hemisphere that initially received the image. If AIT plays a role in object recognition, and the input information is confined to the hemisphere that initially received the image, then deficits in recognition should be evident when the image is projected to the lobectomized hemisphere. The magnitude of these deficits would depend on the importance of AIT for object recognition. Alternatively, if recognition is accomplished prior to AIT (and the AIT section does not result in retrograde ipsilateral loss of critical tissue), or the transfer of information to the opposite hemisphere is completely efficient (in both rate and quality), then no deficit would be expected. If the transfer was only partially efficient then some deficit would be expected when the image is projected to the lobectomized hemisphere.

In the next two sections we review the role of AIT in object recognition, first in the monkey and then the human. Object recognition is defined as the immediate activation of a representation of the shape of a stimulus,

¶ Address for correspondence: Hedco Neuroscience Bldg, MC 2520, University of Southern California, Los Angeles, CA 90089-2520, U.S.A.; e-mail: ib@rana.usc.edu; tel.: 213-740-6094; fax: 213-740-5687.

sufficient for access to a basic level class [50] if one exists for that stimulus. Following that review, we report the results of two experiments with individuals with unilateral AIT sections, in which we compared object recognition performance for images projected to the hemisphere that did or did not have its AIT. To anticipate the results, we found no effect of AIT removal: Recognition performance was equally good for images projected to the lobectomized and intact hemispheres, in both a same–different matching task of depth-rotated objects and a naming task.

Is the anterior inferotemporal lobe (AIT) required for high level shape recognition in monkeys?

There is widespread acceptance of Ungerleider and Mishkin's [60] proposal that two different streams, the ventral and the dorsal, are involved in the processing of shape. Both streams have their origins in the primary visual projection area, V1, but the ventral pathway, V1 → V2 → V4 → TEO → TE (with the latter two stages in the inferior temporal [IT] cortex), is generally assumed to subservise object recognition. The other, the dorsal stream, extends from V1 to PP (posterior parietal cortex) and has been characterized as mediating spatial representations. Our interest in the present investigation is with the ventral stream.

Bilateral AIT removal in the monkey. Holmes and Gross [28] showed that bilateral removal of IT† results in marked deficits in learning object discriminations when the stimuli are different in shape. These investigators also documented a deficit in the relearning of discriminations learned preoperatively. These are shape-discrimination deficits rather than general deficits in visual learning because Holmes and Gross [29] also showed that IT lesioned animals were equivalent to controls when discriminating stimuli that were of the same shape but differing in size or in large angles, i.e. >60°, in their 2D orientation. (Curiously, the ability to perform an orientation discrimination task was lost in IT monkeys when the stimuli were simple lines or gratings.)

Additional evidence for involvement of IT in object processing derives from the observation of Holmes and Gross [28] that monkeys with bilateral sections of the temporal lobes seem to employ different sets of cues than normal controls, such as whether the lower section of the stimulus is parallel with the edge of the display rather than the actual shape of the stimulus. Gross [22] and Holmes and Gross [28] suggest that the success of these IT monkeys at performing a visual shape discrimination task may thus derive from contributions of other cortical

areas, such as those specifying locations in space. The sectioning of the posterior region of IT, TEO, tends to produce deficits in the discrimination of single objects whereas the anterior region, TE, produces deficits in the performance of several simultaneous discriminations [32].

Single unit recording in the monkey. Consistent with an IT locus of higher level vision is the finding that cells in TE in the monkey show tuned responding to visual shape (e.g. Refs [1], [35], [37], [45], [58] and [59]). These cells, in general, show the position and size invariance that are characteristic of human object recognition performance [5, 6, 18].

Bayliss *et al.* [1] recorded from neurons in a subarea of TE and found that a sizable proportion were not only exclusively visual, but also responsive only to static (rather than moving) objects. Only a small proportion of these cells responded to faces or to oriented gratings. The majority of these cells exhibited tuned responding in 90–150 msec, making them prime candidates for mediating object recognition, in that humans can readily recognize object pictures from exposures of that duration [31]. These regions were just posterior (up to 20 mm) to the sphenoid (the anterior terminus of the superior temporal fissure). If the same typology characterized humans, a temporal lobectomy would likely remove the homologous tissue.

Although early reports suggested that, aside from cells tuned to faces and hands, tuning of IT cells is non-specific in that a broad range of objects could stimulate them [14], more recent reports suggest highly specific tuning to shape, sometimes with a conjoined surface property such as color or shading (e.g. Refs [20] and [58]). Tanaka and his colleagues [59, 20, 35] have documented selectivity of these AIT cells in the anesthetized monkey to object-like features of medium complexity, such as a trihedral angle or a small sphere on top of a larger sphere. These features were determined by gradually stripping away the other features of far more complex objects, such as models of animals or foods, that were initially employed to discover what could strongly excite the cell. The cells are insensitive to simpler stimuli that would convey the major spatial frequency components of the shape or only one of the components. A columnar organization is apparent in these regions of TE, in that cells prefer quite different stimuli in adjacent columns but show differential tuning to metric variations of the feature within a column. An examination of the preferred stimuli suggests that adjacent columns appear to be tuned to stimuli that differ in viewpoint invariant properties (e.g. straight, curved, vertex type) and it is possible that a particular object would generally continue to activate its column(s)—and not others—as it was rotated in depth as long as its major features were present in the image. For example, one cell (K. Tanaka, personal communication, January, 1995) responded to a circle and a variety of horizontal and vertical ellipses, which would be projections of a circle at different orientations in depth. This cell did not respond

† IT ablations in monkeys typically spare the temporal pole, the region anterior to the terminus of the superior temporal sulcus. This region is removed in the temporal lobe surgeries in the present investigation.

to a triangle, square or star. (It would be interesting to determine the extent to which a preferred stimulus would continue to fire a cell as the stimulus was rotated in depth.) An important characteristic of human object recognition performance is the ability to employ viewpoint-invariant differences to recognize objects at arbitrary orientations in depth, despite dramatic variations in the two-dimensional image [7, 8].

It is now quite evident that after intensive training with a set of visual stimuli, TE cells in monkeys show tuned responding to the experimental stimuli. In awake monkeys, Miyashita and his colleagues (e.g. Refs [45], [51] and [52]) reported increased firing rates of TE cells to novel fractal patterns presented for memory (as well as to shape preserving planar rotations and their reflections). Altering the amplitude or phase parameters of the fractal patterns (so they look different in shape) reduced the firing rates. Logothetis *et al.* [37], using objects resembling bent paper clips and ameoids, also demonstrated that some cells responded to a particular orientation of a studied object (but often were invariant over changes in position or size).

Desimone and his associates [36, 42], in a variety of match-to-sample and working memory tasks, have also demonstrated alterations in firing of IT cells to pictures of objects, though in this case more often the familiar or memorized stimulus was associated with a decrease in firing rate, presumably reflecting competitive inhibition from other units self-organized to respond to the stimulus.

The demonstrations in Miyashita, Logothetis and Desimone's laboratories that AIT cells can become tuned to experimentally familiarized stimuli were performed with awake monkeys where recording was done over an interval in which the stimulus was, presumably, actively rehearsed or visualized so it could be discriminated from distractors. Kobatake and Tanaka (described in Ref. [58]) have shown increased responding to experimentally trained stimuli by TE cells in the anesthetized monkey. They trained two monkeys to discriminate 28 'moderately-complex' shapes with a delayed matching-to-sample task in which the monkeys had to select the correct shape from five alternatives. After a year of training, one-third of the cells gave a maximum response to some of the shapes used in the training (a set of animal and plant models that had been used to determine the critical features in naive monkeys). Only 5% of the cells in untrained monkeys responded maximally to these shapes. Unfortunately, the response of the cells were not analyzed further to determine if they were tuned to the same kind of complex stimuli that had been discovered in the naive monkeys.

Over 50% of the cells in the Miyashita and Logothetis experiments were found to be tuned to the experimental stimuli (i.e. fractal patterns, wire objects or ameoids), a rate somewhat higher than the 33% incidence of Tanaka [58]. It is possible that some of the higher rates of tuning in the Miyashita and Logothetis experiments might

reflect their weaker criterion for tuning of *some* responsiveness rather than the criterion of maximal responsiveness specified by Tanaka. Alternatively, the difference might reflect active processing in the awake monkeys performing the Miyashita and Logothetis tasks. In the Logothetis experiments, of those units that could be driven by the experimental stimuli, 11.6% were selective to the presented view of the target object.

The evidence seems clear that area TE in the macaque is tunable to arbitrary stimuli. Tanaka's [60] finding that tuning was present in the anesthetized monkey rules out active maintenance strategies as a cause of at least some of the tuning. The implications of this tuning will be considered in the Discussion.

Is the anterior inferotemporal lobe (AIT) required for high level shape recognition in humans?

Unilateral temporal lobe removal and free viewing. Unilateral temporal lobectomy does not result in noticeable deficits in object recognition tasks either in monkeys [49] or in humans [43, 44]. Individuals with right AIT lobectomies do reveal, however, modest deficits in their recognition of visual stimuli that cannot be easily labeled, such as faces and geometric designs, as well as in the interpretation of anomalous scenes [43, 44]. Damage to the left middle temporal area (Brodmann's area 21) has been implicated in at least two reports of severe deficits in object naming, but not in object recognition [12],[19]. The anterior portions of Brodmann's area 21 were removed by the surgery performed on the subjects in the present investigation.

It should be noted that the evidence for the lack of deficit in basic-level object identification with AIT lobectomies has come from tasks that allowed free viewing of the stimuli in which the individual could examine an object at his or her own pace (e.g. Refs [43] and [44]). By far, this has been the predominant mode of assessing visual deficits in patient populations. Uncontrolled are the loci, number, and duration of fixations of the image. Inspection (eye movement) strategies in the free viewing situations could have allowed the information critical for recognition to be processed by the intact hemisphere, obscuring a possible deficit that would have been manifested if the information was solely projected to the lobectomized hemisphere. It is also the case that measurement of response latency was rarely undertaken in these studies, so any difference in recognition latency between the intact and lobectomized hemispheres would have passed unnoticed. Remarkably, we could find no experiments assessing visual recognition deficits in patient populations with brief, lateralized presentations. In Martha Farah's monograph, *Visual Agnosia* [16], which reviews 99 case reports in the literature on object agnosia, not a single study with brief lateralized presentations is presented, despite the evidence for differential hemispheric effects and the potential that brief lateralized presentations hold

for evaluating the claims that unilateral temporal lobectomy does not affect recognition.

Posterior locus for object recognition in humans? On the basis of free viewing and drawing performance in patient populations (though with appropriate attention to the time requirements), Farah [16] argued that the capacity to represent multiple shapes, such as the letters making up a word or the parts of an object that can be decomposed into parts, depends on the left inferior temporal–occipital region (i.e. the fusiform). The representation of the parts themselves, she maintained, involves the temporal–occipital cortex bilaterally. Perceiving the overall shape of the object from unconventional orientations or lighting conditions is dependent on the posterior, inferior region of the right parietal lobe. These deficits, according to Farah, would thus stem from lesions to areas far posterior to AIT. These areas would thus be spared by temporal lobectomies.

Imagery investigations. Several PET investigations [13, 25, 40, 55] are consistent with Farah's contention that occipital regions, specifically regions near the temporo-occipital junction and those on the ventral medial surface of the occipital–temporal region, are involved in object recognition. Both the Haxby *et al.* and Sergent *et al.* studies indicated that the areas of greatest rCBF during object recognition were in the ventral regions of the occipital lobe. Haxby *et al.* [25] measured PET activity during performance of the Benton Face Recognition Test. Compared to the control condition (a simple sensory-motor task involving the detection of three squares), most of the task-associated activity was observed in the ventral region of the occipital cortex; none was evident in the temporal lobe. However, it is arguable whether the Benton is really a shape recognition test. Rather it is an extremely difficult discrimination test in that the choice faces are highly similar to the target and to each other. Performance is often dependent on discovering a small, distinguishing feature such as the exact shape of an eyebrow, and it may be this requirement that taxes the occipital cortex.

Sergent *et al.* [55] obtained PET imagery during three experimental tasks; an object classification task in which subjects made 'living' vs 'non-living or man-made' categorization judgments, a gender discrimination task, and a face discrimination task (actor vs non-actor) involving recognition of a set of famous faces. They employed three control conditions: fixation of a central point, passive looking at faces, and discrimination of the orientation of horizontal and vertical gratings by pressing buttons. Noteworthy about this study was the employment of MRI for the individual subjects so that rCBF loci could be more precisely localized with respect to cortical structures.

After subtraction of the activity from the gratings task, the object task showed significant residual activity at the left inferior temporal region (Brodmann's area 20), fusiform (37, in its more posterior regions), middle temporal (21) and middle occipital (19) gyri. However, there

was no evidence of increased activity in the anterior temporal cortex.

Consistent with the Haxby *et al.* [25] results was the finding that the gender discrimination task yielded significant differential activity (though primarily on the right side) in the ventral regions of the occipital cortex, in areas 18 and 19. Like the Benton, gender discrimination, in the absence of easy features such as hair, makeup, and jewelry, may require relatively fine discriminations, compared to those typically required for object recognition. Support for this interpretation comes from the finding in the Sergent *et al.* [55] study that, compared to discriminating the orientation of gratings, all the significant differential rCBF associated with the gender discrimination task was posterior to the significant differential rCBF associated with the object classification task.

Even if one accepts the argument that object agnosias are produced by posterior lesions and that object recognition increases rCBF in the occipital cortex, it does not necessarily mean that the shape representations mediating recognition are activated at those regions. Instead, it is possible that the posterior regions are a necessary (and a more energy-demanding) prerequisite to more anterior processing in the temporal lobes.

The more recent investigations of Martin *et al.* [40] and Damasio *et al.* [13] assessed whether different loci would be found in PET activity when viewing or silently naming man-made objects and animals. In addition, Damasio *et al.* [13] investigated whether these areas would be consistent with the naming deficits resulting from cortical lesions.

The Martin *et al.* [40] study examined silent naming of objects and animals with a subtraction control of passive viewing of nonsense objects. Relative to naming tools, naming animals produced activity in the left middle temporal gyrus (Brodmann's area 21). However, this region would be largely spared in the temporal lobe surgery. Martin *et al.*'s results are generally consistent with those of Damasio *et al.* [13] in showing a more posterior locus of tool naming (area 19) compared to that of animal naming. In addition, Martin *et al.* [40], like Sergent *et al.* [55], found increased anterior fusiform activity in the object naming tasks.

The lesions in the present investigation spared the areas that show the most PET activity for object recognition and only marginally affected, if at all, those areas that are active for animal recognition.

Present investigation: using brief, lateralized presentations to assess the role of cortical structures

The present investigation was designed to evaluate the role in higher level shape recognition of those regions of the temporal lobe that had been removed through surgery. To prevent the eye fixation strategies available in free viewing, the images were presented to the left or right of fixation so that they would be initially projected

to the contralateral hemisphere, at durations too brief to permit multiple fixations of the picture. We also measured the reaction time (RT) of the response so that differences in real-time efficiency could be assessed. In addition, the stimuli were line drawings, lacking surface cues of color and texture, as we wanted to evaluate whether shape information from contours corresponding to the significant orientation and depth discontinuities critical for object recognition [2, 9], rather than surface cues, is processed in these regions.

The images were presented in the left or right visual fields to eliminate callosal transfer to the opposite hemisphere by afferents along the vertical midline as far as possible in the ventral stream. In a landmark investigation, Clarke and Miklossy [10] were able to study callosal connections by silver-impregnating degenerating axons in sections of six human brains from individuals who had suffered recent unilateral occipital infarctions. They determined that in humans, there are no callosal connections from area 17 in one hemisphere to area 17 in the other. The connections are all from 17 to 18, and then only at the vertical midline. In the monkey, a 1° separation is sufficient to keep the activation in the ipsilateral temporal lobe until after V4, as cells in V4 rarely have receptive fields that extend beyond 1° in the contralateral field, although Desimone *et al.* [15] has shown that color-specific suppressive effects in V4 can be recruited from considerably greater distances in the contralateral hemisphere. Since V4 in the macaque has callosal connections along the horizontal meridian (D. Felleman, personal communication, 1991), the pictures were also presented above or below the horizon, so that they were in one of the four quadrants, to eliminate this possible source of inter-hemispheric transfer in case the homologous area to V4† in the human has the same callosal connectivity as in the macaque. We thus assume that shape information would have been confined to the hemisphere contralateral to the visual field in which the image was shown until at least an area that might be homologous with V4. After V4 in the monkey, callosal connections at the posterior temporal lobe would allow transfer to the contralateral temporal lobe [47]. Using autoradiographic measures, Pandya and Seltzer showed that all the interhemispheric connections at AIT are by way of the anterior commissure. If AIT is critical for the high level object recognition demanded by these tasks, the present experiment allowed a test of whether the interhemispheric connections in the temporal lobe are sufficient for real-time object recognition.

Two tasks, both using line drawings, were used. In Experiment 1, subjects had to judge whether two sequentially presented pictures (with an intervening mask) depicted the same or different object. Some trials included those where an object was depicted at either the same or different orientations in depth, so the actual contours in

the image could change. Different trials were of objects with the same name but a different shape. This task was run using both real objects and a highly similar set of nonsense objects (which, of course, did not have names). The ability to classify an image on the basis of its three dimensional shape is taken as an essential requirement for object recognition (e.g. Ref. [39]). In Experiment 2, subjects named pictures, with a second block run to assess priming. We view these tasks as 'high level' shape recognition activities; one engaging the verbal system for basic-level classification, the other requiring the determination of equivalence from non-identical images. Correct performance on these tasks could not be accomplished by use of a simple classification of contour, as, e.g. straight vs curved or long vs short.

Some of the individuals were missing the left AIT, others the right. It was not clear whether to expect any perceptual differences depending upon which temporal lobe was removed in that prior experiments in our laboratory [5] with normal human subjects did not reveal any hemifield differences in naming of object pictures to the left or right visual field. Furthermore, Nelson *et al.* [46] tested subjects prior to temporal lobe surgery using visually evoked potentials and found no differences between left- and right-damaged individuals. On the basis of the reports by A. Damasio *et al.* [12], H. Damasio *et al.* [13], and Flude *et al.* [19], however, deficits in naming might be expected from those individuals who had left temporal lobectomies. Further, Sergent *et al.* [55] and Martin *et al.* [40] found activation of the left temporal areas during their object recognition task.

Experiment 1: Same-different shape judgments of depth-rotated shapes

Method

Subjects. The patients were seven temporal lobectomized individuals (four left and three right; two males and five females) residing in the Twin Cities Metropolitan area who had undergone anterior temporal lobectomy at The University of Minnesota Medical Center. The surgical resident described the procedure as follows:

The extent of the resection posterior to the temporal pole is dictated to some degree by the findings from the electrocorticography recordings during surgery, but usually it will be about 4.5–5.0 cm in a left dominant hemisphere and 5.0–7.0 cm in a right non-dominant hemisphere. There are no anatomical landmarks for the resection line as such. In this center the superior temporal gyrus is not preserved. Middle and inferior temporal, and fusiform and parahippocampal gyri are removed as far back as the resection line. Mesially the lateral two-thirds of the amygdala are removed and the anterior 2–3 cm of the hippocampus. Occasionally, when the Wada test has suggested the importance of the mesial structures for memory function, the amygdala and hippocampus are not removed (K. G. Davies, personal communication, 18 July 1991).

Figure 1 shows a diagram made by Dr Davis illustrating the extent of the resection in the left hemisphere. Unfortunately, we had no post-operative MRI data that would allow us to quantify the degree of temporal lobe resection for individual

† Possible homologues in humans of macaque V4 will be considered in the discussion.

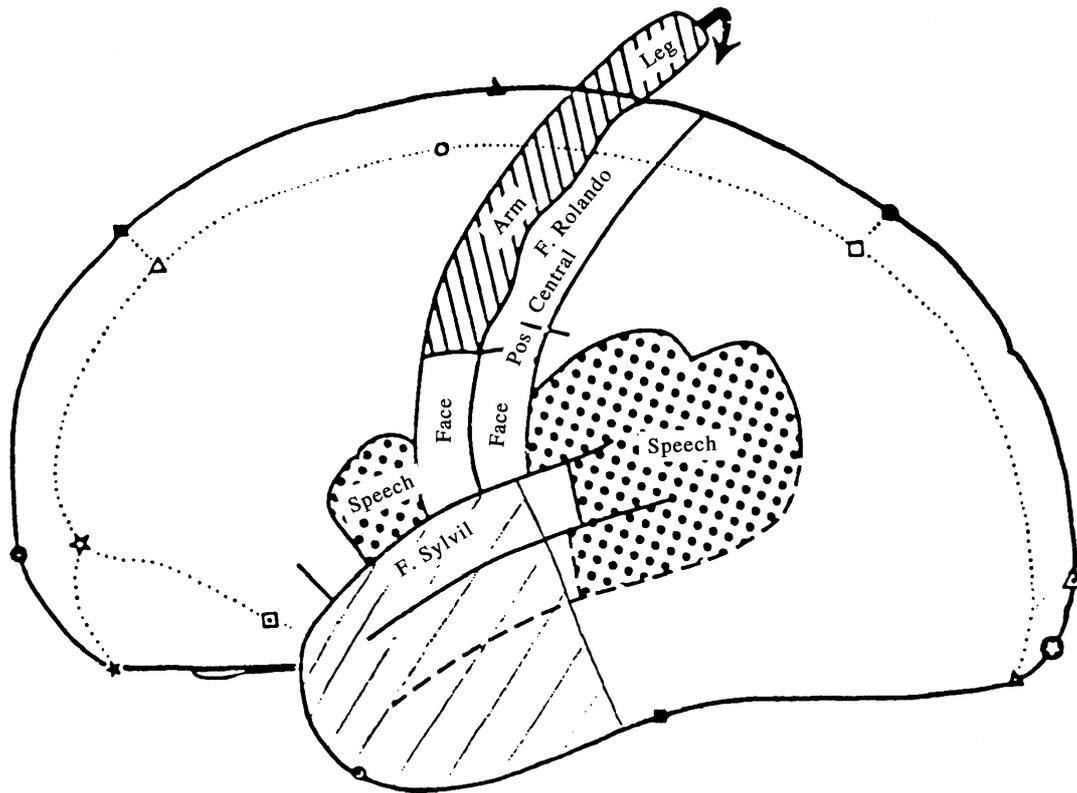


Fig. 1. Diagram illustrating the extent of the surgery (hand drawn diagonal lines) on a left hemisphere patient with the sparing of the speech areas. Lines drawn by K. G. Davis, M.D.

subjects. In addition, inspection of the neuropathology specimens revealed little structural specificity regarding exactly what structures were removed.

The patients were tested in a 2-hour session at the University of Minnesota. The control subjects were eight individuals, all right handed, tested at the University of Southern California. Table 1 shows the specifications of these subjects with respect to age, sex, education, handedness, and, for the patients, the dominant hemisphere, as determined by Wada testing [61], and the age at which the first seizure was experienced. The most striking difference in the histories between the right and left temporal lobe patients was in the age of the first seizure. For those with right lobectomies, the first seizure was in young adulthood, at age 18 (range 15–21) but all the left lobectomy individuals had been experiencing seizures since birth. This distinction, however, was merely a characteristic of our sample, rather than a general characteristic of this patient population [46].

The 1.9-year advantage in years of post-secondary education of the controls compared to the patients underestimates the real differences in educational level between the two groups. Six of the eight controls were university students, one was a Cal Tech post-doctoral student in computational neurobiology and one was a secretary with 2 years of post-secondary education. In all cases but one, the post-secondary school education for the patients was at Minnesota vocational technical schools that are comparable to community colleges. It is thus highly likely that the general ability level of the controls was considerably higher than that of the patients.

In this particular sample, the general ability level of the patients with right temporal lobe removal was, on average, considerably higher than those for the left. Table 2 presents a more detailed specification of the neuropsychological testing of the patients, performed prior to the recruitment of the subjects for the study. A general superiority in functioning of the right temporal lobe group is evident. This superiority does not appear

to be a consequence of the surgery as these differences can be seen with the preoperative IQ scores, shown in parentheses, as well as a number of the neuropsychological tests. Patient R3, who was prone to severe migraine headaches, suffered one while taking the verbal fluency test which could have accounted for his unusually low score.

Stimuli. Each subject viewed line drawings displayed on a high resolution screen controlled by a Macintosh IIx. Each picture could be inscribed in a circle with a diameter of 3.3° with a center that was 3.32° from the center of the fixation point. The closest possible distance between the center of the fixation point and the nearest contour of the object picture was, therefore, 1.67° , which was a minimum of 1.18° off the vertical and horizontal midlines. The stimuli were 18 nonsense objects composed of three or four simple volumes (organized into two sets of nine objects each, as shown in Fig. 2) in the nonsense objects task on Experiment 1, and 48 familiar objects (24 pairs of exemplars having the same name, e.g. two chairs) in the familiar objects task. The pictures were drawn in Cricket Draw and Adobe Illustrator with a line thickness of 2 pixels.

Sequence of events on a trial. Following the presentation of a central fixation dot (a disk that was 0.75° in diameter), subjects were given a 150 msec exposure to a picture of an object in one of the four quadrants of the screen, followed by a mask of random lines presented for 133 msec, as shown in Figs 3 and 4. The mask, which covered the stimulus location, was accompanied by a single digit at the fixation point. A second object was then presented for 200 msec in the same location as the first object. These durations were too brief for the subject to saccade to the picture and then return in time to catch the digit. Moreover, the flash of the digit attracted attention to itself. Subjects made verbal 'same'-'different' object judgments followed by a report of the digit. Response times were recorded using a Scientific Prototype voice key. A 'same' pair could be identical or rotated by 30° or 60° in depth.

Table 1. Specifications of patients and controls

	Age	Sex	Ed	Patients		Age of first seizure
				Dominant hand	Dominant hemisphere	
L1	27	F	2	R	L	Most of life
L2	21	M	1	R	L	Most of life (10 months)
L3	21	F	2	L	R	Most of life
L4	23	F	2	R	L	Entire life
R1	30	F	0	R	L	17
R2	35	F	6	R	L	21
R3	43	M	2	R	L	19
Mean	28.5		2.1			
Controls						
	Age	Sex	Ed			
A1	29	F	5			
A2	34	F	3			
A3	19	M	2			
A4	21	F	4			
B1	31	M	10			
B2	19	M	2			
B3	21	F	4			
B4	21	F	2			
Mean	24.4		4			

Notes: For the patients, L and R refer to side of lobectomy. For the controls, A and B refer to arbitrary matches to left and right lobectomized subjects, respectively. Education refers to years of post secondary schooling. 'Vocational Tech' was entered as 2 years (beyond high school), a MA as 6, and post-doctoral fellowship as 10. For the patients, the dominant hemisphere was determined by Wada testing. The phrases describing the age of first seizures are quotes from the testing report.

Table 2. Detailed specifications of neuropsychological testing of patients

	IQ	Rey-Osterrieth			VF	BNT	Months since surgery	Extent (cm) removed
		Time	NC	REC				
L1	96 (86)	132	31	12	27	45	4	5.0+ H
L2	99	221	31	19.5	<4	44	10	3.5+ H
L3	93 (85)	126	24	12	12	49	32	6.8+2.5 PTL + H
L4	110 (est.)	85	25	13	52	51	48	5.0
L Mean	99.5	141	27.8	14.1	23.7	47.2	21	5.7
R1	107 (101)	192	32	23.5	62	47	8	5.5+ H
R2	126 (117)	120	34	30.5	82	60	35	7.0+ H
R3	120 (118)	94	36	18	17	49	42	5.5+ H
R Mean	117.6	168.6	34	24	53.6	52	28.3	6.0
L + R Mean	107.3	138.6	30.1	18.3	36.6	49.3	25.5	5.8

Notes: IQ is the WAIS full scale score, administered postoperatively. The value in parentheses is the preoperative IQ score, when noted in the neuropsychological evaluations. Copy Time, NC, and REC all refer to the Rey-Osterrieth Complex Figure Test consisting of 36 segments to be copied; Copy Time is the number of seconds needed to copy the figure, NC refers to the number of segments (out of 36) correctly copied, REC is the number of segments recalled immediately after copying. VF refers to the percentile rank on a Verbal Fluidity test as assessed by controlled association. BNT is the number of pictures correctly named in the Boston Naming Test. In this test, a correct response requires that a sub-classification term be used, e.g. nozzle, not hose. Months since Surgery is the time, in months, between the patient's surgery and participation in the experiment. Extent Removed refers to the length (cm), as specified in the pathologists' reports, of temporal cortex removed from the anterior tip extending toward the occipital lobe. PTL refers to the posterior temporal lobe. H designates that the hippocampus was removed, though this was inferred, in two cases, from the tissue as coming from the 'mesial surface.' It is likely that H was removed in patient L4 but it was not specified in the pathologist's report.

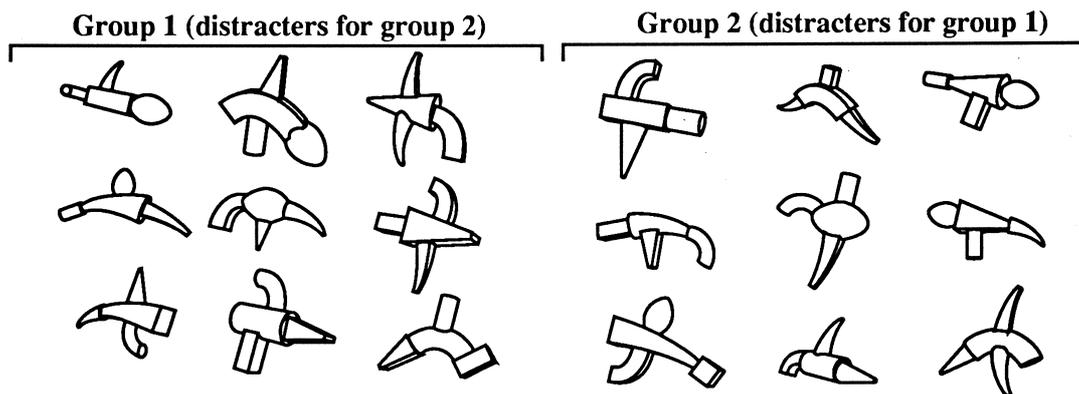


Fig. 2. The set of nonsense objects used in the same-different task. A 'different' trial would pair one of the objects from Group 1 with one from Group 2 (at 0°, 30° or 60° rotation in depth). Both images on a 'same' trial would be of the identical object but the object could be rotated in depth by 0°, 30° or 60°.

Nonsense objects task. This task used the two groups (1 and 2) of objects shown in Fig. 2 and was run prior to the familiar objects task. Each object was comprised of a large central geon [2] and three peripheral geons attached at (approximately) the 12, 3, 6, and 9 o'clock positions. Figure 3 shows an example of a 'same' trial with a 60° rotation in depth of a nonsense object. Ninety-six trials were run. On 'different' trials one nonsense object was drawn from each of the two groups.† The objects in each group were comprised of the same set of nine geons. Each geon was the central geon once in each group but shared no peripheral geons and, as far as possible, had different attachment positions with its cohort in the other group. Subjects could not base their response, therefore, on any individual geon or relation. The division of the objects into the two groups was arbitrary in that any one object could have been placed in the other group by swapping it with its same-central-geon cohort.

Familiar objects task. The 'different' trials on this task used the two members of an exemplar pair, so that the different objects always had the same name, as illustrated in Fig. 4. Subjects could not, therefore, base their response on the name or basic level classification of the images. The task with familiar objects (288 trials) was run immediately after the task with nonsense objects. Subjects were given 24 practice trials with the nonsense objects and 12 with the familiar objects. Experimental objects were not used for practice in either task.

Subjects were instructed to respond as quickly and accurately as possible and to provide the name of the fixation digit after their response. After each trial, feedback was provided to the subject as to the accuracy of the response and, if correct, the RT for that trial was displayed. One quarter of each of a subject's object images were displayed in each of the four quadrants, randomly ordered. The dual task requirement proved to be the most trying aspect of the procedure, and it was sometimes difficult for subjects to refrain from fixating the position where the object had been displayed, thereby missing the digit.

† This feature of the design, in which different trials of the nonsense objects came from different sets, was undesirable (and unnecessary), in that subjects could have used their learning of the different sets to distinguish same from different trials. However, as shown in Fig. 3, the sets were so similar that it would have been unlikely for such learning to have occurred over the 96 trials during this part of the experiment. As evidence for this last point, overall mean RTs in the second half of the block, where the differences would have been learned, were 6 msec longer than they were in the first half. There was only a modest reduction of 3.9% in error rates which would have been a consequence of all sources of learning, not just of the stimuli.

However, the objects were presented too briefly for the subjects to actually fixate the object while it was on and return to catch the digit. Trials on which a subject failed to accurately report the central number were dropped from the analysis as we could not be confident that the subject had maintained fixation on that trial. If a response was not made within 3 sec, the trial was considered an error.

Design. The basic design was 4 (quadrants) × 2 (responses, same or different) × 3 (orientations: 0°, 30° or 60°) for each subject. In the unfamiliar-objects task, four different objects (of the 18 total) were presented as the first stimulus in each of these 24 combinations of conditions to yield the 96 trials. In the familiar-objects task, 12 objects were presented in each of these combinations of conditions. In both tasks, order was counter-balanced such that for each subject who saw a particular sequence, a matched subject viewed the same sequence in reverse order. A different sequence, determined by Latin Square, with different objects assigned to the various conditions, was used for every pair of subjects. Because there were eight controls but only seven patients, there was a patient with an identical sequence for each control subject except for the eighth control subject. The consequence of the balancing was that all conditions, objects, and quadrants appeared equally often with each of the objects at the same mean serial position for the controls (save for a slight departure from complete balancing within the patient group because eight subjects were required for a completely balanced design).

Results

Figure 5 shows the reaction times (RTs) (Upper panel) and sensitivity (Lower panel), computed as Hits minus False Alarms (H-FA), for both patients and controls, as a function of rotation angle and hemisphere in the same-different experiment, collapsed over whether the stimuli were familiar or nonsense objects. As noted previously, the differences among the three groups (left patients, right patients, and controls) may have been present prior to the surgery: The one control subject who might have been closest to the patients in overall ability based on age and educational level, performed at a level that was within the range of the patients, as measured by both RT and accuracy.

The error bars in the RT figures were calculated by

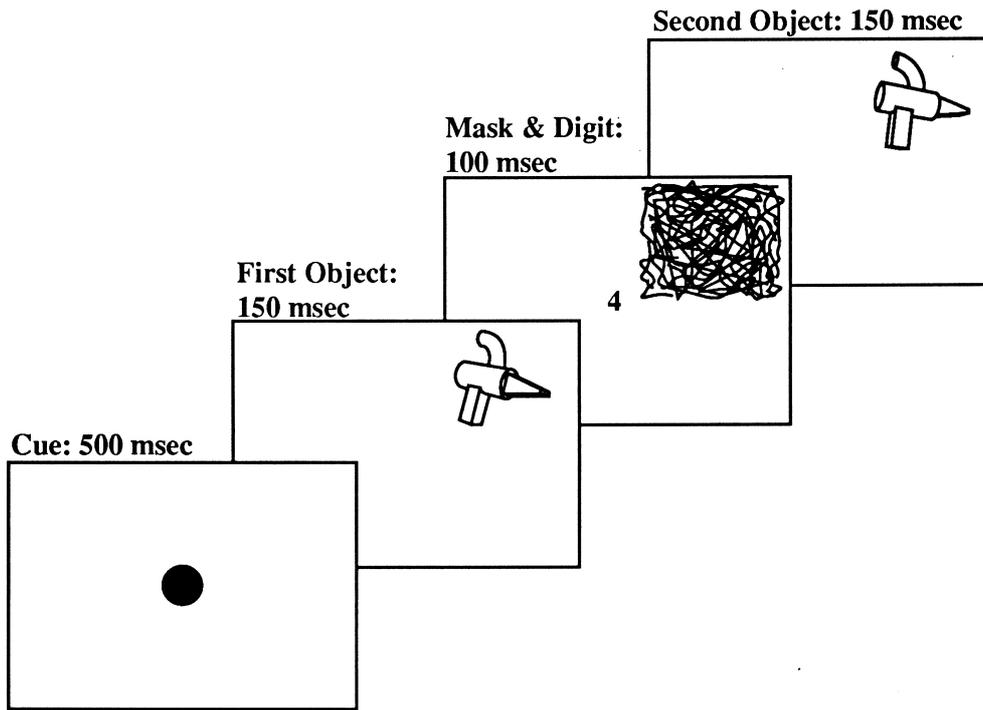


Fig. 3. Sequence of events on a 60° orientation difference, 'same' trial with nonsense objects in the same-different task. A correct response on this trial would require the response 'same-4' in that the small digit at the fixation point had to be accurately reported along with the judgment of the objects as same or different.

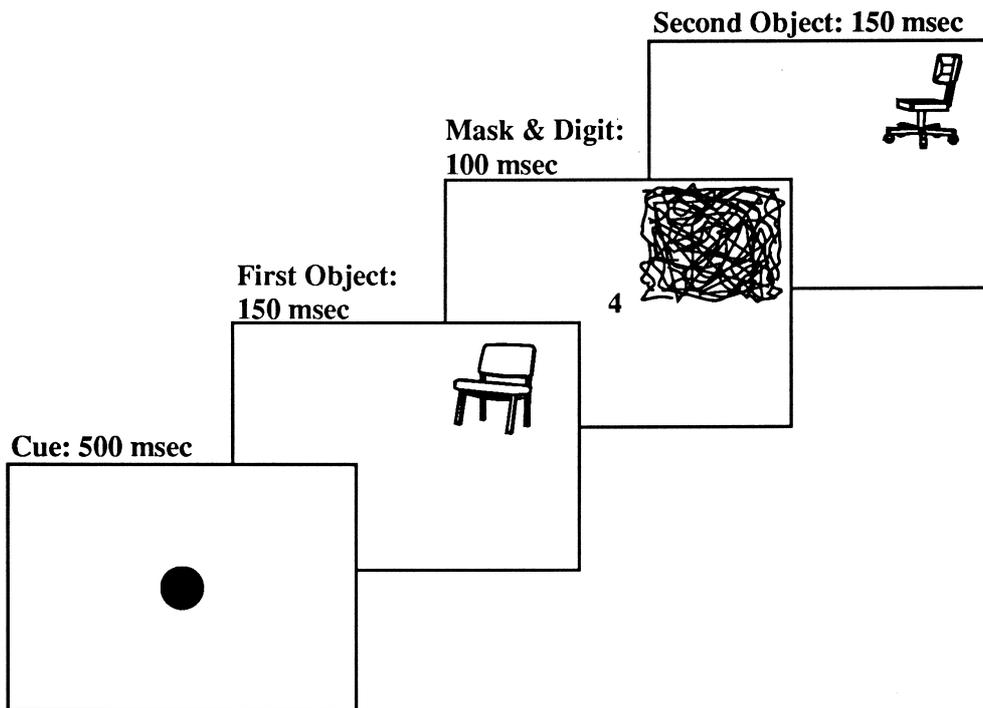


Fig. 4. Sequence of events on a 0° orientation difference, 'different' trial with familiar objects in the same-different task. The two objects in this task always had the same name (and basic level classification). Only if the two exemplars were identical, ignoring any differences of orientation in depth, were the subjects to respond 'same'. This is a different trial because the two exemplars of the class chair were not identical. (The designation of orientation on different trials was arbitrary.)

expressing each data point as a deviation score from the mean of that stimulus' overall RT so that variability between subjects and stimuli were removed. The remaining variance is attributable to interactions between

subjects, stimuli, and the interactions of subjects and stimuli with conditions. The error bars on the sensitivity values were calculated in the same manner.

For the patients, the overall levels of performance, as

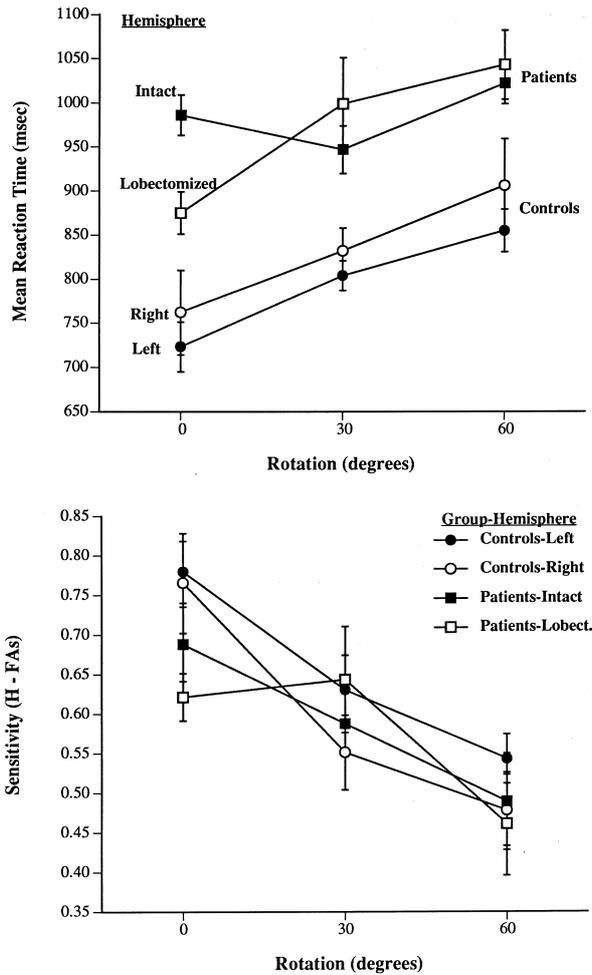


Fig. 5. (Upper panel). Mean correct RTs for 'same' responses as a function of subject group, angle of rotation, and hemisphere to which the image was projected (intact or sectioned for the patients; right or left for the controls) collapsed over object familiarity (familiar vs nonsense objects). Error bars were calculated after removal of between subject and between stimuli variance. (Lower panel). Mean sensitivity values (Hit Rate minus False Alarm Rate) for 'same' responses as a function of subject group, angle of rotation, hemisphere to which the image was projected (intact or sectioned for the patients; right or left for the controls) collapsed over object familiarity (familiar vs nonsense objects).

well as the effects of rotation angle, were virtually identical for images projected to the lobectomized and normal hemispheres. Although RTs were shorter and error rates lower for the controls compared to the patients, the slopes over rotation angle were virtually identical. The differences between patients and controls were in the intercept.

Table 3 shows the mean correct RTs and H-FA rates for a finer breakdown of the data, showing the effects of object familiarity and hemisphere of lesion. The patients with left lobectomies performed more slowly than those with right lobectomies, but the error rates were approximately the same. The incidence of errors caused by exceeding the three second limit were very rare. They occurred only 13 times in 2688 trials for the patients

(<1%) and only once in over 3000 trials for the controls. The secondary task of reporting the digits was relatively more difficult for the patients than the controls, constituting 7.7% of the total trials (6% when only the digit was in error; 1.7% when both digit and response was in error) of the patients and only 1.3% of the total trials (1.1% when only the digit was in error and 0.2% when both were in error) for the controls.

We found similar effects for familiar and nonsense objects: No difference between lobectomized and intact hemispheres was apparent. Moreover, the effects of rotation and familiarity were approximately the same in patient and control populations, suggesting that the deficits revealed by the patients were not a function of a distributed object recognition system that had undergone graceful degradation, i.e. a mass action effect. Such a degradation would be expected to produce greater costs from rotation or unfamiliarity or any variable that increased object recognition difficulty.

Experiment 2: Naming

In this task, which was always performed prior to the same-different task, subjects merely had to name objects. It is presented second as Experiment 2 because its procedure can be described as a subset of the same-different task studied in Experiment 1.

Method

The first three panels of Fig. 4 illustrate the sequence of events on a typical trial in this task, except: (a) the picture was exposed for 100 msec (rather than 150 msec) and (b) the mask and digit were exposed for 125 msec (rather than 100 msec). Note that only one object was presented on each trial; there was no object following the mask in Experiment 2. Two blocks of trials were run, each consisting of 128 pictures, including 40 pairs of objects with the same name but a different shape, such as a grand piano and an upright piano. The remaining 48 objects did not have a same name partner among the stimuli. A second block of trials was run with the same pictures, but in a different order, to assess priming (the magnitude of facilitation from the first to the second block). Stimuli were balanced for each subject such that stimuli which appeared in one quadrant on the first block were equally divided amongst all four quadrants in the second block. Half of these pictures were in the same left-right orientation as they were on the first block. The other half were mirror reversals of the first block images. These pictures generally differed from those used in the same-different task. Subjects were given 24 practice trials with pictures not in the experiment proper. The subjects and other aspects of the stimuli and procedures were identical to those in Experiment 1.

Responses synonymous with the feedback term, such as 'automobile' for 'car', were counted as correct. A response was recorded as an error if it was not the basic level name of the object presented or if the number was not accurately reported. False starts and responses that occurred more than three seconds after the object was displayed were also considered errors.

Table 3. Mean correct 'same' reaction times (msec), hits–false alarms, and rotation slopes (°/sec) as a function of subject group, hemisphere of presentation, and object familiarity

Group	Familiar objects Hemisphere				Nonsense objects Hemisphere			
	Left		Right		Left		Right	
	RT	H–FAs	RT	H–FAs	RT	H–FAs	RT	H–FAs
Controls	842	0.567	787	0.626	825	0.629	800	0.675
Left patients	1046	0.525	1066	0.625	1053	0.587	1019	0.600
Right patients	896	0.509	854	0.630	919	0.617	883	0.570
Slopes (°/sec)								
Controls	357		455		500		455	
Left patients	272		952		208		980	
Right patients	∞ *		472		1282		2584	

*This group showed no effect of rotation for stimuli presented to the left hemisphere. H-FAs is the hit rate minus the false alarm rate.

Results

Figures 6 and 7 show the overall RTs and error rates as a function of hemisphere for normal and control subjects. The overall effects of sectioning the temporal lobe in a hemisphere were slight: When an image was projected to the intact hemisphere, RTs were only 8 msec shorter and error rates only 2.4% lower than when the image was projected to the sectioned hemisphere. For the control subjects, the data are shown separately for left and right

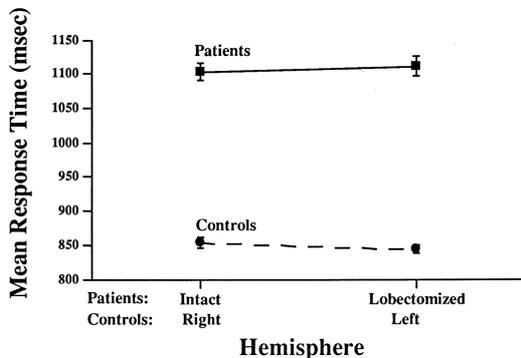


Fig. 6. Mean correct picture naming RTs as a function of subject group and hemisphere. Error bars were calculated after removal of between subject and between stimuli variance.

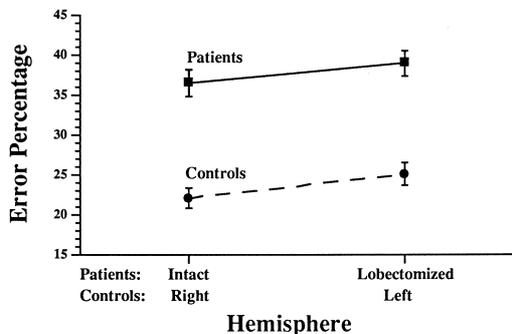


Fig. 7. Mean error rates in naming as a function of subject group and hemisphere. Error bars were calculated as the standard error for a proportion.

hemispheres. Unlike Experiment 1 (where there was an advantage for stimuli projected to the left hemisphere), there was virtually no effect of visual field on naming latency. Responses to images projected to the right hemisphere were only 10 msec faster, but with a 3.0% higher error rate, than images shown to the left hemisphere.

Second block data and priming. There was substantial facilitation in performance from the first to the second trial blocks: RTs were 92 msec lower and error rates were reduced by 14.2% for the patients and 57 msec and 13.5% lower for the controls, as shown in Table 4. That is, the magnitude of priming, when assessed as differences in performance from block 1 to block 2, was equivalent for patients and controls.

Figures 8 and 9 show second block RTs and error rates as a function of hemisphere. There was virtually no effect of whether the image was projected to the same or different hemispheres on the two blocks, nor whether the hemisphere was intact or sectioned.

As in Experiment 1, subjects with left temporal lobe sections performed somewhat worse than those with right temporal lobe sections, with overall RTs that were 64 msec longer and error rates that were 3.3% higher. Table 4 presents a finer breakdown of these data by block, hemisphere and group. Trials that were counted as errors because the three second limit was exceeded were relatively infrequent in both groups. For the patients with left temporal lobectomies, they occurred on 2.3% of the trials on the first block and only 1.4% on the second block. For the patients with right temporal lobectomies, the corresponding values were 3.6% and 1.0%. The incidence of these errors was independent of whether the image was presented to the intact or sectioned hemisphere. The patients erred on reporting the digit on 15.57% of the trials; 10.38% when only the digit was missed and 5.19% when both the digit and the object were missed. The controls missed the digit on a total of 8.93% of the trials; 6.15% when only the digit was incorrect and 2.78 when both the digit and the object were incorrect.

Table 4. Mean correct reaction times (msec) and error rates (%) in object naming as a function of group, hemisphere of projection, and block in Experiment 2

Patients	Response times (msec)			Errors (%)		
	Intact	Lobect.	Means	Intact	Lobect.	Means
Block 1	1163	1143	1153	43.5	46.2	44.9
Block 2	1043	1079	1061	29.5	31.7	30.6
Means	1103	1111		36.5	39.0	
Side of Lobectomy						
Right	1055	1086	1071	36.5	35.1	35.8
Left	1141	1128	1135	36.5	41.8	39.2
Controls						
Block 1	Right	Left	Means	Right	Left	Means
	875	879	877	29.1	31.6	30.4
Block 2	833	808	821	15.0	18.6	16.8
Means	854	844		22.1	25.1	

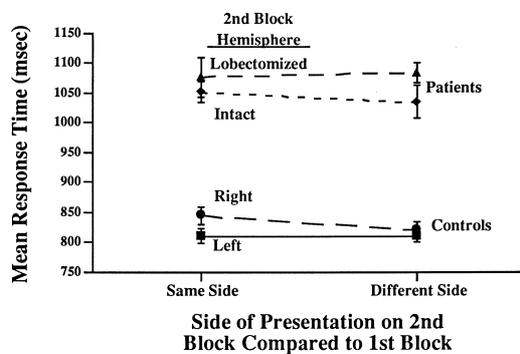


Fig. 8. Mean correct picture naming RTs on the second block as a function of subject group, hemisphere, and whether the image was projected to the same or different hemisphere on the first block. Error bars were calculated after removal of between subject and between stimuli variance.

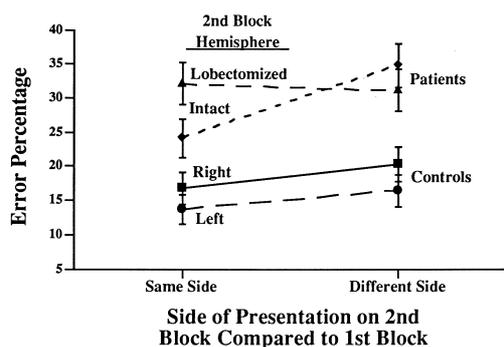


Fig. 9. Mean correct picture naming error rates on the second block as a function of subject group, hemisphere, and whether the image was projected to the same or different hemisphere on the first block. Error bars were calculated as the standard error for a proportion.

Discussion

An extreme, but possible, outcome of this experiment would have been that high level object recognition was

impossible when the image was projected to a hemisphere without the anterior regions of its temporal lobe, with normal performance if the image was projected to the intact hemisphere. Not only did we not obtain the extreme result, there was no evidence for the weaker effect: Projecting an image to the lobectomized hemisphere did not even result in a decrement in performance (in RTs or accuracy) relative to that obtained when the image was presented to the intact hemisphere. Even though only seven patients were run, these results are highly reliable in that each patient was used as his or her own control and the tasks were tightly controlled so stimulus variability could be removed. Consequently, the error bars in these experiments are substantially reduced from what they would be in the typical patient study (where error bars are rarely displayed).

In this discussion we consider three questions: (a) Why was there a difference between the patients and controls and a difference between left and right lobectomy groups? (b) Why was there no deficit in recognition when an image was projected to a lesioned, as compared to an intact, hemisphere? (c) If object recognition is performed prior to AIT, then what function is served by AIT?

Differences between groups

The performance of patients on these tasks was worse than that of the controls and the performance of the left lobectomy patients was worse than that of the right lobectomy patients. We conclude that these differences represent pre-experimental ability levels.

Patients vs controls. Although the patients performed at a markedly lower level on these speeded tasks than the controls, their disadvantage was not increased by factors that would increase difficulty in the same-different task, viz., stimulus unfamiliarity or rotation angle, nor did they evidence reduced priming in the naming task. Jensen [33]

has shown that performance on novel perceptual-motor tasks, such as those used here, is predictable from IQ scores. Thus the group differences, both between patients and controls and between left and right sectioned patients, likely reflect differences (a) that existed prior to the surgery, and (b) occurred in stages subsequent to object recognition, such as decision and response selection. A somewhat less parsimonious possibility, which we cannot definitively rule out at this time, is that the lesioned tissue was participating in a distributed (or mass action) system that manifested graceful degradation for shape recognition and that other areas subserved rotational invariance and priming. This alternative is termed less parsimonious in that it posits additional tissue for priming and rotational matching that would be independent of an object recognition system.

Dual-task performance. Much of the difficulty in performing these tasks, as noted previously, derived from the dual task methodology that was employed to induce the subjects to remain on central fixation. This was certainly the subjective impression of the authors when they performed the task. Even when the digit was fixated, sometimes it was forgotten during the brief interval when attention was directed toward making the object response. Consequently, there is little doubt that performance on these tasks was lowered and variability increased because of the requirements for report of the fixation digit. Nonetheless, the data do not suggest that removal of AIT in one hemisphere differentially affected performance on either the shape recognition or the digit reporting tasks.

We note here a potential benefit to be derived from the secondary task requirement of reporting the fixation digit. Without a secondary task, a subject might be able to devote excess 'attentional capacity' to counteract perceptual deficits. For example, the reduction in intelligibility of artificial speech compared to natural speech is not readily observed unless a secondary task is imposed that prevents the listener from attending more to the artificial speech [38]. In the present case, the presence of the secondary task would have reduced the attentional capacity that the subject might have devoted to make up for less efficient perceptual processing, if it existed, in the sectioned hemisphere.

Lower performance with the left lobectomy patients compared to the right. Recall that the review of the PET and human lesion results indicated that naming was more likely to be associated with damage to the left hemisphere than the right. Although overall performance was worse for those patients with left lobectomies, the deficit compared to those with right lobectomies was modest: only a 3.3% higher error rate and 64 msec increase in naming RTs. As noted earlier, it is likely that the worse performance of the left lobectomy patients was a function of a lower level of functioning in general, as their pre-operative performance was lower than that for the right sectioned patients, rather than an object recognition deficit, in that they were still able to perform the task

and reported little difficulty in their day to day object recognition activities. They also performed worse than the right lobectomy patients on the depth-rotation matching task, which did not require naming. Our results suggest that to the extent that lesions to the left middle temporal area result in severe naming deficits, the deficit may be a consequence of damage to the posterior regions of this cortical area (which were spared in the present surgery) or collateral damage to the speech areas. The presence of a right homonymous hemianopia and naming difficulties in the patient tested by Flude *et al.* [19] suggest that his lesion may have extended to these more posterior temporal regions and to areas in the occipital lobe.

Although we cannot definitively rule out a general deficit in naming from left AIT lesions, we find little support for it with these particular patients performing these tasks.

Webster *et al.* [62] have documented a considerable degree of developmental plasticity associated with lesions of TE in that such lesions in infant monkeys resulted in the preservation of transient connections from TEO to the amygdala that are normally pruned as the animal matures. Not only were these connections preserved but areas of the amygdala that are normally served by area TE were occupied by connections that sprouted from TEO. Of the patients in the present investigation, seizures were evidenced at a very young age only by the left AIT sectioned patients, not by the right. There would have thus been more opportunity for infant plasticity to compensate for the lesion producing the seizures in the left sectioned group. If this did occur, then the results suggest that these connections were not critical for performance on the present tasks, as the right AIT sectioned subjects, especially R2 whose surgery was a result of a tumor rather than an infant lesion, performed quite well with images projected to the right hemisphere.

How could there have been substantial object processing when an image was projected to the lobectomized hemisphere?

We consider, and accept, two non-exclusive answers to the title question of this section: (a) The anterior halves of the temporal lobes are not required for high level object recognition of the kind assessed here, and (b) object recognition can be accomplished by the remaining temporal lobe through efficient callosal transfer.

Object recognition in humans is achieved by regions posterior to AIT. This position has been previously advanced by Farah [16] and our results are consistent with her hypothesis.

Lesions to V4 in the macaque result in severe deficits in shape discrimination [53], as do lesions to IT [22]. There is now ample evidence that V4 is critical for achieving invariance over variations in position, size, occlusion, and contour noise [53], and the capacity to respond to a less salient attribute value, i.e. something that is smaller

or of lower contrast [54].[†] In contrast, lesions to IT have no effect on transferring an originally learned discrimination to one in which the stimuli are changed in size or orientation [29].

Given that V4 supports the computation of a representation that is invariant over changes in viewpoint, occlusion, noise, and lighting, it is possible that TEO supports the derivation of a reduced code, such as that specifying a structural description, that can be efficiently associated with an arbitrary response. Whereas the deficits in V4 are apparent in performing a simultaneous discrimination, such as making a fixation to a match-to-sample [53] or 'odd man out' [54], the deficits in IT lesions are revealed in learning to associate a stimulus to an arbitrary response. By this account, TEO lesions might produce smaller deficits in performing the simultaneous discrimination tasks developed by Schiller.

What area in humans might serve an equally critical role as IT and AIT for the representation of object shape? Two candidates are the dorsolateral regions of area 19, just rostral to area 18, and the fusiform gyri, area 37. Both would have been spared by the temporal lobectomies.

Kaas [34], on the basis of anatomical parallels, has argued that V4 in humans may be on the dorsolateral surface of the occipital cortex with IT at the occipital-temporal border, near the superior temporal region. This is the region that Farah [16] has proposed as a site critical for the representation of object parts. It is also the locus of one of the lesions of D. F. [23, 24] who had normal color and motion perception as well as a capacity for motor interactions with objects but no capacity for either judging the orientation of a segment or determining its identity on the basis of its shape. This area also includes the region of the middle occipital gyrus where increased PET activity (though only on the left side) in object classification tasks has been found.

Almost all studies of PET activity during object classification show involvement of the fusiform gyri. This is the region (though only on the left side) that Farah [16] had argued is involved in the representation of multiple shapes, such as the different parts of an object or letters of a word. The experiment by Martin *et al.* [40] reported

[†] Although we are unaware of direct comparisons of IT and V4 lesions for shape, a somewhat opposite picture is apparent for color discriminations than the one we are proposing for shape discriminations. Heywood *et al.* [26, 27] showed that color discriminations are only moderately affected by V4 lesions but severely affected by IT lesions just anterior to V4. These IT lesions largely spare achromatic discrimination abilities. Zeki [63] and Zeki *et al.* [64] proposed that because lesions of the lingual and posterior fusiform gyri produce cerebral achromatopsia, then these gyri might be homologous to V4 in the macaque, as a large portion of V4 cells show strong color tuning. Heywood *et al.* [26, 27] suggest that the ventromedial lesions to the lingual and fusiform gyri that produced cerebral achromatopsia might have had their effects either on tissue that is functionally equivalent to IT or on the fibers from earlier visual areas to IT. It is possible that the color tuning evident in V4 is involved in the calculation of color constancy.

bilateral activation of both the left and right fusiform gyri in silent naming of both animal and tools.

Even if recognition was accomplished at these stages that are earlier than AIT, then worse performance in a lesioned hemisphere might have been produced because of retrograde degeneration of feed forward fibers from earlier visual areas to later visual areas or direct destruction of backward fibers, which have been well documented in the case of AIT → V4 in the monkey [17]. It is of interest that no effects of such degeneration, if it did occur, were apparent in the subjects' performance.

Object recognition was accomplished by the intact temporal lobe. To consider this possibility, we must address two related questions: What is the pathway for real-time object (shape) recognition? What are the inter-hemispheric connections of this pathway? These questions are critical for interpreting the lateralized presentations of the current investigation for, as noted in the Introduction, even if AIT was required for object recognition, a completely efficient transfer of information to the intact temporal lobe prior to AIT would render unilateral removal of AIT ineffective (or less effective) for producing sizable deficits in performance.

In V4 in the macaque, the receptive fields rarely cross more than a degree or so into the ipsilateral visual field [15].[†]

The experiments by Schiller [53] and Schiller and Lee [54] can be interpreted as providing direct evidence against efficient callosal transfer of stimulus information in V4 or at an earlier stage in the macaque. In these studies a portion of V4 that coded either the lower left or lower right portion of the visual field was lesioned. The Schiller and Lee [54] experiment employed circular displays of eight items (positioned as the hours on a clock face). Monkeys were trained to make a saccade from a central fixation point to an item that differed from the rest of the items in the display (all of which were the same). The monkeys showed decrements in their ability to respond to low saliency targets (e.g. smaller, lower contrast) when such targets were in the receptive field of the lesion. If there was efficient transfer of stimulus information to the intact hemisphere then these deficits should not have been evidenced. Similar reason can be applied to the Schiller [53] study, although there the monkey had to make a saccade to a peripheral stimulus that matched a central one.

[†] Although the classic receptive fields only rarely extend across the midline, and then only for less than a degree, there are long-range suppressive interactions in V4 such that a large bar, of the same color and orientation as the preferred stimulus, but located up to 16° from the midline in the opposite visual field, can inhibit responding [15]. Desimone *et al.* [15] theorize that these inhibitory effects, mediated by the corpus callosum, may play a role in color constancy and figure ground segregation.

How much visual processing for real-time object recognition remains after V4? Extensive callosal connections suggest that a perceptual description has already been achieved

What amount of visual processing remains to be accomplished in areas following V4? In the macaque, although the receptive fields in V4 are larger than they are in V1 and V2, they still retain a retinotopic organization and relatively simple spatial tuning functions (though with variation in the coordinate spaces in that some units are tuned to patterns that are defined by polar and hyperbolic coordinate) [21]. Moreover, after V4, in IT, the receptive fields are enormous, averaging 26°, and retain little, if any, retinotopic correlation. There are extensive connections in TEO after V4 in the macaque, as well as through the anterior commissures [47], and such connections most likely characterize human posterior temporal cortex as well. These connections, assuming that they are the same in humans and macaques, would allow transfer of information prior to AIT, in the posterior inferior temporal lobe (PIT).

Biederman [3] has proposed that the existence of widespread callosal efferent connections from a given stage are an indication that a reduced code has been activated. For example, a small number of units are all that are required for specifying a viewpoint invariant, structural description of an object (e.g. Ref. [30]). The callosal connections that are present in early stages, viz., V1, V2, and V4, in the macaque, are presumed to facilitate grouping across the midline but nothing else. Widespread inter-hemispheric transmission of image information prior to V4 would be of spatial filter activation values that would only add noise to the cells in the contralateral hemisphere. Once a reduced code can be activated, its transmission to the opposite hemisphere allows it to have access to a high level perceptual description of the ipsilateral visual field.

If priming is mediated by representations in AIT, then the results of the naming task suggest that the inter-hemispheric transfer of information required for these representations is completely efficient in that there was as much priming for images presented to the lesioned hemisphere as to the intact hemisphere. The transfer would have to be by way of the callosum, at PIT, rather than at the anterior commissures, in that the lesioned area was that served by the anterior commissures. The independence of the magnitude of priming on whether the stimuli are presented to the same or different hemispheres has been observed previously in normal subjects [4, 11].

If not object recognition, what visual functions are performed by the anterior regions of the temporal lobes?

A question that is raised by a posterior locus for object recognition concerns just what visual functions are served by the massive amounts of tissue removed in temporal

lobectomies. It is unlikely that the anterior reaches of the temporal lobes are on some vestigial path (bookends for the hippocampus?) in that they show enormous expansion relative to infrahuman primates. One possibility is that they are part of a distributed memory system and the poorer performance revealed by our patients relative to the controls reflects a form of graceful degradation characteristic of such systems. Somewhat inconsistent with this interpretation is that the patients did not reveal greater deficits than the controls with unfamiliar or rotated objects, as would be expected from a less efficient system. There is little doubt, however, that temporal-lobe patients demonstrate some longer term memory loss for visual information (e.g. Refs [41], [48], [56] and [57]).

Rather than real-time object recognition, it is possible that AIT is part of a system for representing visual episodes and integrating them with prior knowledge, a hypothesis suggested by Sergent *et al.* [55]. This hypothesis can explain how it is possible to have so many cells tuned to arbitrary experimental stimuli in the Logothetis and Miyashita experiments, without any noticeable deficit (we presume) in responding to new stimuli. The estimates ($\approx 11\%$ in the Logothetis *et al.* [37] experiment) of the proportion of cells showing preferential tuning to the specific experimental stimuli could suggest a calculation that the monkey could only represent the stimuli from about 9–10 experiments before it would run out of coding capacity. The Li *et al.* [36] results suggest that many of these AIT cells are participating in competitive interactions so that the net result would be that relatively few cells would be coding any one episode but that the intense experimental training is providing many current episodes. The competitive aspect of AIT coding could explain why it is so rare for AIT to show any activity in PET imagery. It would be instructive to test for tuning to experimental stimuli after a new set of experimental stimuli has been learned.

Another possible function for AIT, suggested by the difficulty that temporal patients have in interpreting picture anomalies [43, 44], is that temporal lobes play a role in the quick access to semantic information derived from the arrangement of multiple objects into scenes. In constructing an extraordinarily large number of arbitrary but predictable environments, humans have a far greater demand for scene representations than monkeys. Thus, capacity for scene representation may ultimately prove to be a special case of episodic integration but scene representations are likely to be one of the more significant ways in which human and monkey visual capabilities differ.

Conclusion

The recognition of images projected to a hemisphere lacking AIT does not suffer relative to the recognition of images projected to an intact hemisphere. Two non-exclusive implications follow from this result: (a) Object

recognition may be completed prior to the lesioned tissue, perhaps at the occipital-temporal border or in a more posterior region of the temporal lobe with no retrograde interference effects from the lesioned tissue, and (b) completely efficient callosal transfer allows the intact hemisphere to achieve recognition as effectively as when it is projected to the ipsilateral hemisphere. Rather than real-time object recognition, AIT may subserve the representation of visual episodes and scenes.

Acknowledgements—This research was supported by the Fesler-Lampert Endowment Fund at the University of Minnesota. We thank Drs. Ilo Lepick and Robert Gumnit of the Minnesota Comprehensive Epilepsy Program for referring patients to this study and Moshe Bar and Peter Kalocsai for their helpful comments on the manuscript.

References

1. Bayliss, G. C., Rolls, E. T. and Leonard, C. M. Functional subdivisions of the temporal lobe neocortex. *The Journal of Neuroscience* **7**, 330–342, 1987.
2. Biederman, I. Recognition-by-components: A theory of human image understanding. *Psychological Review* **94**, 115–147, 1987.
3. Biederman, I. What is transferred across the callosum? Unpublished manuscript. University of Southern California, 1996.
4. Biederman, I. and Cooper, E. E. Object recognition and laterality: Null results. *Neuropsychologia* **29**, 685–694, 1991.
5. Biederman, I. and Cooper, E. E. Evidence for complete translational and reflectional invariance in visual object priming. *Perception* **20**, 585–593, 1991.
6. Biederman, I. and Cooper, E. E. Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance* **18**, 121–133, 1992.
7. Biederman, I. and Gerhardstein, P. C. Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance* **19**, 1162–1182, 1993.
8. Biederman, I. and Gerhardstein, P. C. Viewpoint-dependent mechanisms in visual object recognition: Reply to Tarr and Bülthoff (1995). *Journal of Experimental Psychology: Human Perception and Performance* **21**, 1506–1514, 1995.
9. Biederman, I. and Ju, G. Surface vs edge-based determinants of visual of visual object recognition. *Cognitive Psychology* **20**, 38–64, 1988.
10. Clarke, S. and Miklossy, J. Occipital cortex in man: Organization of callosal connections, related myelo- and cytoarchitecture, and putative boundaries of functional visual areas. *The Journal of Comparative Neurology* **298**, 188–214, 1990.
11. Cooper, E. E., Biederman, I. and Hummel, J. E. Metric invariance in object recognition: A review and further evidence. *Canadian Journal of Psychology* **46**, 191–214, 1992.
12. Damasio, A., Damasio, H., Tranel, D. and Brandt, J. P. The neural regionalization of knowledge access: Preliminary evidence. *The Brain*. Cold Spring Harbor Symposium, 1990.
13. Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D. and Damasio, A. R. A neural basis of lexical retrieval. *Nature* **380**, 499–505, 1996.
14. Desimone, R., Albright, T. D., Gross, C. G. and Bruce, C. Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of Neuroscience* **4**, 2051–2062, 1984.
15. Desimone, R., Moran, J., Schein, S. J. and Mishkin, M. A role for the corpus callosum in visual area V4 of the macaque. *Visual Neuroscience* **10**, 159–171, 1993.
16. Farah, M. J. *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision*. MIT, Cambridge, MA, 1990.
17. Felleman, D. J. and Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* **1**, 1–47, 1991.
18. Fiser, J. and Biederman, I. Size invariance in visual object priming of gray scale images. *Perception* **24**, 741–748, 1995.
19. Flude, B. M., Ellis, A. W. and Kay, J. Face processing and name retrieval in an amonic aphasic patient: Names are stored separately from semantic information about familiar people. *Brain and Cognition* **11**, 60–72, 1989.
20. Fujita, I., Tanaka, K., Ito, M. and Cheng, K. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* **360**, 343–356, 1992.
21. Gallant, J. L., Braun, J. and Van Essen, D. C. Selectivity for polar, hyperbolic, and Cartesian gratings for macaque visual cortex. *Science* **259**, 100–103, 1993.
22. Gross, C. G. How inferior temporal cortex became a visual area. *Cerebral Cortex* **4**, 455–469, 1994.
23. Goodale, M. A., Milner, D. A., Jakobson, L. S. and Carey, D. P. A neurological dissociation between perceiving objects and grasping them. *Nature* **349**, 154–156, 1991.
24. Goodale, M. A., Humphrey, G. K., Milner, D. A., Jakobson, L. S., Servos, P. and Carey, D. P. Object versus picture identification in a patient with visual form agnosia. Paper presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL, May, 1991.
25. Haxby, J. V., Grady, C. L., Horowitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B. and Rapoport, S. I. Dissociation of object and spatial processing in human extrastriate cortex. *Proceedings of the National Academy of Sciences, U.S.A.* **88**, 1621–1625, 1991.
26. Heywood, C. A., Gadotti, A. and Cowey, A. Cortical area V4 and its role in the perception of color. *The Journal of Neuroscience* **12**, 4056–4065, 1992.
27. Heywood, C. A., Gaffan, D. and Cowey, A. Cerebral achromatopsia in monkeys. *European Journal of Neuroscience* **7**, 1064–1073, 1995.
28. Holmes, E. J. and Gross, C. G. Effects of inferior temporal lesions on discrimination of stimuli differ-

- ing in orientation. *Journal of Neuroscience* **4**, 3063–3068, 1984.
29. Holmes, E. J. and Gross, C. G. Stimulus equivalence after inferior temporal lesions in monkeys. *Behavioral Neuroscience* **98**, 898–901, 1984.
 30. Hummel, J. E. and Biederman, I. Dynamic binding in a neural network for shape recognition. *Psychological Review* **99**, 480–517, 1992.
 31. Intraub, H. Identification and naming of briefly glimpsed visual scenes. In *Eye Movements: Cognition and Visual Perception*, D. F. Fisher, R. A. Monty and J. W. Senders (Editors), pp. 181–190. Lawrence Erlbaum, Hillsdale, NJ, 1981.
 32. Iwai, E. and Mishkin, M. *Vision, Memory, and the Temporal Lobes*. Elsevier, New York, 1990.
 33. Jensen, A. R. Spearman's hypothesis tested with chromometric information-processing tasks. *Intelligence* **17**, 44–77, 1993.
 34. Kaas, J. H. Do humans see what monkeys see? *Trends in Neuroscience* **15**, 1–3, 1992.
 35. Kobatake, E. and Tanaka, K. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology* **71**, 856–867, 1994.
 36. Li, L., Miller, E. K. and Desimone, R. The representations of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology* **69**, 1918–1929, 1993.
 37. Logothetis, N. K., Pauls, J. and Poggio, T. Shape representation in the inferior temporal cortex of monkeys. *Current Biology* **5**, 552–563, 1995.
 38. Luce, P. A., Feustel, T. C. and Pisoni, D. B. Capacity demands in short-term memory for synthetic and natural speech. *Human Factors* **25**, 17–32, 1983.
 39. Marr, D. *Vision*. Freeman, San Francisco, 1982.
 40. Martin, A., Wiggs, C. L., Ungerleider, L. G. and Haxby, J. V. Neural correlates of category-specific knowledge. *Nature* **379**, 649–652, 1996.
 41. McCarthy, R. A. and Warrington, E. K. *Cognitive Neuropsychology: A Clinical Introduction*. Academic Press, San Diego, 1990.
 42. Miller, E. K., Li, L. and Desimone, R. Activity of neurons in anterior inferior temporal cortex during a short term memory task. *The Journal of Neuroscience* **13**(4), 1460–1478, 1993.
 43. Milner, B. Psychological deficits produced by temporal-lobe excision. *Proceedings of the Association for Research in Nervous and Mental Disorders, Vol. 36, The Brain and Human Behavior*, pp. 244–257. Hafner, New York, 1958.
 44. Milner, B. Visual recognition and recall after right temporal-lobe excision in man. *Neuropsychologia* **6**, 191–209, 1968.
 45. Miyashita, Y. and Chang, H. S. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* **331**, 68–70, 1988.
 46. Nelson, C. A., Collins, P. F. and Torres, F. P300 brain activity in patients preceding temporal lobectomy. *Archives of Neurology* **48**, 141–148, 1991.
 47. Pandya, D. N. and Seltzer, B. The topography of commissural fibers. In *Two Hemispheres—One Brain*, F. Lepore, M. Ptito, H. H. Jasper (Editors), pp. 47–73. Alan R. Liss, New York, 1986.
 48. Pigott, S. and Milner, B. Memory for different aspects of complex visual scenes after right anterior temporal lobectomy. *Neuropsychologia* **31**, 1–15, 1993.
 49. Plaut, D. C. and Farah, M. J. Visual object representation: Interpreting neurophysiological data within a computational framework. *Journal of Cognitive Neuroscience* **2**, 320–343, 1990.
 50. Rosch, E., Mervis, C. B., Gray, W., Johnson, D. and Boyes-Braem, P. Basic objects in natural categories. *Cognitive Psychology* **8**, 382–439, 1976.
 51. Sakai, K. and Miyashita, Y. Neuronal tuning to learned complex forms in vision. *NeuroReports* **5**, 829–832, 1994.
 52. Sakai, K., Naya, Y. and Miyashita, Y. Neuronal tuning and associative mechanisms in form representation. *Learning and Memory* **1**, 83–105, 1994.
 53. Schiller, P. H. Effect of lesions in visual cortical area V4 on the recognition of transformed objects. *Nature* **376**, 342–344, 1995.
 54. Schiller, P. H. and Lee, K. The role of the primate extrastriate area V4 in Vision. *Science* **251**, 1251–1253, 1991.
 55. Sergent, J., Ohta, S. and Macdonald, B. Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain* **115**, 15–36, 1992.
 56. Shallice, T. *From Neuropsychology to Mental Structure*. Cambridge University Press, Cambridge, U.K., 1988.
 57. Smith, M. L. Memory disorders associated with temporal-lobe lesions. In *Handbook of Neurology*, F. Boller and J. Grafman (Editors), Vol. 3, pp. 91–106. Elsevier, Oxford, 1989.
 58. Tanaka, K. Inferotemporal cortex and object vision: Stimulus selectivity and columnar organization. *Annual Review of Neuroscience* **19**, 109–139, 1996.
 59. Tanaka, K. Neuronal mechanisms of object recognition. *Science* **262**, 685–688, 1993.
 60. Ungerleider, L. G. and Mishkin, M. Two cortical visual systems. In *Analysis of Visual Behavior*, D. J. Ingle, M. A. Goodale and R. J. W. Mansfield (Editors), pp. 549–586. MIT, Cambridge, MA, 1982.
 61. Wada, J. and Rasmussen, T. Intracarotid injection of sodium Amytal for the lateralization of cerebral speech dominance: Experimental and clinical observations. *Journal of Neurosurgery* **17**, 266–282, 1960.
 62. Webster, M. J., Ungerleider, L. G. and Bachevalier, J. Lesions of inferior temporal area TE in infant monkeys alter cortico-amygdalar projections. *NeuroReport* **2**, 769–772, 1991.
 63. Zeki, S. A century of cerebral achromatopsia. *Brain* **113**, 1721–1777, 1990.
 64. Zeki, S., Watson, J. D. G., Lueck, C. J., Friston, K. J., Kennard, C. and Frackowiak, R. S. J. A direct demonstration of functional specialization in the human visual cortex. *The Journal of Neuroscience* **11**, 641–649, 1991.