Behavioral/Cognitive

Mild Perceptual Categorization Deficits Follow Bilateral Removal of Anterior Inferior Temporal Cortex in Rhesus Monkeys

Narihisa Matsumoto,[1,2]* ⬤Mark A.G. Eldridge,[2]* Richard C. Saunders,[2] Rachel Reoli,[2] and ⬤Barry J. Richmond[2]

[1]Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305-8568, Japan, and [2]Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892

In primates, visual recognition of complex objects depends on the inferior temporal lobe. By extension, categorizing visual stimuli based on similarity ought to depend on the integrity of the same area. We tested three monkeys before and after bilateral anterior inferior temporal cortex (area TE) removal. Although mildly impaired after the removals, they retained the ability to assign stimuli to previously learned categories, e.g., cats versus dogs, and human versus monkey faces, even with trial-unique exemplars. After the TE removals, they learned in one session to classify members from a new pair of categories, cars versus trucks, as quickly as they had learned the cats versus dogs before the removals. As with the dogs and cats, they generalized across trial-unique exemplars of cars and trucks. However, as seen in earlier studies, these monkeys with TE removals had difficulty learning to discriminate between two simple black and white stimuli. These results raise the possibility that TE is needed for memory of simple conjunctions of basic features, but that it plays only a small role in generalizing overall configural similarity across a large set of stimuli, such as would be needed for perceptual categorical assignment.

Key words: behavior; categorization; generalization; lesion; perception

---

Significance Statement

The process of seeing and recognizing objects is attributed to a set of sequentially connected brain regions stretching forward from the primary visual cortex through the temporal lobe to the anterior inferior temporal cortex, a region designated area TE. Area TE is considered the final stage for recognizing complex visual objects, e.g., faces. It has been assumed, but not tested directly, that this area would be critical for visual generalization, i.e., the ability to place objects such as cats and dogs into their correct categories. Here, we demonstrate that monkeys rapidly and seemingly effortlessly categorize large sets of complex images (cats vs dogs, cars vs trucks), surprisingly, even after removal of area TE, leaving a puzzle about how this generalization is done.

Introduction

Categorizing visual stimuli through generalization of visual features seems effortless, but has been difficult to replicate in artificial systems (Poggio et al., 2012). In the canonical view of

hierarchical convergent processing in the visual system, the neural representation of complex objects emerges as visual information is integrated through a set of hierarchically organized processing stages ending in the primate inferior temporal lobe (Fukushima, 1980; Mishkin et al., 1983; Riesenhuber and Poggio, 2000). This view explains the results of many behavioral and physiological studies that, taken together, established that the anterior inferior temporal cortex (area TE) is essential for visual perception and memory of complex whole objects (Iwai and Mishkin, 1968; Cowey and Gross, 1970; Gross et al., 1972; Desimone et al., 1984; Weiskrantz and Saunders, 1984; Tanaka, 1996; Sugase et al., 1999; Sigala and Logothetis, 2002; Tsao et al., 2003; Afraz et al., 2006; Kiani et al., 2007; Bell et al., 2009; Sugase-Miyamoto et al., 2014). It seems reasonable to infer that visual perceptual categorization, the process by which complex, multifaceted objects are grouped by physical similarity, ought to require the integrity of area TE. Indeed, rhesus monkeys with lesions of area TE are impaired in discriminating between

pairs of objects (Iwai and Mishkin, 1968; Cowey and Gross, 1970). Also, single neurons in area TE show selectivity for visual stimuli, most notably for faces or complex objects (Gross et al., 1972; Desimone et al., 1984; Tanaka, 1996). Furthermore, physiological recordings (multiple single-cell recordings, optical recordings, and functional magnetic resonance imaging) suggest that there are localized patches within TE specialized for analyzing category-like stimuli, e.g., face and nonface patches in area TE (Sugase et al., 1999; Sigala and Logothetis, 2002; Tsao et al., 2003; Kiani et al., 2007; Bell et al., 2009; Sugase-Miyamoto et al., 2014). Additionally, the microstimulation of face patches has been shown to influence face categorization (Afraz et al., 2006). Finally, monkeys with TE lesions are impaired in generalizing across new views of objects (Weiskrantz and Saunders, 1984) and in remembering eight previously unseen objects concurrently (Iwai and Mishkin, 1968; Buffalo et al., 1999).

Here we test visual categorization in monkeys before and after TE removals. In the first experiment, monkeys were tested using an implicit measure of categorization (Task 1). To our surprise, TE removals produced only a minor impairment. We performed a second experiment (Task 2), in which the subjects made an explicit choice between categories. Again, there was only a mild impairment. Given the mild impairment, it seemed possible that the impairment might be explained by a deficit in visual acuity or procedural learning, but, as Task 3 demonstrates, both were intact. Finally, the question arose, would our TE lesions produce similar impairments in visual discrimination learning seen classically? In Task 4, the TE-lesioned monkeys were tested on a series of visual discriminations, which showed a pattern of performance that was similar to the deficits reported in past studies. It appears that our lesions, and resulting visual behavioral deficits, match those reported in the past.

Materials and Methods

Subjects and surgeries
All experimental procedures conformed to the *Institute of Medicine Guide for the Care and Use of Laboratory Animals* and were performed under an Animal Study Proposal approved by the Animal Care and Use Committee of the National Institute of Mental Health. For Tasks 1 and 4, we used three experimentally naive male rhesus monkeys (*Macaca mulatta*; weight, 3.9–5.75 kg). For Tasks 2 and 3, we added three unoperated control (normal) male monkeys.

Surgeries to remove TE bilaterally using aspiration were performed in a veterinary operating facility using aseptic procedures. Anesthesia was induced with ketamine hydrochloride (10 mg/kg, i.m.) and maintained using isoflurane gas (1.0–3.0% to effect). Vital signs were monitored
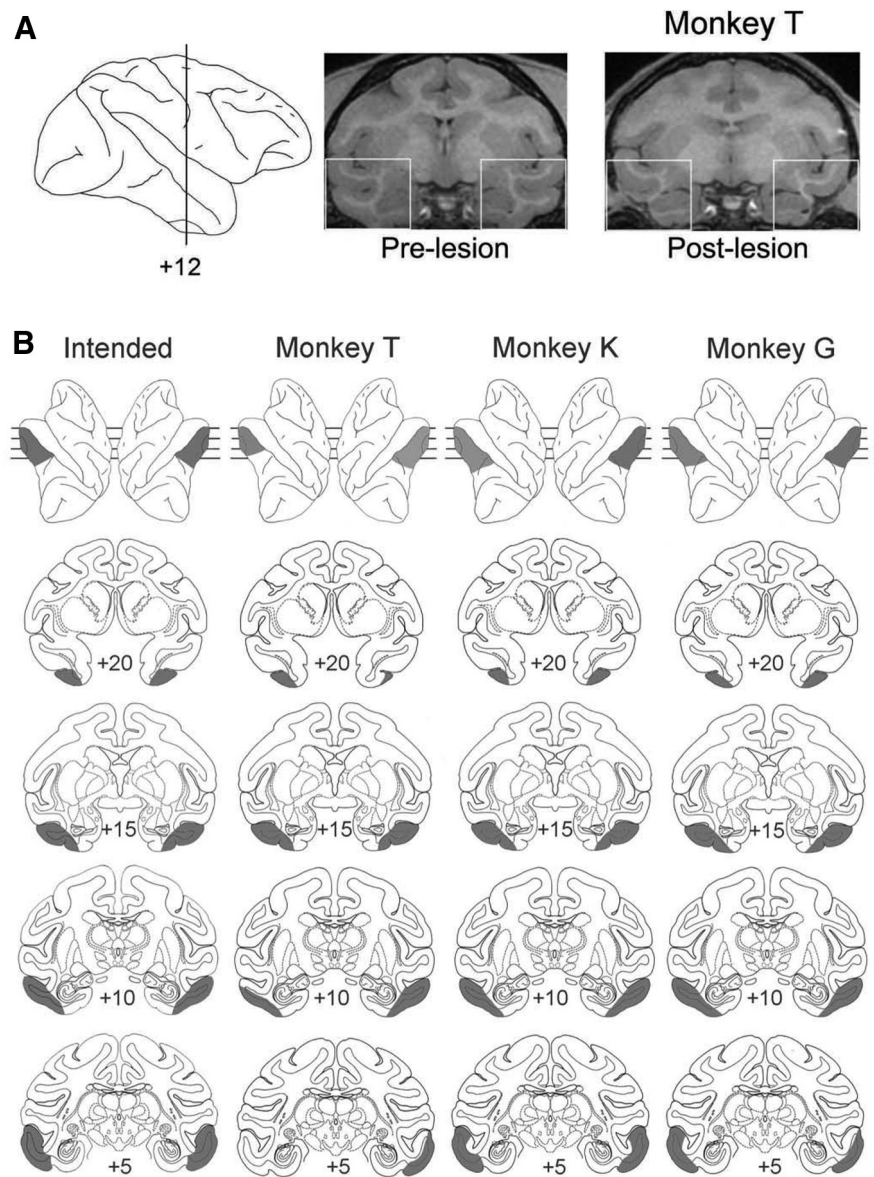
Figure 1. *A*, MR images of prelesion and postlesion brain (white rectangles, right) at the vertical line on a surface view (left). *B*, Lesion reconstructions for the three monkeys with TE removals.

throughout the surgeries. The extents of the lesions were reconstructed from magnetic resonance (MR) scans (Fig. 1A). The intended area TE lesion extended from 1 cm from the anterior limit of the inferior occipital sulcus rostrally to the temporal pole, and from the fundus of the superior temporal sulcus around to the lateral bank of the occipitotemporal sulcus. The ventral–medial border extended rostrally from the occipitotemporal sulcus ending midway between the anterior middle temporal sulcus and the rhinal sulcus (Fig. 1B). The area TE removals were largely as intended.

Behavior
An overview is provided in Table 1.

Initial training. Each monkey was seated in a primate chair in a sound-attenuated dark room facing a computer monitor. Stationary random black and white dots ("noise") were displayed on the monitor (800 × 600 pixels resolution) in front of the monkey as a visual background. Monkeys were not head-fixed, so eye position data were not recorded. The monkeys were trained to grasp and release an immovable, touch-sensitive bar in the chair to get a liquid reward. Then, they were trained to perform a sequential, red–green color discrimination (Fig. 2A). A trial began when the monkey grabbed the touch-sensitive immovable bar. A

Table 1. Behavior overview

TE status	Task	Stimulus set	Number of sessions
Presurgery	Procedural training	Red + green cues	Until >80% accuracy for 3 d
Presurgery	Precategory training	2 black and white cues	Minimum = 6
Presurgery	Dog/cat discrimination	40 training cues	4
Presurgery	Human/monkey discrimination	40 training cues	4
Postsurgery (TE aspiration lesion)	Dog/cat discrimination	40 training cues	4
Postsurgery	Dog/cat categorization	480 novel cues	1
Postsurgery	Human/monkey discrimination	40 training cues	4
Postsurgery	Human/monkey categorization	480 novel cues	1
Postsurgery	Car/truck discrimination	40 training cues	4
Postsurgery	Car/truck categorization	480 novel cues	1
Postsurgery	WGTA: discrimination	Plus and square cards	Until >90% accuracy for 3 d
Postsurgery	WGTA: discrimination	3D objects	Until >90% accuracy for 3 d

black and white image (within a 13° square region) appeared immediately, at the center of the monitor; 0.4 s later, a red target (0.5° on each side) appeared at the center on the image. After another 0.5–1.5 s, the target changed to green. If the monkey released the bar within 3 s after the green target appeared (correct response), the target changed to blue. At this point a drop of juice was delivered after every correct response, i.e., after every release in response to the green target. After each correct trial, there was a 2 s intertrial interval before the monkey could initiate a new trial. If the monkey released the bar before the green target appeared or after the green target had disappeared, the trial was scored as an error, the cue disappeared, and a 4 s "time-out" period was imposed.

Task 1: implicit categorization task. When the monkeys reached 80% correct in the red–green color discrimination for three consecutive daily training sessions, trials were divided into two equal groups, of high-incentive (3 drops of liquid) and low-incentive (1 drop of liquid) trials. At this point, the image that had been the same in every trial was replaced with one of two images. For training, the two images were two black and white block patterns, 13° on a side. These images were incentive-related cues; that is, one was presented only in high-incentive trials, and the other was presented only in low-incentive trials. The task requirement did not change, i.e., release the bar when the target was green. We know from previous work that in low-incentive trials the monkeys increase the number of trials in which they release either before the green target appears or after it disappears; that is, the number of "errors" increases in low-incentive trials (Minamimoto et al., 2010). After any error, either in a high-incentive or low-incentive trial, the same visual cue appeared in the next trial after a 4 s time-out period. Our dependent measure for assessing the effect of the cues was the number of errors. The monkeys practiced with the two black and white cues for 6, 10, and 27 sessions (Monkeys T, K, and G, respectively), a period over which the monkeys developed a stable difference in behavior between the high- and low-incentive conditions.

After the monkeys became acclimated to the incentive difference between the two black and white cues, these two cues were replaced with two sets of cues, i.e., high-incentive (1 drop of liquid) and low-incentive (no drop) sets (Fig. 2A), where the incentive cue in each trial came from one of two categories, such as dogs or cats (Fig. 2B). If the monkeys distinguished between categories, we expected fewer errors for the category associated with the high incentive and more errors for the category associated with the low incentive. Again, the dependent measure was the number of errors, but now for category one versus category two. The monkeys were tested preoperatively in Task 1 for

four sessions using the 20 dog/20 cat training set, and four sessions with the human/monkey training set.

Task 2: explicit categorical choice task. We transformed the task into an explicit asymmetrically rewarded forced choice task. Each trial began as above with presentation of the incentive cue, followed by the appearance of the red target. The choice was indicated by releasing during presence of the red target for one category, e.g., a cat, and during presence of the green target for the other, e.g., a dog. If the monkey released while the red target was present for either cue, the trial was ended and a new cue was picked at random. If the monkey released during the green target when the high-incentive cue, a dog, was present, the monkeys received a 1 drop liquid reward (Fig. 2A). If the monkey released while the green target was present when the low-incentive cue, a cat, was present, no reward was delivered and there was a 4–6 s time-out. There was never a reward for releasing while the red target was present. The optimal behavior is to release during the presentation of the red target for the low-incentive cue, essentially skipping on to the next randomly selected trial, and release during the presentation of the green target for the high-incentive cue to obtain the reward. In this task the monkey is not making a simple red–green judgment. Rather, this design is a two-interval forced choice (2-IFC) task, with asymmetrical reward, in which the color of the central target indicates the current "choice window." The behavior that maximizes reward rate at a given point in the trial is determined by the categorical cue. For training, two black and white block pattern cues were used (different from those used in the initial training described above). As is shown in Results, the monkeys mastered this task quickly. The performance of three monkeys with TE removals was compared with that of a group of three intact control monkeys in this task. The red target was presented for 2–3 s to provide ample opportunity for the monkeys to release on red.

Experimental control and data acquisition was done using a REX behavioral paradigm (Hays et al., 1982). The visual stimuli, targets and cues, were presented with Presentation (Neurobehavioral Systems), where stimuli and timing were specified from REX.

Visual cues for categorization (*Tasks 1 and 2*). For the computer-controlled Tasks 1 and 2, all visual cues were jpeg or pcx format photos (200 × 200 pixels). The cue sets of dogs/cats and cars/trucks used in this study are the same as in our previous report (Minamimoto et al., 2010). The training set for the cat/dog category and 40 of the 480 trial-unique exemplars are shown in Figure 2B. There was a mixture of sizes and views for the dogs, cats, cars, and trucks. The 20 human faces were obtained at http://www.computervisiononline.com/dataset/georgia-tech-face-database (Nefian and Hayes, 2000). The members of the training set of 20 human faces were in frontal view and neutral expression. The 240 trial-unique human faces were obtained at http://vis-www.cs.umass.edu/lfw/ (Huang et al., 2007). This set includes different views and expressions. The training set of 20 monkey faces and 240 trial-unique monkey faces were obtained by Dr. K. Gothard (The University of Arizona College of Medicine, Tucson, AZ) (Gothard et al., 2004). Some of the trial-unique monkey faces were obtained from the Internet. All of the 20 training set monkey faces were frontal view and with a neutral expression. The members of the set of 240 trial-unique monkey faces include different views and expressions. The backgrounds of all images in all sets were erased by hand using Adobe Photoshop. The monkeys were not shown any image from these sets before the behavioral testing.

Task 3: test of visual acuity. To be sure that the mild deficits in categorization observed in Tasks 1 and 2 did not result from impaired visual acuity, the acuity of control and TE groups was assessed by presenting full-screen sine wave gratings [i.e., the local intensity was modulated by a one-dimensional (vertical) sine wave across the screen] that covered a range of frequencies (16, 8, 4, 2, 1, 0.5, 0.25, and 0.125 cycles/degree) and contrasts (1, 0.64, 0.32, 0.16, 0.8, 0.4, 0.2, 0.1, 0.05). Contrast was calculated as: $L_P - L_T/(L_P + L_T)$, where L_P represents peak luminance and L_T trough luminance. The space-average luminance was kept constant across stimuli. The task took the form of a signal detection paradigm, whereby the monkey was required to release the lever immediately if a grating was detected (during the presentation of the red target) to obtain a reward, or otherwise to continue to hold the lever until the target turned

green, and then to release the lever to obtain a reward. This is a two-interval forced choice task, similar to Task 2, but with symmetrical reward value (the absolute size of the reward available for correct releases during the second response window was larger than that for the first to take into account the delay discounting function of the monkey, and hence to balance reward value between both response intervals). Thus, the absence of an impairment in this task would indicate not only that the visual acuity of the monkeys with TE lesions is intact, but also that they have no difficulty learning the incentive values for Task 2. Gratings were presented for 500 ms on 50% of trials. If the monkey released the lever during the presence of the red target when no grating had been presented or released on green when a grating had been presented (i.e., both incorrect responses), a 4−6 s time-out was incurred.

Task 4: Wisconsin General Testing Apparatus discrimination learning task. For the purpose of comparing the effects of our bilateral TE ablations on discrimination behavior with those historically reported, classical visual discrimination tasks were performed in a standard Wisconsin General Testing Apparatus (WGTA). Initially the monkeys were trained to obtain food from two food wells. Then, they were trained to discriminate between two patterns (plus and square) presented on paper cards. If they displaced the card with the plus sign, they obtained a banana-flavored pellet food reward. Selection of the square was not rewarded. The two cards were renewed frequently, and their left/right placement over the food wells was changed pseudorandomly on successive trials. Monkeys were trained for 20 trials per session, five sessions per week, with an intertrial interval of 20 s, until they attained a criterion performance of 90% or better in each of three consecutive sessions. Once they learned the pattern discrimination they were trained on two additional discrimination paradigms. The second discrimination was between two three-dimensional (3D) junk objects, and, as with the pattern discrimination, one stimulus was always rewarded after a correct choice and the other not. All other parameters were the same as for the plus/square discrimination. The third test was a concurrent learning paradigm in which the animals had to learn to discriminate between eight pairs of 3D objects, the presentation of which was interleaved. All parameters were the same as for the first two tasks.

Data analysis

Most data analyses were performed using standard tests included in the MATLAB Statistical Toolbox (MathWorks). For some behavioral conditions, where appropriate, we used a generalized linear mixed model (GLMM) from the "lme4" library in R (R Development Core Team, 2004). The generalized linear modeling test was applied to behavioral data that contained repeated measures within subjects (fixed effects) and
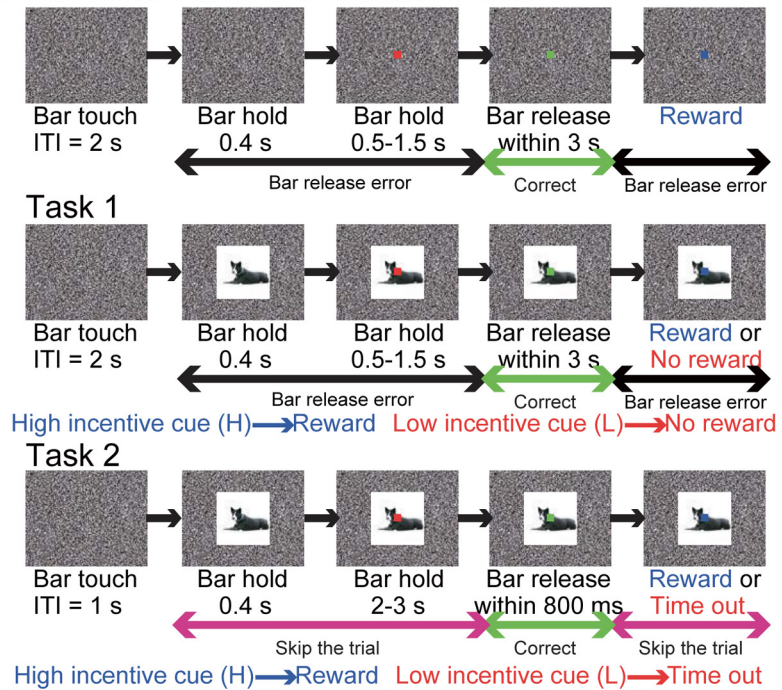
Figure 2. *A*, Sketch of the sequence of events during one trial. *B*, Examples of visual cue sets. Top, Twenty dogs and 20 cats. Bottom: Twenty dogs and 20 cats out of 240 trial-unique dogs and 240 trial-unique cats. ITI, Intertrial interval.
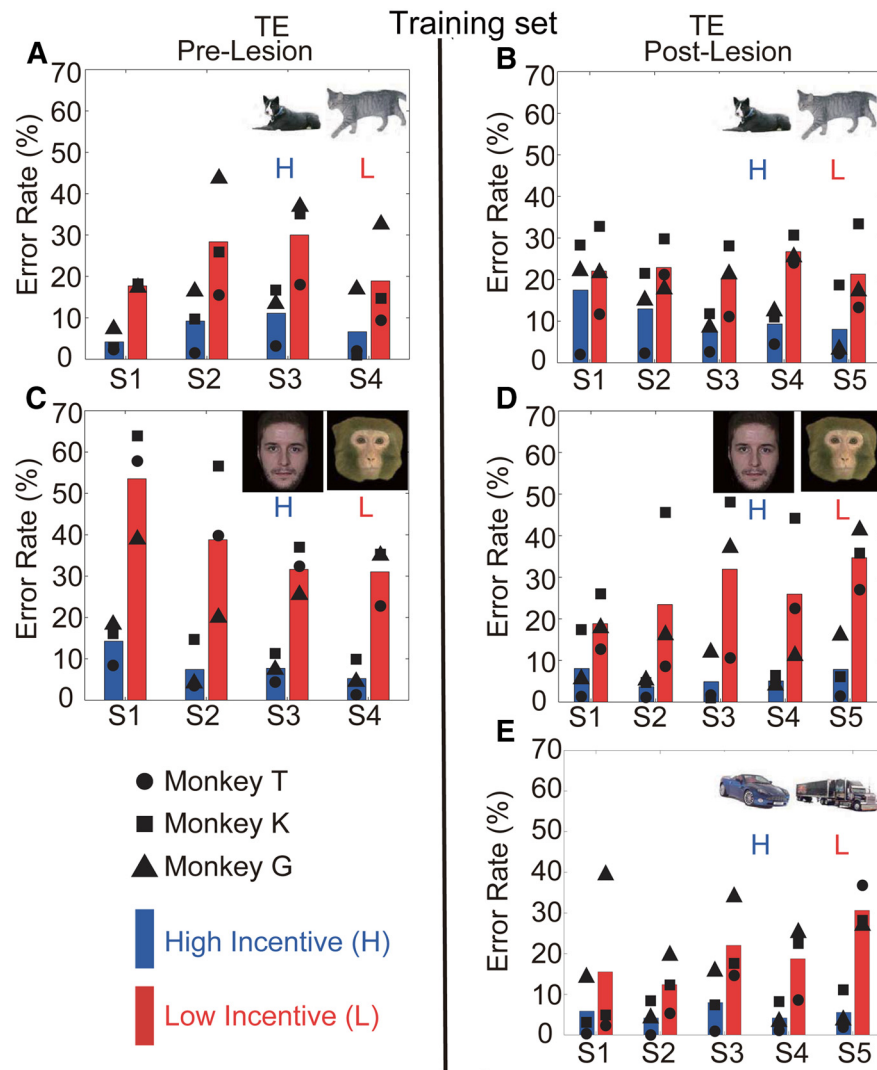
Figure 3. Error rates in Task 1 for high- (H) and low-incentive (L) cues in each session. The inset images indicate which cue sets were used in each session. Blue and red bars represent mean error rates for the high and low-incentive cues, respectively. Symbols indicate error rates for individual monkeys. **A, C**, Results for the prelesion tests. **B, D, E**, Results for the postlesion tests.

Results

Task 1: implicit task

The monkeys were tested preoperatively in Task 1 for four sessions using the 20 dog/20 cat training set and four sessions with the human/monkey training set. As has been demonstrated previously, the behavioral performance of normal monkeys is related to the categorical cue (Minamimoto et al., 2010). Before the TE removals, all three monkeys made significantly more errors on low-incentive trials than high-incentive trials in dog/cat and human face/monkey face categorization (Fig. 3A,C). The difference in error rates between categories was significant for all monkeys, both when discriminating between cats and dogs (χ^2 test, $p < 0.05$; $p = 1.0 \times 10^{-10}$, $p = 4.0 \times 10^{-10}$, and $p = 6.5 \times 10^{-5}$ for Monkeys T, K, and G, respectively) and when discriminating between human and monkey faces ($p < 0.05$; $p = 0.0$, $p = 0.0$, and $p = 0.2 \times 10^{-5}$ for Monkeys T, K, and G, respectively).

Postoperatively, the monkeys were tested for four sessions using the same dog/cat and human/monkey face cue sets used preoperatively. On the dog/cat cue set, one monkey (T) made significantly more errors on low-incentive trials (χ^2 test, $p < 0.05$; $p = 1.8 \times 10^{-7}$) in the first postoperative session, indicating intact categorical discrimination. Monkey K made significantly more errors on low-incentive trials beginning in the second session ($p < 0.05$; $p = 0.011$), and Monkey G made significantly more errors on low-incentive trials beginning in the third postoperative session ($p < 0.05$; $p = 6.6 \times 10^{-5}$). Thus, by the end of three postoperative sessions, all three monkeys had clearly relearned/recalled the discrimination between dogs and cats (Fig. 3B). For Set 2, the human/monkey faces, all three monkeys demonstrated differences in error rates between high-incentive (human faces) and low-incentive (monkey faces) trials beginning in the first postoperative testing session ($p < 0.05$; $p = 2.6 \times 10^{-8}$, $p = 5.5 \times 10^{-5}$, $p = 1.1 \times 10^{-5}$ for Monkeys T, K, and G, respectively; Fig. 3D).

When we compared the performance across the four preoperative testing sessions to the that across the four postoperative testing sessions for the dog/cat categorization using a GLMM to model the binomial variance structure (see Materials and Methods), there were large significant effects of dog versus cat, a small but significant difference due to the lesion on overall performance (GLMM, $z = -2.51$, group and monkey as random effects, $p = 0.012$), and a significant interaction between category and lesion ($z = 6.22$, $p = 4.8 \times 10^{-11}$). There was no interaction between the lesion and the face categories ($z = 1.70$, $p = 0.089$). Thus, there appears to be a mild impairment for the dog/cat set after the ablations, but not the human/monkey set.

across subjects and/or groups (random effects) with dependent measures that did not follow a normal distribution (yes–no data such as our correct vs error trials), in our case a binomial distribution (Longford, 1993). We compared both random intercept (constrains covariance between any pair of repeated measurements to be equal) and random slope and intercept (includes covariance between repeated measurements as an additional parameter) models using ANOVA, choosing the random slope and intercept model only if it provided a significantly better fit.

Lesion localization

The size and location of lesions were confirmed by T1-weighted MR images collected during scans on a 1.5-T General Electric Signa unit, using a 5 inch surface coil and a three-dimensional volume spoiled gradient recalled acquisition in steady state pulse sequence. Slices were taken every 1 mm with in-plane resolution of 0.4 mm. Coronal slices from the MR images were matched to plates in a stereotaxic rhesus monkey brain atlas. The extent of the visible lesion was plotted on each plate and subsequently reconstructed (Fig. 1B), showing that area TE was largely removed in all three animals. The monkeys were used in further experiments, so histological analysis is not available.

To examine the possibility that the monkeys were memorizing some of the visual cues, since the same 40 stimuli within a set were used across all four preoperative and four postoperative sessions, we examined the error distribution to determine whether the errors were concentrated in relation to a few stimuli. The errors for all three monkeys were distributed across all of the visual cues (shown in stacked bars for individual cues in each session in Fig. 4A–D), suggesting monkeys did not perform the task by remembering just a few the visual cues.

A better test to demonstrate category generalization is to use never-before-seen exemplars. We used 240 new exemplars each of dogs and cats with many different views (Fig. 2B, 20 examples each of dogs and cats). These trial-unique stimuli were taken from a variety of sources, with image sizes adjusted to cover the same height and width, and put onto a uniform white background. The images included different views and a wide set of colors that strongly overlapped across the dogs and cats so that it is unlikely that any simple feature, such as color, view, or size, distinguished the cats from dogs (for further analysis supporting this conclusion, see results from the linear classifier in "Image analysis," below). Each stimulus of this trial-unique set was presented once in a single testing session of 480 trials. The group mean error rates and individual cumulative error counts for the high- versus low-incentive cues for each monkey were significantly different (χ^2 test, $p < 0.05$ for Monkeys T, K, and G; Fig. 5A,B).

We then tested the monkeys with 480 trial-unique exemplars from our sets of monkey/human faces. The group mean error rates and individual cumulative error counts for the high- versus low-incentive cues were significantly different (χ^2 test, $p < 0.05$ for Monkeys T, K, and G; Fig. 5C,D, left, middle, right columns). However, unlike the cat/dog set, it seems plausible that the monkey/human face distinction could be achieved by looking for simple, low-level characteristics such as facial hair or overall reflectance. Our results are consistent with the view that area TE is unlikely to be required for discriminations requiring such basic comparisons.

At this point it appeared that the monkeys with the TE ablations were only mildly affected in the retention of the dog/cat–reward associations, and not affected for the faces. To test whether the monkeys would be able to learn a new pair of categories after the ablations, a new stimulus set of 20 cars (low incentive) and 20 trucks (high incentive) was introduced. Monkeys T and G committed significantly more errors for cars in Session 1 (χ^2 test, $p < 0.05$; $p = 0.02$ and $p = 3.6 \times 10^{-11}$, respectively), and Monkey K achieved significant discrimination in Session 3 ($p < 0.05$; $p = 1.1 \times 10^{-3}$; Fig. 3E). Again, the errors were spread across the whole stimulus set (Fig. 4E). When the ability of the monkeys to generalize was tested with 240 trial-unique exemplars each of cars and trucks, the group mean error rates and individual
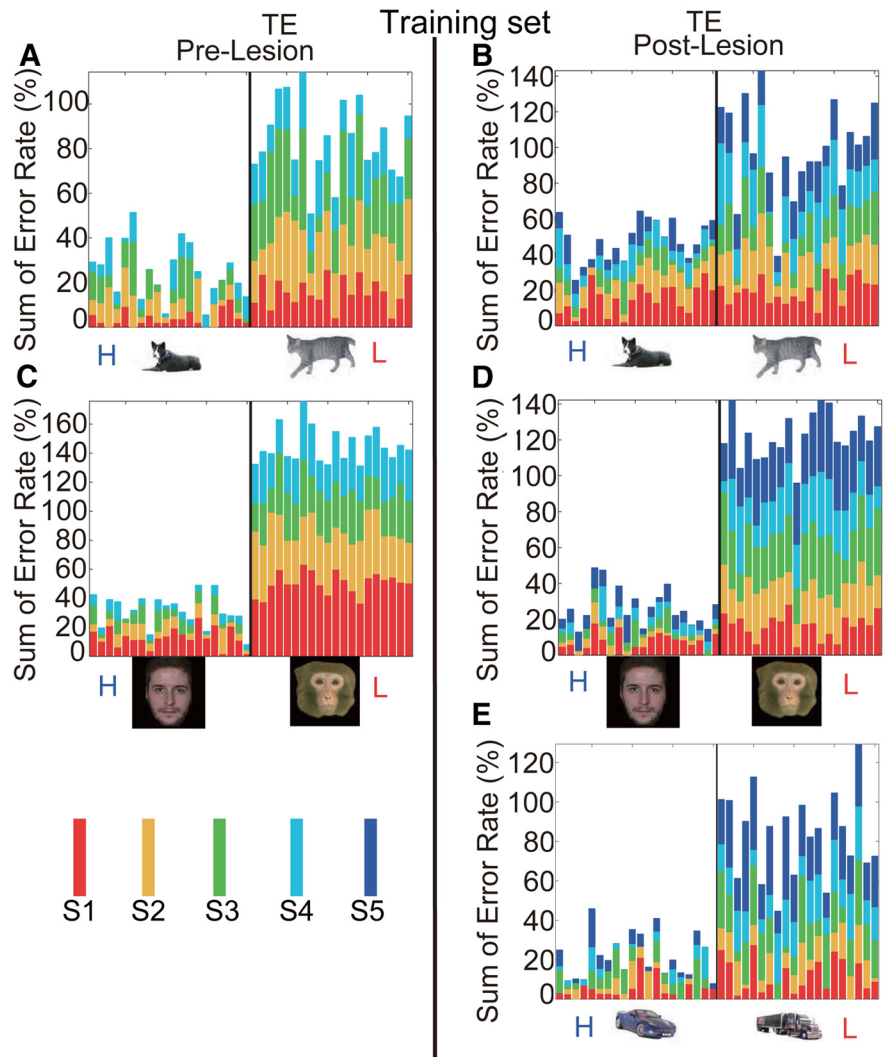

Figure 4. Summation of error rates for individual stimuli in Task 1. **A, C,** Results for four sessions of prelesion tests (each represented by a different color in the stacked bar). **B, D, E,** Results for five sessions of postlesion tests. The thin black vertical line delineates the categorical boundary between high- (H) and low-incentive (L) cues. Mean error rates of the three monkeys in sessions were stacked.

cumulative error counts for the high- versus low-incentive cues, for each monkey, were significantly different ($p < 0.05$, $p = 3.2 \times 10^{-12}$, $p = 2.6 \times 10^{-11}$, and $p = 1.5 \times 10^{-3}$ for Monkeys T, K, and G; Fig. 5E,F, left, middle, right columns, respectively). Thus, not only do these monkeys generalize to categories learned before their TE ablation, but they also quickly learn to distinguish between members of a new pair of categories.

Task 2: explicit choice task
The results above demonstrate that the monkeys with TE ablations are sensitive to the categorical differences in the cue sets; for example, they make more errors in the unfavorable trials as signaled by the appearance of a cat than in the favorable ones, even though the most advantageous behavior would be to perform all of the trials correctly. The categorization task (Task 1) used above did not require the monkeys to make an explicit choice; we only observed their errors in the face of greater and lower incentives. If categorization for implicitly learned discriminations were to occur using different neuroanatomical substrates than for explicitly learned tasks, or if the option of aborting trials with undesirable outcomes increased attention and/or the motivation of the mon-
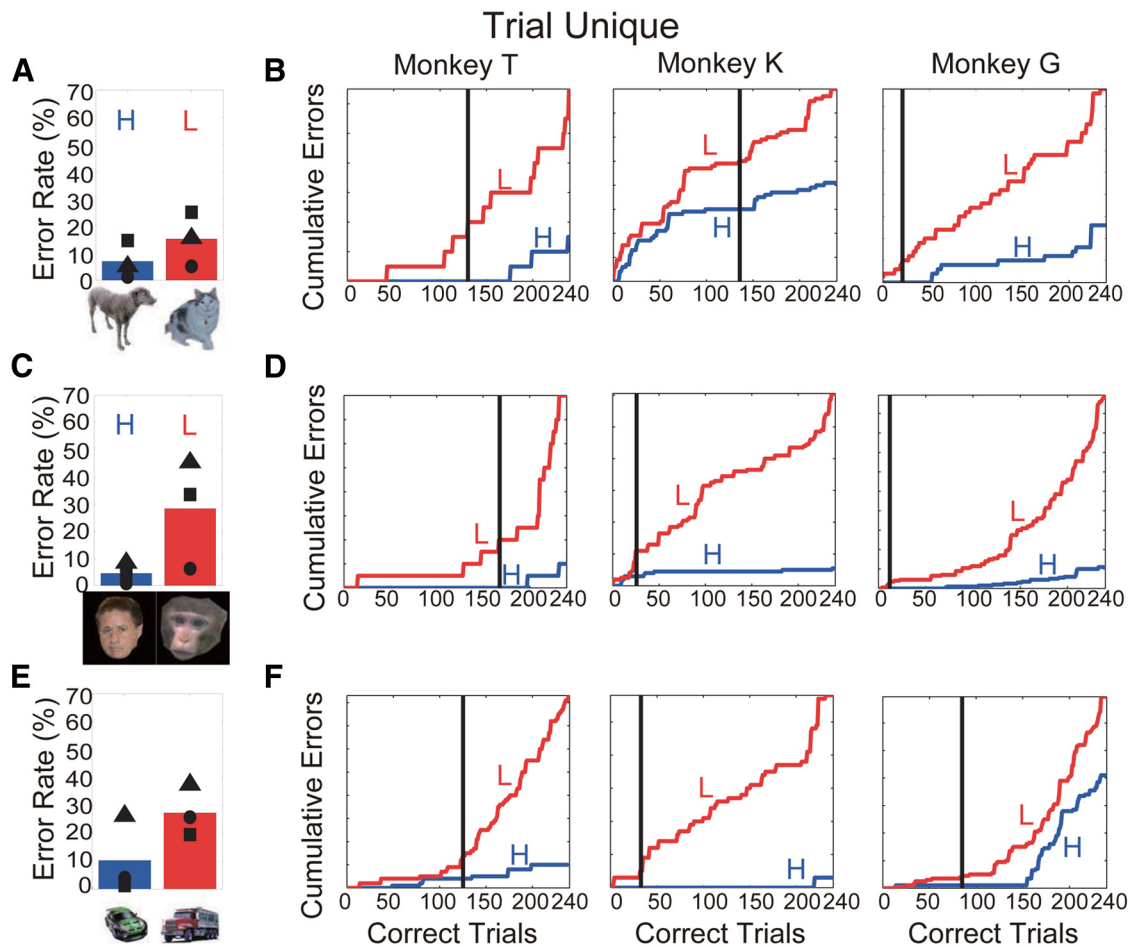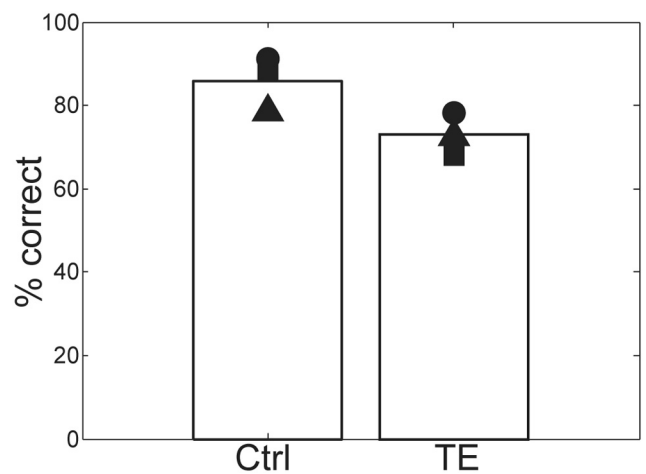
Trial Unique

Figure 5. Error rates and cumulative error counts for high- (H) and low-incentive (L) trial-unique cues in Task 1. ***A, C, E***, Error rates for the high- and low-incentive cues. The meanings of bars, symbols, and images are the same as in Figure 3. ***B, D, F***, Cumulative error counts for the high- and low-incentive cues. The scale of cumulative error counts change substantially. The figures in the left, middle, and right columns show the results for Monkeys T, K, and G, respectively. The vertical lines show the trial at which the difference in cumulative errors for the high- and low-incentive cues becomes significant ($p < 0.05$ by χ^2 test for 5 consecutive trials).

keys to discriminate, testing the monkeys in an explicit task might yield a different result.

We transformed our implicit task into an explicit 2-IFC task with asymmetrical reward, comparable to those used throughout the psychophysical literature (Green 1964; Farah 1985; Magnussen et al., 1991; Yeshurun et al., 2008). The two alternative choices were, (1) if a cat appeared release during the red target (Interval 1), to terminate the current trial and move on to the next, and (2) if a dog appeared release during the green target (Interval 2), to collect the reward. In the simplest terms, that is, the task was to release on red for cat and on green for dog.

Within seven sessions, all six monkeys (three with TE removals, three unoperated controls) learned that the most advantageous response pattern was to release on red when the low-incentive cue appeared and to wait and release on green to collect the liquid reward for the high-incentive cue. Two TE monkeys and one control monkey learned this in four sessions, the other TE monkey in five sessions, and the other two control monkeys in seven sessions (the last two control monkeys were naive, never having worked with cues before). The monkeys were then trained on the set of 20 dogs and 20 cats for four consecutive sessions before being exposed to sets of trial-unique cats and dogs, with 240 cats and 240 dogs each session for four consecutive sessions (total, 960 cats, 960 dogs). For the control monkeys, the average performance was high: 85% correct response (Fig. 6). For the TE

Figure 6. Correct rates in Task 2: explicit choice task for trial-unique dogs/cats on four consecutive sessions for control (Ctrl) and TE-lesioned (TE) monkeys. Symbols and bars indicate correct rates for individual monkeys and their averages, respectively.

monkeys, the average performance was significantly lower: 73% correct (Fig. 6; GLMM, group and monkey as random effects, $z = -13.36$, $p < 2.0 \times 10^{-16}$). Thus, when compared to controls in this sequential response task, the TE monkeys show a mild deficit

in assigning the pictures correctly to the categories. Nonetheless, they generalize; that is, they discriminate at well above chance levels (50%) between images from different categories.

Task 3: contrast sensitivity test
To control for the possibility that any impairment in the TE-lesioned monkeys' performance on the categorization and discrimination learning tasks could be accounted for by a simple loss of ability to perceive fine-grained visual attributes, visual acuity was measured in the form of contrast sensitivity across a range of sine wave grating frequencies and compared to that of controls. Consistent with the human literature (Blakemore and Campbell, 1969), control monkeys had difficulty detecting the very lowest and several of the highest frequencies (Fig. 7); that is, the monkeys could detect only the higher contrast gratings at these frequencies. The visual acuity of TE-lesioned monkeys was indistinguishable from that of controls (ANOVA, group, $p > 0.05$; group by frequency, $p > 0.05$).

Task 4: Wisconsin General Testing Apparatus discrimination learning task
The postoperative ability to learn a new categorization rapidly and the ability to generalize all the categorizations on which we tested the monkeys was especially surprising to us, given that TE is thought to play a critical role in discrimination and memory for complex visual stimuli. One explanation for this would be if our ablations did not match those reported in the past where pattern discrimination was severely impaired after bilateral TE removals (Iwai and Mishkin, 1968; Buffalo et al., 1999). Therefore, we tested our three monkeys with TE removals using three additional visual discrimination tasks traditionally used to characterize this type of deficit (Iwai and Mishkin, 1968; Cowey and Gross, 1970). This testing was performed in a Wisconsin General Testing Apparatus. Our monkeys with TE lesions showed a pattern of behavior across the three discrimination tasks that was consistent with the pattern reported previously (Iwai and Mishkin, 1968; Buffalo et al., 1999): the monkeys had difficulty in learning a single-pair pattern discrimination (plus vs square) and a concurrent eight-object discrimination, compared with learning a single-pair, three-dimensional object discrimination (Fig. 8).

Image analysis
The relative difficulty of each pair of category discriminations was assessed by training a linear support vector machine (SVM) classifier on 90% of the stimuli of the trial-unique sets from each category pair and using the outcome to perform a linear classification of the remaining 10% of trial-unique stimuli (MATLAB, svmtrain and svmclassify). Stimulus information was passed to the SVM as a pixelwise decomposition in RGB (red–green–blue) color space. The random sampling cross-validation was repeated
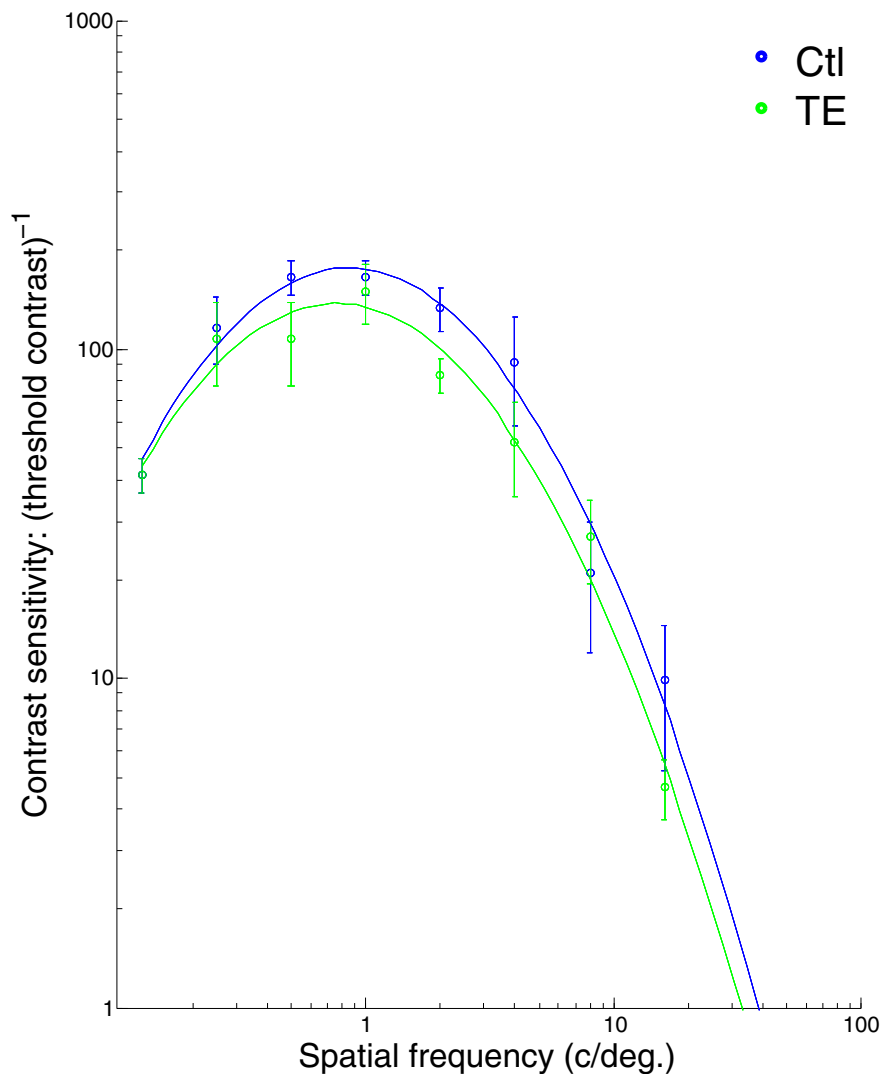

Figure 7. The contrast sensitivity functions for control (Ctl) and TE-lesioned (TE) monkeys. Contrast sensitivity is plotted on a logarithmic scale against spatial frequency. The fitted curves are quadratic functions.

with 100 random splits to permit all stimuli to be used in the training and test sets multiple times. The resulting linear classifiers categorized the cat/dog stimuli with a mean accuracy of 65% (95th quantile, 75%), car/truck stimuli with a mean accuracy of 74% (95th quantile, 85%), and human/monkey faces with an accuracy of 96% (95th quantile, 100%). Comparing the above results to the classification performance of monkeys in Task 2, it is apparent that normal monkeys perform considerably better than the SVM, and that monkeys with TE lesions perform at the upper limit of what might be expected from the classifier.

Discussion
To arrive at an explanation of how categorization based on visual similarity occurs in the brain requires first the identification of brain regions that are necessary for this process. As a conjecture, the region or regions that are essential for categorization should receive high-level visual input and be involved in visual memory. The prefrontal cortex is a candidate region for categorization (Freedman et al., 2001), yet monkeys with large lesions of lateral prefrontal cortex are not impaired in generalizing preoperatively learned categories, or in acquiring and generalizing postoperatively encountered categories (Minamimoto et al., 2010). Area
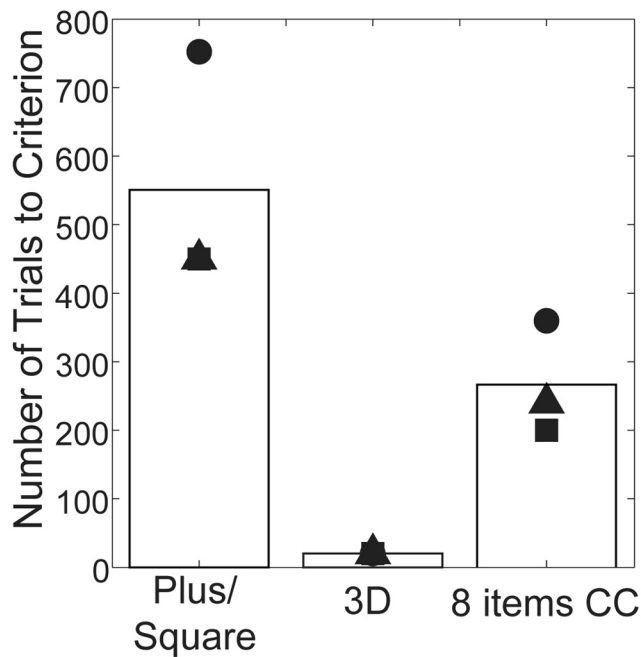
Figure 8. Number of trials to criterion in Task 4: WGTA tests. Symbols and bars indicate the numbers of trials for individual monkeys and their averages, respectively.

TE, a lateral temporal brain region, has also been proposed as a candidate region for categorization because it plays a central role in high-level visual processing, object recognition, and object discrimination (Iwai and Mishkin, 1968; Cowey and Gross, 1970; Gross et al., 1972; Desimone et al., 1984; Weiskrantz and Saunders, 1984; Tanaka, 1996; Sugase et al., 1999; Sigala and Logothetis, 2002; Tsao et al., 2003; Afraz et al., 2006; Kiani et al., 2007; Bell et al., 2009; Sugase-Miyamoto et al., 2014). The results presented here, however, seem to, at least partially, refute this prediction.

In Tasks 1 and 2, monkeys showed only a mild impairment in visual perceptual categorization after TE removal, rather than a severe impairment as might be expected. As the postoperative test with the car/truck set demonstrates, the monkeys were able to learn new categories as quickly as they had preoperatively. Nevertheless, these monkeys were severely impaired in pattern discrimination learning and eight-item concurrent learning, as has been reported for monkeys with TE lesions (Iwai and Mishkin, 1968; Buffalo et al., 1999). Thus, our monkeys with TE lesions have the same pattern of visual learning deficits that have been classically reported for monkeys with TE lesions.

The contrast sensitivity analysis serves a twofold purpose as a control for the explicit test of categorization: it indicates, first, that the visual acuity of the TE-lesioned monkeys is unimpaired, and second, that their ability to grasp the procedural components of the explicit task (Task 2) is intact, including high-incentive versus low-incentive contingencies and the two-interval forced choice design. Thus, what we have found is that monkeys with bilateral TE lesions are, in our view, surprisingly good at visual perceptual categorization (training task, 20 high-incentive/20 low-incentive cues) and that they retain the ability to generalize and categorize novel items (trial-unique stimuli). That the results for the explicit task (Task 2) parallel those for the implicit task suggests that the mild deficit observed after TE lesions in the implicit task is unlikely to arise from a reduction in attention or motivation, or from a general ceiling effect on the error rate as a result of the lack of an explicit choice option in the implicit task.

One important issue for interpreting these results is whether the monkeys learned to use some single key feature to perform this task. Our cats, dogs, cars, and trucks came from different sources on the Web, included a wide variety of colors and different color combinations, were of different views (side views, head-on views), and different amounts of stimuli were visible [from close-ups restricting the view to a face (cats/dogs) or just the front (cars/trucks) to long lateral views; Fig. 2B]. Because these images comprise such a broad, and overlapping, array of parameters, it seems unlikely that any simple feature could be used to distinguish the cats from dogs. To support this qualitative conclusion, we performed an analysis using a linear SVM classifier to discriminate the dogs and cats. Normal monkeys performed considerably better than a powerful linear classifier (control monkeys attained 85% classification accuracy, the SVM classification accuracy at the 95th quantile = 75%), and the monkeys with complete TE removals performed at the upper limit of what might be expected from the classifier (TE-lesion monkeys, 73% classification accuracy). This supports our supposition that discrimination is not based on a "simple" feature.

Although it is widely acknowledged that feedback and other interactions are important for object-based visual recognition, it would probably still be assumed that a system with a feedforward circuit leading into area TE and then perhaps to frontal cortex has a central role in the type of visual perceptual categorization studied here (Kravitz et al., 2013). Even tasks with fixed reward contingencies, such as those reported here, would be expected to recruit TE if the perceptual discriminations required were sufficiently challenging. The effects of aspiration lesions must always be interpreted in light of the possibility of damage to underlying fibers of passage. In this study, however, the main finding is the relative ineffectiveness of the TE removal for impairing visual generalization; thus, potential damage to fibers of passage does not represent a confound to the interpretation of these results.

There are at least four explanations for why lesions of lateral prefrontal cortex (Minamimoto et al., 2010) or TE do not devastate visual perceptual categorization. First, categorization might be a distributed function, so there is no single critical area. In that case, there should be no single lesion that catastrophically disrupts visual perceptual categorization. Second, if there is a single critical brain area, it has not been yet identified. Candidates that might yet be consistent with the hierarchical model are the perirhinal cortex, which follows TE in the visual processing hierarchy and is important for stimulus–stimulus and stimulus–reward associations (Liu et al., 2004; Murray et al., 2007), generalization to new views (Weiskrantz and Saunders, 1984; Buckley and Gaffan 2006), and visual discriminations between stimuli with a high level of feature overlap (Bussey et al., 2003); area TEO, which precedes TE and is known to contribute to pattern discrimination (Kobatake and Tanaka, 1994); and amygdala, which has direct inputs from TEO (Webster et al., 1991; Stefanacci and Amaral, 2002) and is known to carry neural representations of visual stimulus–reward associations and abstract categories (such as different classes of facial expression; Mosher et al., 2010). The difficulty with invoking a vital role for perirhinal cortex is that our TE lesions should have deafferented perirhinal cortex from nearly all of its primary visual input, although it is possible that sufficient information arrives from other areas with major projections to perirhinal cortex, such as TEO and TF (Suzuki and Amaral, 1994). Third, image analysis within the brain for visual pattern generalization might occur before fine pattern analysis, a result hinted at in single neuronal physiology, where coarse pattern information appears earlier in responses of TE neurons than

the information used for fine discrimations (Sugase et al., 1999; Matsumoto et al., 2005). Finally, it might be that the image analysis used for categorization arises in parallel with and separately from the image analysis that underlies the final steps of pattern discrimination, which would provide an explanation of how rodents might do visual generalization (Zoccolan et al., 2009). Up to this point, there does not seem to be any evidence for separation of image analysis and image categorization; they are generally treated as properties of the same ventral stream process and, at least implicitly, as properties of the same stages of visual processing (Poggio et al., 2012; Poggio and Ullman, 2013). If, as the work described here implies, image analysis and pattern discrimination might be subserved by different tissue, there would need to be a revision of most theoretical models of primate visual processing, which posit that classification and pattern recognition occur concurrently in a mutually dependent manner (Poggio et al., 2012).

Previous studies showed that monkeys with TE lesions are markedly impaired in a delayed nonmatch-to-sample (DNMS) task even with a short delay period (Mishkin, 1982; Buffalo et al., 1999). A categorical DNMS task requires two types of visual memory: longer-term memory for the category and a shorter-term or working memory for a sample cue to compare with the test stimulus that appears after a delay period (Freedman et al., 2003). Tasks 1 and 2 only require memory for the category; the cue remains on throughout the trial until the monkey responds, and hence there is no working memory demand.

Our results raise questions about how categorization based on visual generalization is perfomed. A pair of categories that was discriminated accurately by a powerful linear classifier (human/monkey faces) was classified well by monkeys also, both before and after TE removal. When a perceptually more ambiguous distinction, dogs versus cats, was presented to the linear classifier, it struggled, whereas normal monkeys did well. The surprise is that monkeys with TE removal performed better than the linear classifier ~95% of the time, and the monkeys did this within the first session of exposure. The result with the linear classifier makes it unlikely that there is any single so-called "low-level" feature available for the dog/cat discrimination. It remains surprising to us that the monkeys with TE lesions were able perform the cat/dog categorization at all. It is possible that the monkeys adopted an alternative strategy based on the functions of cortical areas that remained intact, but a question remains: For what level of ambiguity in categorical discrimination is area TE critical?

References

Afraz SR, Kiani R, Esteky H (2006) Microstimulation of inferotemporal cortex influences face categorization. Nature 442:692–695. CrossRef Medline

Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG (2009) Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. J Neurophysiol 101: 688–700. Medline

Blakemore C, Campbell FW (1969) On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J Physiol 203:237–260. CrossRef Medline

Buckley MJ, Gaffan D (2006) Perirhinal cortical contributions to object perception. Trends Cogn Sci 10:100–107. CrossRef Medline

Buffalo EA, Ramus SJ, Clark RE, Teng E, Squire LR, Zola SM (1999) Dissociation between the effects of damage to perirhinal cortex and area TE. Learn Mem 6:572–599. CrossRef Medline

Bussey TJ, Saksida LM, Murray EA (2003) Impairments in visual discrimination after perirhinal cortex lesions: testing 'declarative' vs. 'perceptual-

mnemonic' views of perirhinal cortex function. Eur J Neurosci 17: 649–660. CrossRef Medline

Cowey A, Gross CG (1970) Effects of foveal prestriate and inferotemporal lesions on visual discrimination by rhesus monkeys. Exp Brain Res 11: 128–144. Medline

Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4:2051–2062. Medline

Farah MJ (1985) Psychophysical evidence for a shared representational medium for mental images and percepts. J Exp Psychol Gen 114:91–103. CrossRef Medline

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. Science 291:312–316. CrossRef Medline

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci 23:5235–5246. Medline

Fukushima K (1980) Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36:193–202. CrossRef Medline

Gothard KM, Erickson CA, Amaral DG (2004) How do rhesus monkeys (Macaca mulatta) scan faces in a visual paired comparison task? Anim Cognit 7:25–36. CrossRef

Green DM (1964) General prediction relating yes-no and forced-choice results. J Acoust Soc Am 36:1042–1042.

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. J Neurophysiol. 35: 96–111. Medline

Hays AV, Richmond BJ, Optican LM (1982) A Unix-based multiple process system for real-time data acquisition and control. Paper presented at WESCON proceedings conference, Annaheim, CA, September.

Huang GB, Ramesh M, Berg T, Learned-Miller E (2007) Labeled faces in the wild: a database for studying face recognition in unconstrained environments. Amherst, MA: University of Massachusetts Technical Report 07–49.

Iwai E, Mishkin M (1968) Two visual foci in the temporal lobe of monkeys. In: Neurophysiological basis of learning and behavior (Yoshii N, Buchwald NA, eds) pp 23–33. Osaka: Osaka UP.

Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J Neurophysiol 97:4296–4309. CrossRef Medline

Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. J Neurophysiol 71:856–867. Medline

Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends Cogn Sci 17:26–49. CrossRef Medline

Liu Z, Richmond BJ, Murray EA, Saunders RC, Steenrod S, Stubblefield BK, Montague DM, Ginns EI (2004) DNA targeting of rhinal cortex D2 receptor protein reversibly blocks learning of cues that predict reward. Proc Natl Acad Sci U S A 101:12336–12341. CrossRef Medline

Longford, NT (1993) Random coefficient models. Oxford, UK: Oxford UP.

Magnussen S, Greenlee MW, Asplund R, Dyrnes S (1991) Stimulus-specific mechanisms of visual short-term memory. Vision Res 31:1213–1219. CrossRef Medline

Matsumoto N, Okada M, Sugase-Miyamoto Y, Yamane S, Kawano K (2005) Population dynamics of face-responsive neurons in the inferior temporal cortex. Cereb Cortex 15:1103–1112. Medline

Minamimoto T, Saunders RC, Richmond BJ (2010) Monkeys quickly learn and generalize visual categories without lateral prefrontal cortex. Neuron 66:501–507. CrossRef Medline

Mishkin M (1982) A memory system in the monkey. Philos Trans R Soc Lond B Biol Sci 298:83–95. Medline

Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: two cortical pathways. Trends Neurosci 6:414–417. CrossRef

Mosher CP, Zimmerman PE, Gothard KM (2010) Response characteristics of basolateral and centromedial neurons in the primate amygdala. J Neurosci 30:16197–16207. CrossRef Medline

Murray EA, Bussey TJ, Saksida LM (2007) Visual perception and memory: a new view of medial temporal lobe function in primates and rodents. Annu Rev Neurosci 30:99–122. CrossRef Medline

Nefian AV, Hayes MH (2000) Maximum likelihood training of the embed-

ded HMM for face detection and recognition. Paper presented at IEEE International Conference on Image Processing, Vancouver, BC, Canada, September.

Poggio T, Ullman S (2013) Vision: are models of object recognition catching up with the brain? Ann N Y Acad Sci 1305:72–82. CrossRef

Poggio T, Mutch J, Leibo J, Rosasco L, Tacchetti A (2012) The computational magic of the ventral stream: sketch of a theory (and why some deep architectures work). Technical report MIT-CSAIL-TR-2012-035. Cambridge: MIT.

R Development Core Team (2004) A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Riesenhuber M, Poggio T (2000) . Models of object recognition. Nat Neurosci 3 [Suppl]:1199–1204.

Sigala N, Logothetis NK (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. Nature 415:318–320. CrossRef Medline

Stefanacci L, Amaral DG (2002) Some observations on cortical inputs to the macaque monkey amygdala: an anterograde tracing study. J Comp Neurol 451:301–323. CrossRef Medline

Sugase Y, Yamane S, Ueno S, Kawano K (1999) Global and fine information coded by single neurons in the temporal visual cortex. Nature 400: 869–873. CrossRef Medline

Sugase-Miyamoto Y, Matsumoto N, Ohyama K, Kawano K (2014) Face in-

version decreased information about facial identity and expression in face-responsive neurons in macaque area TE. J Neurosci 34:12457–12469. CrossRef Medline

Suzuki WA, Amaral DG (1994) Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. J Comp Neurol 350:497–533. CrossRef Medline

Tanaka K (1996) Inferotemporal cortex and object vision. Annu Rev Neurosci 19:109–139. CrossRef Medline

Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB (2003) Faces and objects in macaque cerebral cortex. Nat Neurosci 6:989–995. CrossRef Medline

Webster MJ, Ungerleider LG, Bachevalier J (1991) Connections of inferior temporal areas TE and TEO with medial temporal-lobe structures in infant and adult monkeys. J Neurosci 11:1095–1116. Medline

Weiskrantz L, Saunders RC (1984) Impairments of visual object transforms in monkeys. Brain 107:1033–1072. CrossRef Medline

Yeshurun Y, Carrasco M, Maloney LT (2008) Bias and sensitivity in two-interval forced choice procedures: tests of the difference model. Vision Res 48:1837–1851. CrossRef Medline

Zoccolan D, Oertelt N, DiCarlo JJ, Cox DD (2009) A rodent model for the study of invariant visual object recognition. Proc Natl Acad Sci U S A 106:8748–8753. CrossRef Medline