

## ORIGINAL ARTICLE

# Contributions of the Monkey Inferior Temporal Areas TE and TEO to Visual Categorization

Tsuyoshi Setogawa<sup>1,2,†</sup>, Mark A. G. Eldridge<sup>1,†</sup>, Grace P. Fomani<sup>1</sup>, Richard C. Saunders<sup>1</sup> and Barry J. Richmond<sup>1</sup>

<sup>1</sup>Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA and <sup>2</sup>System Emotional Science, Faculty of Medicine, University of Toyama, Toyama 930-0194, Japan

Address correspondence to Barry J. Richmond, National Institutes of Health, National Institute of Mental Health, Bldg 49, Bethesda, MD 20892, USA.  
Email: [barryrichmond@mail.nih.gov](mailto:barryrichmond@mail.nih.gov)

<sup>†</sup>Tsuyoshi Setogawa and Mark A. G. Eldridge have contributed equally to this work

## Abstract

The ability to categorize images is thought to depend on neural processing within the ventral visual stream. Recently, we reported that after removal of architectonic area TE, the terminal region of the ventral stream, monkeys were still able to categorize images as cats or dogs moderately well. Here, we investigate the contribution of TEO, the architectonically defined region located one step earlier than area TE in the ventral stream. Bilateral removal of TEO caused only a mild impairment in categorization. However, combined TE + TEO removal was followed by a severe, long-lasting impairment in categorization. All of the monkeys tested, including those with combined TE + TEO removals, had normal low-level visual functions, such as visual acuity. These results support the conclusion that categorization based on visual similarity is processed in parallel in TE and TEO.

**Key words:** aspiration, inferior temporal cortex, rhesus monkey, visual categorization

## Introduction

Primates, including humans, can quickly group images based on visual similarity. The ability to perform this type of categorization is thought to arise from the activity of the neurons in the ventral visual stream. The ventral visual stream is a sequentially connected set of visual areas extending from primary visual cortex (V1) through other visual areas including V2 and V4, and ending in the inferior temporal cortex, areas TEO and TE. Representations of images are built up from simple features in V1, through intermediate associations of features in V2 and V4, to information about whole, complex images in inferior temporal cortex (Iwai and Mishkin 1968; Ungerleider and Mishkin 1982; Sigala and Logothetis 2002; Afraz et al. 2006; Kiani et al. 2007; Sato et al. 2013). Neurophysiological and lesion studies implicate IT cortex in high-level visual processing, for example, visual object recognition, object discrimination, and memory of complex objects (Fujita et al. 1992; Tanaka 1996; Buckley et al.

1997; Vogels et al. 1997; Buffalo et al. 1999, 2000; Gainotti 2000; Matsumoto et al. 2016).

Behaviorally, categorization is accomplished quickly, accurately, and seemingly without conscious effort, even for stimuli that have never before been seen. This categorization based on similarity makes it possible to infer the significance of objects, both those that are familiar as well as those never seen before, for example, prey or predator, tasty (yellow banana) versus not so tasty foods (green banana). Monkeys can also categorize pictures of natural objects (e.g., dogs vs. cats), and artificial objects, (e.g., cars vs. trucks) (Vogels 1999; Freedman et al. 2001, 2002; Minamimoto et al. 2010; Matsumoto et al. 2016; Eldridge et al. 2018).

Recently, we showed that removing TE caused only modest impairments in visual categorization using a visually cued two-interval forced choice paradigm (13.0% increase in error rate in the categorization task compared to the control) (Matsumoto

et al. 2016; Eldridge et al. 2018). The partial sparing of categorization after the TE removals surprised us. One possible explanation is that TEO, the architecturally identified region just before TE in the ventral stream hierarchy, can substitute for some of the missing functionality after TE removal. Although TEO is physically smaller than TE, neurons in TEO represent high-level visual properties, and the region has a full representation of the visual field. Allman and Kaas described the re-representation of the visual field as the standard for recognizing a visual area; hence, TEO should be considered a discrete visual area (Allman and Kaas 1974). It contains neurons with large receptive fields, although still smaller than those typically reported for area TE, and unlike TE, TEO seems to be retinotopically organized. It has been also suggested that TEO is important for visual feature analysis and integration, whereas perhaps TE plays a more important role when memory for a whole object is required (Iwai and Mishkin 1968). This led us to speculate that TEO may make a critical contribution to visual categorization. To test this hypothesis, we compared performance on a visual categorization task across four groups of monkeys: those with TEO removals, those with TE removals (data reproduced from Eldridge et al), those with TE + TEO removals, and unoperated controls. The TEO-removal group showed mild impairments that disappeared after 1–3 days of practice (13.4% increase in error rate in the categorization task compared to the control). The TEO plus TE group was severely impaired (31.7% increase in error rate in the categorization task compared to the control). The degree of deficit was approximately equal to the sum of the effects for each of TEO and TE.

## Materials and Methods

### Subjects

Subjects were eight adult rhesus monkeys (*Macaca mulatta*). Three monkeys (one male; weighing 11.6 kg, two females; weighing 5.4 and 9.3 kg) received bilateral aspiration removals of area TEO (Supplementary Fig. 1). Two monkeys (one male; weighing 9.6 kg, one female; weighing 5.5 kg) received bilateral aspiration removals of areas TE and TEO (Supplementary Fig. 2). After collecting behavioral data from monkeys with TEO removals, one of the three monkeys (monkey M) received additional bilateral aspiration removals of area TE (Supplementary Fig. 2). These five monkeys performed a visual categorization task before and after surgery; the data collected before surgery were used as a within-subject control. The five monkeys with TEO or TE + TEO removals received additional testing in tasks not used prior to surgery (see Results); three unoperated monkeys (three males; weighing 7.8–9.5 kg) were used as controls for these additional experiments. 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.

### Experimental Conditions

Monkeys sat in a primate chair facing a 22-inch computer monitor (Samsung 2233RZ) placed 57 cm from their eyes. A touch sensitive bar was attached to the front panel of the primate chair at the level of the monkey's hand. A water reward was dispensed from a stainless steel tube that was positioned at the monkey's lips. Experiments were conducted in a sound-isolated

dark room. Experimental control and data acquisition were performed using the real-time experimental system "REX" adapted for the QNX operating system (Hays et al. 1982). Visual stimuli were presented by "Presentation" (Neurobehavioral Systems, Inc.) running on a Windows computer.

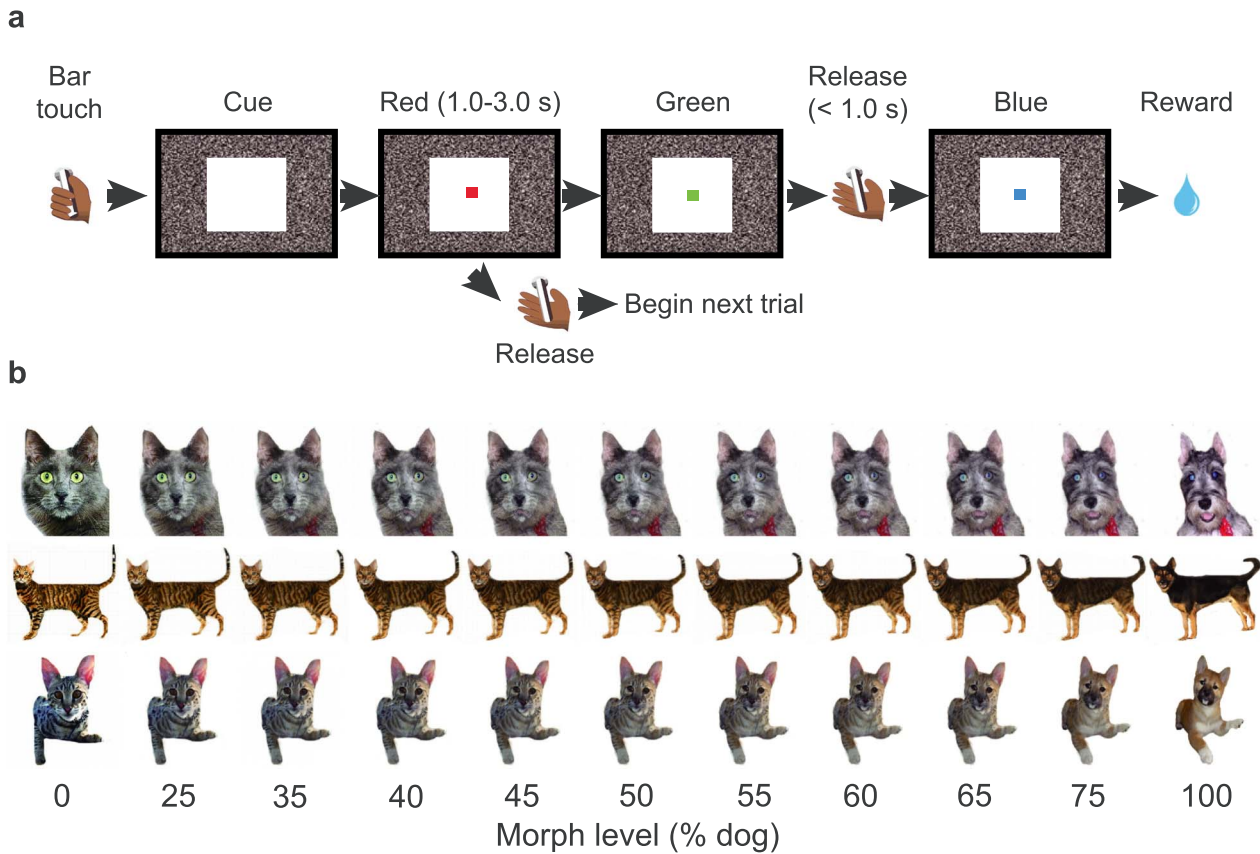
### Task Procedures

Monkeys were initially trained to grasp and release a touch sensitive bar to earn fluid rewards. After this initial shaping, a red/green color discrimination task was introduced (Bowman et al. 1996). The trial began with a bar press, and 100 ms later, a small red target square ( $0.5^\circ \times 0.5^\circ$ ) was presented at the center of the display (overlying a white noise background). Animals were required to continue grasping the touch bar until the color of the target square changed from red to green. Color changes occurred randomly 2000–3000 ms after bar touch. Rewards were delivered if the bar was released between 200 and 1000 ms after the color change; bar releases occurring either before or after this epoch were counted as errors. All correct responses were followed by visual feedback (target square color changed to blue) after bar release and reward delivery 200–400 ms after visual feedback. There was a 2-s intertrial interval (ITI), regardless of the outcome of the previous trial.

After an animal reached criterion in the red/green color discrimination task (two consecutive days with >85% correct performance), the monkeys progressed to category training (Fig. 1). In the first phase of category training, 20 dog and 20 cat images were used. Each trial began when the animal grasped the touch sensitive bar. If the monkey released the bar during the green target when a dog was presented, the monkeys received one drop of liquid reward (Fig. 1a). If the monkey released the bar during the green target when a cat was presented, there was a 4000–6000 ms time-out with no reward. If the monkey released the bar during the red target when either category of stimulus was present, no reward was delivered, and the monkey could initiate a new trial after the standard ITI. Therefore, the optimal behavior is to release during the red target for the trials on which cats are presented, essentially skipping on to the next randomly selected trial, and release during the green target for the trials on which dogs are presented to obtain a reward. This design is effectively a visually-cued two-interval forced choice (2-IFC) task, with asymmetrical reward. The 20 dog and 20 cat stimuli were repeated multiple times per session. In the second phase of category training, the monkeys were presented with four larger sets of trial-unique images (240 cats and 240 dogs), to confirm that the monkeys were able to classify stimuli based on visual perceptual categorization (cat-dog trial-unique task).

For the perceptually challenging tests of categorization, we used 20 sets of morphed stimuli, as in our previous study (Eldridge et al. 2018) (Fig. 1b). For the experiments with morphed stimuli, releasing the bar during the green target resulted in a 4000–6000 ms time-out with no reward if the stimulus was more cat-like (i.e., <50% dog), and a reward if the stimulus was more dog-like (i.e., >50% dog). The outcome of trials on which a stimulus at the category boundary (i.e., =50% dog) was presented was determined probabilistically; 50% of trials resulted in a reward delivery, 50% resulted in a 4000–6000 ms time-out. We collected behavioral data for 10 days using the same set of morphed images.

For the simple discrimination task, two cues were used; these were black and white block ("Walsh") patterns ( $13^\circ \times 13^\circ$ ) (Fig. 4a). These cues signaled whether a release during the green



**Figure 1.** Task and visual stimuli. (a) A single trial of the visual categorization task (rewarded trial). By touching the bar, a morphed cat–dog image was presented. When the dog-like image was presented (shown as Fig. 1a, 50–100% in Fig. 1b), the animal was required to continue grasping the touch bar until the color of the target square changed from red to green to obtain water reward. When the cat-like image was presented (0–50% in Fig. 1b), the animal was required to release the touch bar during the red target period to skip the current trial; otherwise, it received a time-out with no reward (see Materials and Methods). (b) Three examples of the cat–dog morphed images.

target would result in the delivery of a drop of liquid reward, or a 4000–6000 ms “time-out.” Monkeys could avoid the predicted outcome by releasing the lever before the red target transitioned to green; a new trial could then be initiated after the standard ITI. We tested each group of monkeys for one session on this task.

### Visual Cues

All visual cues were jpeg- or pcx-format photos (200 × 200 pixels). The training sets of dogs/cats used in this study are the same as in our previous report (Minamimoto et al. 2010). The images used in the main visual categorization task were generated from a subset of the training images, in which pairs of cats and dogs were used to create cat–dog morph sequences using FantaMorph software (Abrosoft). For the main categorization task, 20 cat and 20 dog images were morphed with the distribution of stimuli concentrated around the category boundary (11 levels, 0, 25, 35, 40, 45, 50, 55, 60, 65, 75, and 100% dog) (Fig. 1b). For the masked stimulus task, the same set of cat–dog morph series was used as in main categorization task, but on four-fifths of trials, the stimuli were overlaid with one of four coarse black-block masks (Fig. 3a). For the trial-unique morphed-stimuli categorization task, a new set of 20 cats and 20 dogs was used to create new cat–dog morph series. In this task, each cat image was morphed with two dog images with equal

distribution of the morph level, and vice versa (11 levels, 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100% dog) (Fig. 3g). We tested the trial-unique morphed-stimuli categorization task for 1 day.

### Data Analysis

The “R” statistical programming language (R Foundation for Statistical Computing, R Development Core Team, 2017) was used for all statistical analyses.

A generalized linear mixed model (GLMM) analysis with a binomial link function was performed for analyzing the categorization performance

$$P = \gamma_0 + \gamma_1 \text{Level} * \gamma_2 \text{Condition} + (1|\text{Subject}), \quad (1)$$

where P is trial-by-trial categorization performance (0 indicating the trial was reported as cat and 1 indicating the trial was reported as dog), “Level” is the morph level, “Condition” is the lesion group (TE + TEO removals, TEO removals, TE removals, and control) or the masked/no masked condition,  $\gamma_0$  is the intercept,  $\gamma_1$  and  $\gamma_2$  are the coefficients estimated by GLMM, and (1|Subject) is the random effect for each monkey.

For analyzing the reaction time during the red target or the green target period, we used before and after surgery data of the main visual categorization task (10 days each) and conducted

a generalized linear model (GLM) analysis with a Gaussian link function as follows:

$$RT = \alpha_0 + \alpha_1 \text{Level} * \alpha_2 \text{Condition}, \quad (2)$$

where RT is logarithm of the reaction time, “Level” is the morph level (we removed the choice reaction time data on the categorization boundary, i.e., 50% dog, for this analysis; 0, 25, 35, 40, and 45% dog were used for the red target period, 55, 60, 65, 75, and 100% dog were used for the green target period), “Condition” is the lesion group (TE + TEO removals, TEO removals, and control),  $\alpha_0$  is the intercept, and  $\alpha_1$  and  $\alpha_2$  are the coefficients estimated by GLM.

For analyzing the processing time, we used a GLM with a Gaussian link function as follows:

$$PT = \beta_0 + \beta_1 \text{Level} * \beta_2 \text{Condition}, \quad (3)$$

where PT is logarithm of the processing time, “Level” is the morph level (0, 25, 35, 40, and 45% dog), “Condition” is the lesion group (TE + TEO removals, TEO removals, and control),  $\beta_0$  is the intercept, and  $\beta_1$  and  $\beta_2$  are the coefficients estimated by GLM.

### Order of Testing

All monkeys received basic categorization training prior to surgery. Monkeys received 10 sessions of the main visual categorization task using morphed stimuli before and after surgery. Three monkeys received bilateral TEO aspiration removals, and two received bilateral TE + TEO removals. After collecting behavioral data from monkeys with TEO removals, one of the three monkeys (monkey M) received an additional bilateral aspiration removal of area TE. Thus, for the main categorization task, pre- and post-op data were compared within-subjects. The subsequent experiments were administered only postoperatively. The performance of the monkeys with aspiration removals on these latter tasks was compared to that of a group of control monkeys that had received parallel training. After the main visual categorization task was tested, we conducted the simple discrimination task (1 day) and a contrast sensitivity task (5 days) for assessing low-level visual functions. Then, the cat-dog trial-unique task (1 day) and the masked stimulus task were tested (10 days). Finally, we tested the trial-unique morphed-stimuli categorization task (1 day). All tests after lesion were conducted within 3 months.

## Results

We tested eight monkeys, three with TEO removals, three with combined TE + TEO removals (including one from the previous group after a second surgery to remove TE bilaterally), and three normal unoperated controls, on a visual categorization task using 20 sets of cat/dog morphs (Fig. 1). Monkeys learned to categorize morphed images as either “cat-like” or “dog-like” to avoid a time-out or to obtain a liquid reward.

### Categorization after TEO Removals

We collected pre- and postoperative behavioral data for 10 days from the TEO-removal group. There was a mild impairment in categorization ability for the first 2 days of postoperative testing (generalized linear mixed model [GLMM], Eq. 1; TEO vs. control,

effect of condition:  $z = -10.48$ ,  $P = 2.0 \times 10^{-16}$ ) (Fig. 2a—red and b—top). By the third day, the performance of all three monkeys had returned to their pre-lesion levels (Fig. 2b—top). This transient impairment in the TEO-removal group was smaller than seen for monkeys with TE removals reported previously (Fig. 2a—orange; Eldridge et al. 2018).

### Categorization after TE + TEO Removals

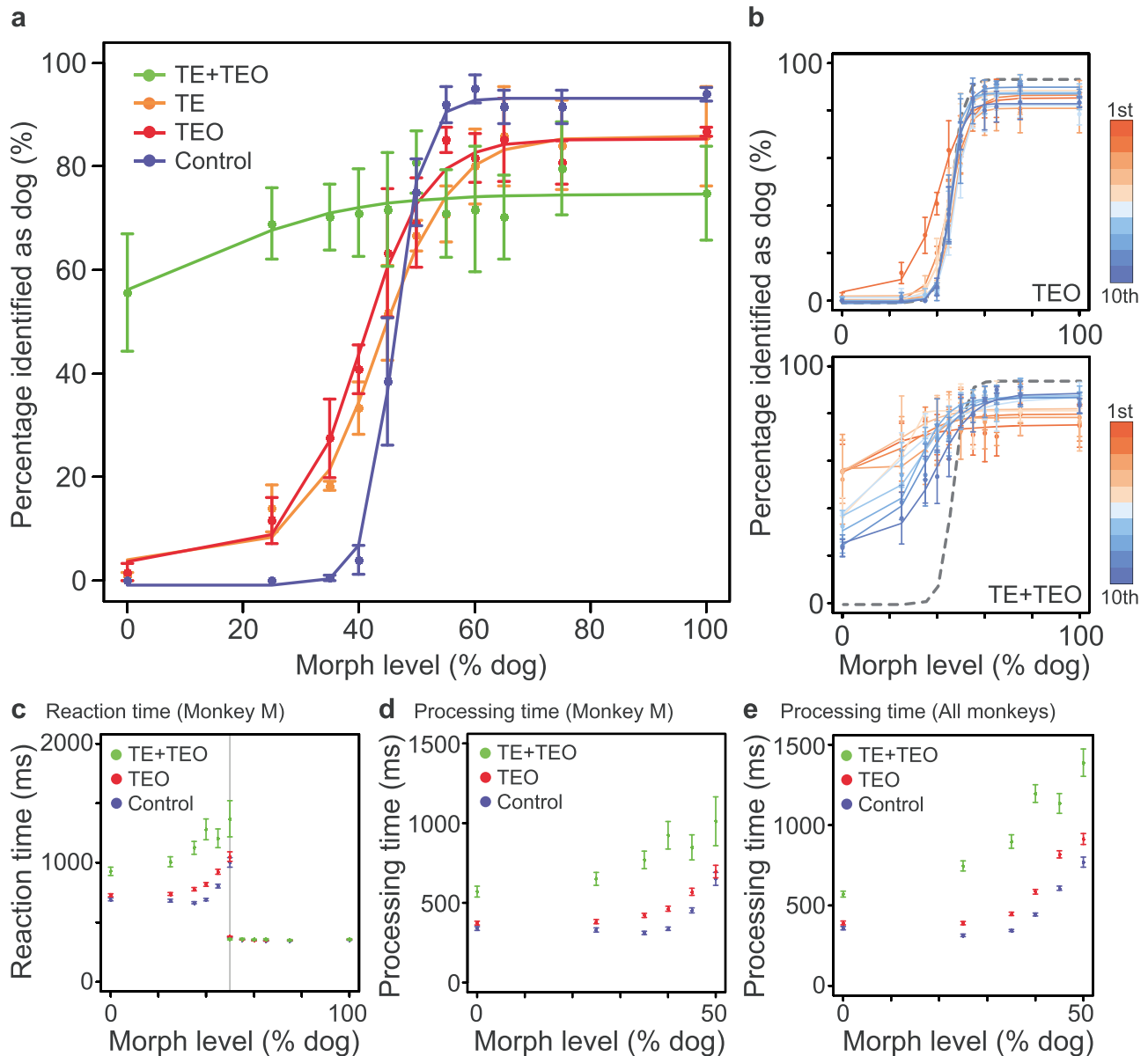
We removed TE and TEO bilaterally in two monkeys and added a TE lesion to one of the monkeys that had previously received a TEO removal (monkey M). The performance of the monkey with the two-stage lesion was indistinguishable from that of the monkeys with one-stage TE + TEO removals in this and all subsequent tasks; hence, the data from these three monkeys were pooled. For the TE + TEO-removal group, categorization performance was severely impaired (GLMM, Eq. 1; TE + TEO vs. control, effect of condition:  $z = -48.35$ ,  $P = 2.0 \times 10^{-16}$ ) (Fig. 2a—green). Performance recovered partially with experience, but remained significantly poorer than the TEO or TE groups by the 10th day of postoperative testing (GLMM, Eq. 1; TE + TEO vs. TE, effect of condition:  $z = 11.16$ ,  $P = 2.0 \times 10^{-16}$ ; TE + TEO vs. TEO, effect of condition:  $z = 11.04$ ,  $P = 2.0 \times 10^{-16}$ ) (Fig. 2b—bottom). As shown in Figure 2a,b, the categorization performance was asymmetrical: higher percentages of cat stimuli were categorized as dog than dog stimuli categorized as cat. This asymmetry likely reflects the asymmetric reward structure of the task; monkeys with impaired ability to categorize—for example, as the result of a lesion—tend to present a bias toward releasing during the green interval because only the dog stimuli are associated with reward.

### Reaction Times

The pattern of reaction times indicates that monkeys in all groups were sensitive to the degree of mixing in the morphs when responding to a cat (Fig. 2c for monkey M). When the cat-like image was presented, the reaction time was defined as the time between onset of the stimulus and bar release. When the dog-like image was presented, the reaction time was defined as the time between onset of the green cue and bar release. The reaction time following the green target was constant across morph level (55–100% dog), presumably because the monkey had already made a decision that the presented stimulus was dog-like before the red target changed to green (generalized linear model [GLM], monkey M, Eq. 2; effect of morph level:  $t = 0.75$ ,  $P = 0.46$ ). We interpret the response time during the green target period as a basic visual-motor reaction time. This visual-motor reaction time was indistinguishable across lesion groups, suggesting that motor skill was not affected by TEO or TE + TEO removals (GLM, monkey M, Eq. 2; effect of task condition:  $t = 0.91$ ,  $P = 0.36$ ).

To characterize the reaction times to more cat-like stimuli (i.e., lever releases during the red target period), we introduced a measure we term “processing time” that was calculated by subtracting the average visual-motor reaction times (55–100% dog) from the reaction times for each cat-like level (0–45% dog) (Fig. 2d—monkey M and e—all monkeys). Because the visual-motor reaction times were different among animals, the processing time provides a normalized measure of the time it takes the monkeys to decide whether an image is more cat-like. The processing time for the TE + TEO group was significantly longer than the control group (GLM, Eq. 3; effect of condition:  $t = 10.5$ ,  $P = 2.0 \times 10^{-16}$ ). The processing times for the





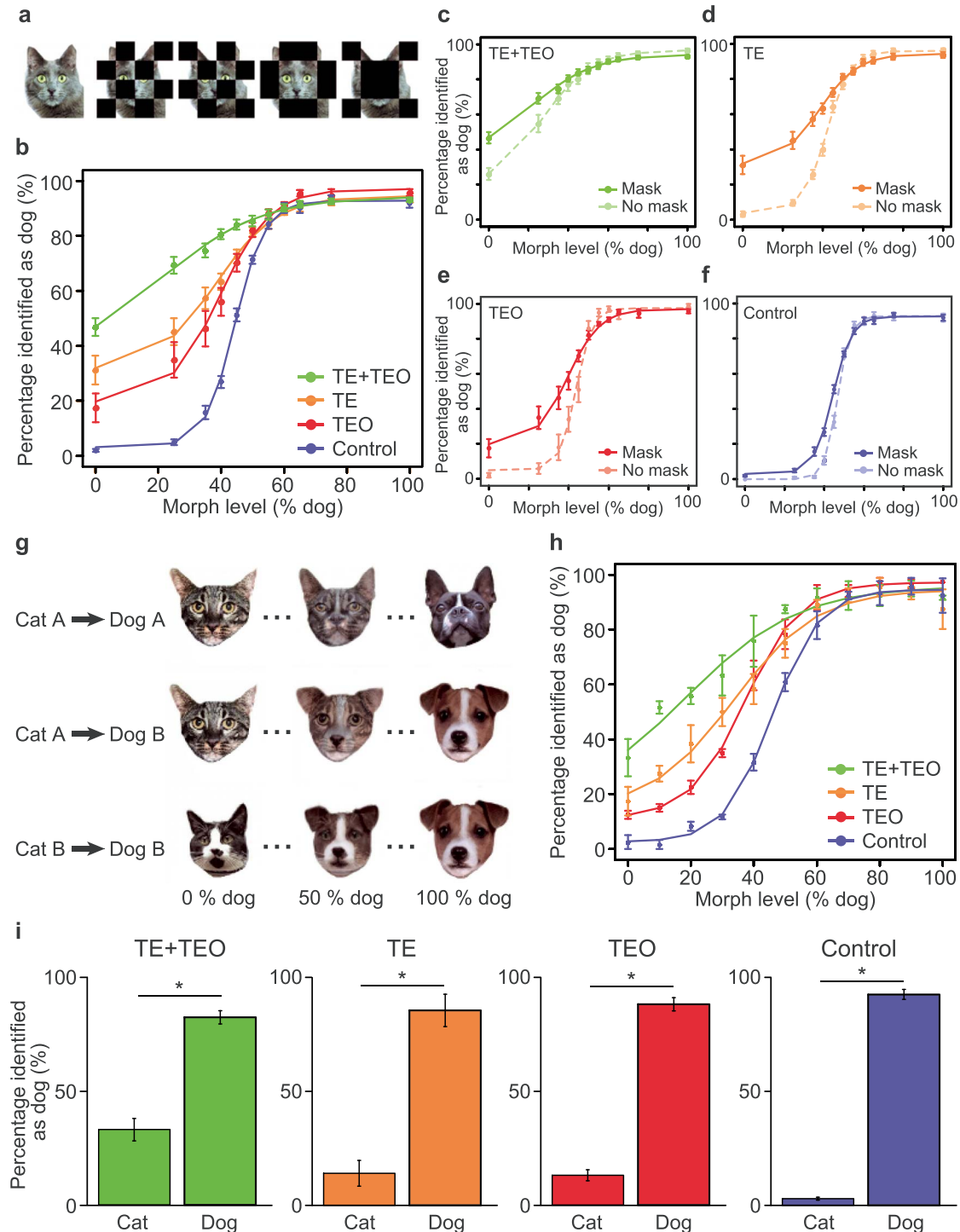
**Figure 2.** Categorization performance. (a) Percentage identified as dog in the TE + TEO, TE, TEO, and control groups on the first experimental day (TE data have been previously reported [Eldridge et al. 2018]). (b) Day-by-day task performance for 10 days in the TE + TEO and TEO groups. A deep red indicates the first day and a deep blue indicates the last day. A gray dashed line indicates the task performance before surgery (blue line in Fig. 2a). (c) An example of the reaction time (monkey M). The vertical gray line indicates the category boundary, that is, 50% dog. (d) An example of the processing time (monkey M). (e) The average processing time of all animals. Error bar: SEM.

TEO and control groups were indistinguishable (GLM, Eq. 3; effect of condition:  $t=0.67$ ,  $P=0.51$ ). These results indicate that the TE + TEO-removal group takes longer to process the stimuli even when the animals have previously seen them.

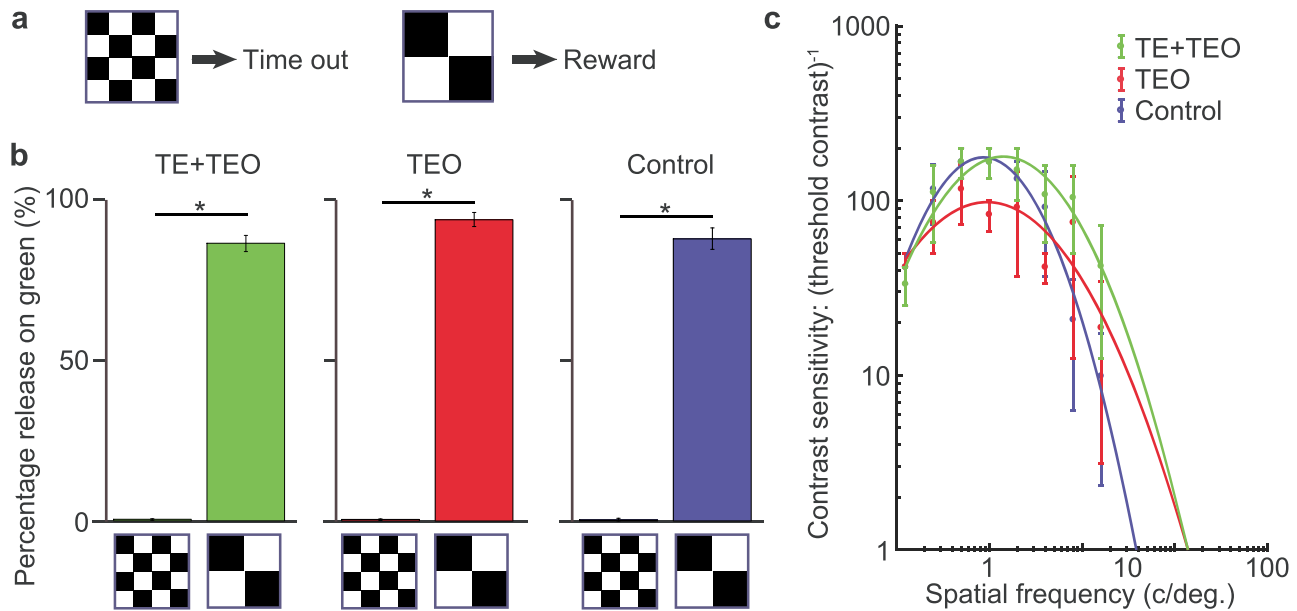
### Role of Experience in Categorization

We tested the possibility that the monkeys with TEO or TE + TEO removals had compensated for impaired visual categorization by memorizing one or more simple features of each morph series (e.g., the “ear” of the stimuli in the first row of Fig. 1b). To examine this, we introduced two manipulations to the categorization task; a masked stimulus set and a trial-unique stimulus set. In the masked stimulus task, the stimuli were

overlaid with one of four coarse black-block masks on four-fifths of the trials (Fig. 3a). If the animals with TEO or TE + TEO lesions rely on a limited set of (or even single) diagnostic features to categorize a presented image, their performance should be impaired by masking. Consistent with our hypothesis, both TEO and TE + TEO-removal groups showed severe impairments in categorizing masked stimuli relative to the interleaved unmasked trials (GLMM, Eq. 1; Mask vs. No mask in TE + TEO, effect of condition:  $z=-8.56$ ,  $P=2.0 \times 10^{-16}$ ; Mask vs. No mask in TEO, effect of condition:  $z=-14.23$ ,  $P=2.0 \times 10^{-16}$ ) (Fig. 3b,c,e). Conversely, the performance of the control group was only mildly affected by masking (GLMM, Eq. 1; Mask vs. No mask in Control, effect of condition:  $z=-7.73$ ,  $P=1.1 \times 10^{-14}$ ) (Fig. 3f).



**Figure 3.** Categorization performance in the mask categorization task and the trial-unique categorization task. (a) Examples of the visual stimuli in the mask categorization task. Four checker-board masks were placed over each of the stimuli used in the main categorization task and presented interleaved with an unmasked version of each stimulus. (b) Categorization performance of the TE + TEO, TE, TEO, and control groups (TE data have been previously reported [Eldridge et al. 2018]) in the mask categorization task. (c–f) Categorization performance on masked (mean of all masks) versus unmasked stimuli for (c) TE + TEO, (d) TE, (e) TEO, and (f) control groups. (g) Examples of the visual stimuli in the trial-unique categorization task. Each cat was morphed with two dogs, and vice versa. Examples at the 0%, 50%, and 100% dog level are shown. (h) Categorization performance of the TE + TEO, TE, TEO, and control groups (TE data have been previously reported [Eldridge et al. 2018]) in the trial-unique categorization task. Error bar: SEM. (i) Categorization performance of the TE + TEO, TE, TEO, and control groups in the cat-dog trial-unique task (no-morphed 240 cats and 240 dogs).



**Figure 4.** Low-level visual functions. (a) The visual stimuli of the cue discrimination task. (b) Discrimination performance of the TE + TEO, TEO, and control group.  $*P < 2.2 \times 10^{-16}$ . (c) The contrast sensitivity curves of the TE + TEO, TEO, and control group. Error bar: SEM.

For the trial-unique stimulus task, we prepared a large set of novel morph images as trial-unique stimuli. A key difference from the stimulus set used for main experiment was that each cat image was morphed with two dog images, and vice versa (Fig. 3g; Eldridge et al. 2018). This manipulation reduces the utility of a strategy focused on a single memorized feature. We tested this trial-unique categorization task for one day (a single session). Consistent with the results of the masked stimulus task, the categorization performance of the TE + TEO-removal group was severely impaired compared to the other groups (e.g., TE + TEO vs. TE, GLMM, Eq. 1; effect of condition:  $z = 5.4$ ,  $P = 5.5 \times 10^{-8}$ ) (Fig. 3h). The degree of impairment in both the masked stimulus and the trial-unique stimulus tasks was consistent with the main experiment (Fig. 2a), the order of impairment from greatest to least was TE + TEO, TE, and TEO (Fig. 3b,h). The processing time in the trial-unique task was also analyzed. For the TE-TEO group, the processing time was significantly longer than the control group (GLM, Eq. 3; effect of condition:  $t = 2.69$ ,  $P = 7.4 \times 10^{-3}$ ) (Supplementary Fig. 3). The processing time for the TEO group was also significantly longer than the control group (GLM, Eq. 3; effect of condition:  $t = 4.09$ ,  $P = 5.1 \times 10^{-5}$ ) (Supplementary Fig. 3). These results indicate that both the group with TE + TEO removals and the group with TEO-only removals take longer to process the stimuli, whether they are trial-unique or familiar.

Because two adjacent morphed stimuli are visually similar (e.g., 35% dog and 40% dog in Fig. 1b), it is possible that the subjects learn stimulus-reward associations within a single session instead of generalizing from previous experience with categorical exemplars. To examine this possibility, we included a single session of the cat-dog trial-unique task (nonmorphed 240 cats and 240 dogs) in which completely novel images were used; the subjects can only solve this task via visual perceptual generalization. We observed the same ranking of results as obtained at the 0% and 100% morph level in the morphed categorization tasks ( $\chi^2$ -test, % identified as cat vs. % identified as dog, TE + TEO:  $\chi^2 = 356.1$ ,  $df = 1$ ,  $P = 2.0 \times 10^{-16}$ ; TE:  $\chi^2 = 719.7$ ,

$df = 1$ ,  $P = 2.0 \times 10^{-16}$ ; TEO:  $\chi^2 = 804.5$ ,  $df = 1$ ,  $P = 2.0 \times 10^{-16}$ ; Control:  $\chi^2 = 1925.5$ ,  $df = 1$ ,  $P = 2.0 \times 10^{-16}$ ) (Fig. 3i). This result confirms that the TE + TEO removal induces a severe impairment in visual categorization, rather than impairment of the stimulus-reward association learning.

#### Low-Level Visual Function after TEO and TE + TEO Removal

Two tasks were used to assess low-level visual functions: cue discrimination and contrast sensitivity. In the cue discrimination task, two different Walsh patterns were used; one cue associated with reward and the other associated with time-out (Fig. 4a). All groups of monkeys distinguished between the rewarded and unrewarded cues ([number of no-reward trials accepted/all no-reward trials]: TE + TEO; 6/1137 (0.4%), TEO; 6/1267 (0.5%), control; 3/686 (0.5%), [number of rewarded trials accepted/all reward trials]: TE + TEO; 1310/1498 (87.4%), TEO; 1342/1433 (93.6%), control; 781/882 (88.5%) (reward vs. no-reward,  $\chi^2$ -test, TE + TEO:  $df = 2$ ,  $P < 2.2 \times 10^{-16}$ , TEO:  $df = 2$ ,  $P < 2.2 \times 10^{-16}$ , control:  $df = 2$ ,  $P < 2.2 \times 10^{-16}$ ) (Fig. 4b). In the contrast sensitivity test (Matsumoto et al. 2016), full-screen sine wave gratings (i.e., the local intensity was modulated by a one-dimensional [vertical] sine wave across the screen) were presented 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, 0.08, 0.04, 0.02, 0.01, 0.005). Contrast was calculated as:  $LP - LT / (LP + LT)$ , where LP represents peak luminance and LT 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, with symmetric reward. 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. Grating contrast sensitivity, a test designed to assess the visual acuity of human subjects (Blackmore and Campbell 1969), was indistinguishable across all three groups (GLM, group:  $t=0.07$ ,  $df=2$ ,  $P=0.95$ ) (Fig. 4c). The contrast sensitivities of the three groups were similar to those of humans (Blackmore and Campbell 1969) and monkeys that received TE (Matsumoto et al. 2016) or rhinal cortex lesion (Eldridge et al. 2018).

## Discussion

Above we have shown that selective bilateral removal of the inferior temporal cortex (TE + TEO combined) interferes with categorical discrimination when tested with sets of trial-unique cats and dogs, visually degraded morphed images, and trial-unique morphed images. The severity of the deficit is different if either of the two subregions of inferior temporal cortex, areas TEO and TE, is removed independently. There was a significant deficit in the categorization of all trial-unique images after TEO removals, and a slightly more severe deficit after TE removals. After removal of either area, the monkeys' performance improved quickly with repetitions of an image set; however, only the TEO-removal group recovered to control levels of performance. When TEO and TE are both removed, the monkeys are severely impaired, and while they show some improvement with additional practice on a single image set, they remain severely impaired. It appears that TEO and TE lesions have an additive effect on the severity of the deficit, consistent with models derived from single unit recordings taken from subregions TEO and TE of IT cortex (Majaj 2015). The improvement with repeated image set presentation raises a difficulty for the experimentalist trying to study perception or perceptual categorization—the only presentation that can be assured to rely solely on perception/categorical memory is the first one. Every subsequent encounter is confounded by the possibility of recollective processes.

The canonical description of visual image processing by the brain posits that simple features, such as oriented edges or lines, are represented in caudal regions, beginning with area V1 (Hubel and Wiesel 1959), and that representations become increasingly more complex as information converges along a ventral pathway in a sequential, feed-forward manner, culminating in the representation of whole objects in inferior temporal cortex (Gross et al. 1972; Ungerleider and Mishkin 1982). Two observations from single neuron recording studies offered strong support to a sequential, feed-forward processing model for image analysis in the ventral stream. First, there were the progressively larger receptive fields, and, second there was the increasing complexity of stimulus selectivity in architectonically separable cortical brain regions as information flowed from caudal (V1) to rostral (ending in area TE). However, whether the ventral visual stream relies exclusively on feed-forward processing has been thrown into doubt by the observation of recurrent and bypass projections in studies of anatomical connectivity (Kravitz et al. 2013; Kar et al. 2019). Now, we add results showing that processing is not always strictly sequential. The observation that bilateral removal of area TEO—the region immediately upstream of area TE—produces milder deficits than those we previously reported after bilateral TE removals (Eldridge et al. 2018) indicates that the visual information used for analyzing images depends on a route from earlier in the ventral stream

stations to TE without passing through area TEO. The most direct path would be from connections to TE arising at earlier stages such as V4 (Distler et al. 1993; Ungerleider et al. 2008), assuming that those connections bypassing TEO have enough bandwidth to carry sufficient information for TE to analyze images. Previous studies have suggested that four-legged animals are more likely to be confused among one another than “simpler” contrasts, such as fruits versus tables or cars versus chairs (Cadieu et al. 2014). We elected cats and dogs as the categories for comparison in the present study on the basis that they were likely to yield high levels of confusion. We have previously demonstrated that a linear classifier performs more accurately on human face versus monkey face and car versus truck comparisons than it does a cat versus dog comparison (Matsumoto et al. 2016). To maximize the perceptual difficulty in the present study, we used morphed pairs of cats and dogs to create even more category-ambiguous intermediate images (Eldridge et al. 2018). Our expectation is that “simpler” comparisons could be performed at earlier stages of the visual system (e.g., bilateral removals of area TE produce no impairment in the ability to categorize human vs. monkey faces [Matsumoto et al. 2016]).

The greatest reduction in categorization accuracy occurred when we used previously unseen stimuli. Because the stimuli presented in this phase of the experiment were new, the only means by which the monkeys could have accurately classified them was to generalize from previously experienced exemplars. The data from using new images show that monkeys with combined TE + TEO removals exhibit a deficit in categorization accuracy that approximates a sum of the deficits observed following removal of either subregion of IT alone. Thus, it appears that TE and TEO work in parallel, and with minimum redundancy, to encode category membership of a novel stimulus. Even the monkeys with complete TE + TEO removals are able to categorize at above-chance levels with practice; on the first test session after the removals, they performed at chance. The rapid increase in performance with practice, plus the increased processing time observed for both lesion groups (see Fig. 2e and Supplementary Fig. 3), suggests that compensatory mechanisms may be invoked that preserve some degree of categorization accuracy at the cost of increased decision time.

Our data also demonstrate that the deficits observed in all treatment groups were ameliorated with increased familiarity to a stimulus set; classification accuracy improved with repeated postlesion exposure to the morphed stimulus set (Fig. 2b,c, and Eldridge et al. 2018). The results here show that the ability of the TEO-removal group to generalize from previously experienced exemplars remains compromised because the deficits in classifying novel stimuli were recorded after the monkeys received repeated exposure to the morphed stimulus set. Thus, the recovery in performance must be supported by the learning of an alternative strategy, presumably one based on the learning of stimulus–reward associations. The rapid and complete recovery of the TEO-removal group with stimulus repetition suggests that other regions (such as TE) can support the stimulus–reward association learning required to support this enhanced performance. The TE-removal group asymptote at a level of performance inferior to that of controls—this indicates that no other area can adequately support the fidelity of stimulus–reward associations needed to compensate for the loss of categorization capability conferred by TE removal. The TE + TEO removals produced the most substantial impairment in categorization accuracy—an initial near-total loss of function,



which recovered with practice to a level of accuracy consistently just above chance. As we proposed for the savings in categorization of novel stimuli discussed above, the residual ability of the TE + TEO group to classify at above-chance levels is likely subserved by projections from earlier in the visual system that bypass IT to subcortical targets. Taken together, these observations indicate that in the TE-removal group, TEO is likely the key substrate for the stimulus–reward associations that confer the ability of this group to improve so substantially with practice.

There are two possible explanations for our result showing slower and less complete recovery in the performance of the TE-removal group versus the TEO-removal group. One is that TEO, although a distinct architectonic and physiologically separate region, contributes to this categorization task as if TEO and TE are one larger functional region. Thus, the difference in recovery is related to the difference in the volume of tissue removed; that is, TE is a larger architectonic region; hence, the deficits observed correspond simply to the quantity of tissue removed, and not from a specific segregation of function. The second possibility is that TEO and TE are functionally one architectonic region and should not be considered as different. If the latter were the case, we would not expect to see the asymmetry in the impairment in categorization that appeared with novel exemplars (Fig. 3). In addition, previous data show that receptive fields are different in TEO and TE, and that TEO contains a full representation of the visual fields, the means by which Allman and Kaas separated functional visual regions (Allman and Kaas 1974). Thus, the weight of the evidence favors considering TEO and TE as different functional regions.

Over the past six decades, many studies have concluded that the inferior temporal cortex is critical for pattern discrimination (Iwai and Mishkin 1968; Cowey and Gross 1970), visual pattern recognition (Butter and Gekoski 1966; Weiskrantz and Saunders 1984), and by inference, visual perceptual categorization (Sigala and Logothetis 2002; Afraz et al. 2006; Kiani et al. 2007). There has remained a disconnect though. The data supporting IT participation in visual perceptual categorization have largely relied on correlations in the selectivity of neurons in physiological recordings. Our data show that both areas TEO and TE contribute to categorization-based behavior when subjects are challenged with novel stimuli but that performance quickly improves with repeated exposure to the same stimuli. Thus, caution must be exercised when interpreting the results of experiments in which stimuli are repeated, both behavioral and electrophysiological, as perceptual generalization can be easily confounded with other processes.

## Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

## Funding

Intramural Research Program; National Institute of Mental Health; National Institutes of Health; Department of Health and Human Services (annual report number ZIAMH002032).

## Notes

We thank Megan Fredericks and Grace Mammarella for assistance with behavioral testing.

## References

- Afraz SR, Kiani R, Esteky H. 2006. Microstimulation of inferotemporal cortex influences face categorization. *Nature*. 442:692–695.
- Allman JM, Kaas JH. 1974. The organization of the second visual area (V II) in the owl monkey: a second order transformation of the visual hemifield. *Brain Res*. 76:247–265.
- Blackmore 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.
- Bowman EM, Aigner TG, Richmond BJ. 1996. Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J Neurophysiol*. 75:1061–1073.
- Buckley MJ, Gaffan D, Murray EA. 1997. Functional double dissociation between two inferior temporal cortical areas: perirhinal cortex versus middle temporal gyrus. *J Neurophysiol*. 77:587–598.
- 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.
- Buffalo EA, Ramus SJ, Squire LR, Zola SM. 2000. Perception and recognition memory in monkeys following lesions of area TE and perirhinal cortex. *Learn Mem*. 7:375–382.
- Butter CM, Gekoski WL. 1966. Alterations in pattern equivalence following inferotemporal and lateral striate lesions in rhesus monkeys. *J Comp Physiol Psychol*. 61:309–312.
- Cadiou CF, Hong H, Yamins DLK, Pinto N, Ardila D, Solomon EA. 2014. Deep neural networks rival the representation of primate IT cortex for Core visual object recognition. *PLoS Comput Biol*. 10:e1003963.
- Cowey A, Gross CG. 1970. Effects of foveal prestriate and inferotemporal lesions on visual discrimination by rhesus monkeys. *Exp Brain Res*. 11:128–144.
- Distler C, Boussaoud D, Desimone R, Ungerleider LG. 1993. Cortical connections of inferior temporal area TEO in macaque monkeys. *J Comp Neurol*. 334:125–150.
- Eldridge MAG, Matsumoto N, Wittig JH Jr, Masseau EC, Saunders RC, Richmond BJ. 2018. Perceptual processing in the ventral visual stream requires area TE but not rhinal cortex. *Elife*. 7:1–16.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2001. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*. 291:312–316.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2002. Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. *J Neurophysiol*. 88:929–941.
- Fujita I, Tanaka K, Ito M, Cheng K. 1992. Columns for visual features of objects in monkey inferotemporal cortex. *Nature*. 360:343–346.
- Gainotti G. 2000. What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: a review. *Cortex*. 36:539–559.
- Gross CG, Rocha-Miranda CE, Bender DB. 1972. Visual properties of neurons in inferotemporal cortex of the macaque. *J Neurophysiol*. 35:96–111.
- Hays AV, Richmond BJ, Optican LM. 1982. Unix-based multiple-process system, for real-time data acquisition and control. *WESCON Conf Proc.*, 2:1–10.
- Hubel DH, Wiesel TN. 1959. Receptive fields of single neurones in the cat's striate cortex. *J Physiol*. 143:574–591.
- Iwai E, Mishkin M. 1968. Two visual foci in the temporal lobe of monkeys. In: *Neurophysiological basis of learning and behavior*.

- Eds.: N. Yoshii and N.A. Buchwald. 1969. Further evidence on the locus of the visual area in the temporal lobe of the monkey. *Exp. Neurol.* Japan: Osaka Univ. Press 1968. 25:585–594.
- Kar K, Kubilius J, Schmid K, Issa EB, DiCarlo JJ. 2019. Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nat Neurosci.* 22:974–983.
- 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.
- 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.
- Majaj NJ, Hong H, Solomon EA, DiCarlo JJ. 2015. Simple learned weighted sums of inferior temporal neuronal firing rates accurately predict human core object recognition performance. *J. Neurosci.* 39:30–35.
- Matsumoto N, Eldridge MAG, Saunders RC, Reoli R, Richmond BJ. 2016. Mild perceptual categorization deficits follow bilateral removal of anterior inferior temporal cortex in rhesus monkeys. *J Neurosci.* 36:43–53.
- Minamimoto T, Saunders RC, Richmond BJ. 2010. Monkeys quickly learn and generalize visual categories without lateral prefrontal cortex. *Neuron.* 66:501–507.
- Sato T, Uchida G, Lescroart MD, Kitazono J, Okada M, Tanifuji M. 2013. Object representation in inferior temporal cortex is organized hierarchically in a mosaic-like structure. *J Neurosci.* 33:16642–16656.
- Sigala N, Logothetis NK. 2002. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature.* 415:318–320.
- Tanaka K. 1996. Inferotemporal cortex and object vision. *Annu Rev Neurosci.* 19:109–139.
- Ungerleider LG, Galkin TW, Desimone R, Gattass R. 2008. Cortical connections of area V4 in the macaque. *Cereb Cortex.* 18:477–499.
- Ungerleider LG, Mishkin M. 1982. Two cortical. In: Systems V, Ingle DJ, Goodale MA, Mansfield RJW, editors. *Analysis of visual behavior.* Cambridge (MA): MIT Press, pp. 549–586.
- Vogels R. 1999. Categorization of complex visual images by rhesus monkeys. Part 1: behavioural study. *Eur J Neurosci.* 11:1223–1238.
- Vogels R, Saunders RC, Orban GA. 1997. Effects of inferior temporal lesions on two types of orientation discrimination in the macaque monkey. *Eur J Neurosci.* 9:229–245.
- Weiskrantz L, Saunders RC. 1984. Impairments of visual object transforms in monkeys. *Brain.* 107:1033–1072.