OXFORD

Visual recognition in rhesus monkeys requires area TE but not TEO

Mark A.G. Eldridge [1,*], Jonah E. Pearl[1,2], Grace P. Fomani[1], Evan C. Masseau[1], J. Megan Fredericks[3], Gang Chen[4], Barry J. Richmond[1]

[1]Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, United States,
[2]Present address: Department of Neurobiology, Harvard Medical School, Boston, MA 02115, United States.,
[3]Present address: Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY 10014, United States.,
[4]Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, United States
*Corresponding author: Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA.
Email: Mark.Eldridge@nih.gov (Mark A.G. Eldridge)

Key resources table: KRT5d7a5dab1d4b4

The primate visual system is often described as a hierarchical feature-conjunction pathway, whereby each level represents an increasingly complex combination of image elements, culminating in the representation of whole coherent images in anterior inferior temporal cortex. Although many models of the ventral visual stream emphasize serial feedforward processing (Poggio et al. 2012; Yamins and DiCarlo 2016) anatomical studies show connections that bypass intermediate areas and that feedback to preceding areas (Distler et al. 1993; Kravitz et al. 2011). Prior studies on visual discrimination and object transforms also provide evidence against a strictly feed-forward serial transfer of information between adjacent areas (Kikuchi and Iwai 1980; Weiskrantz and Saunders 1984; Kar and DiCarlo 2021). Thus, we sought to investigate whether behaviorally relevant propagation of visual information is as strictly sequential as sometimes supposed. We compared the accuracy of visual recognition after selective removal of specific subregions of inferior temporal cortex—area TEO, area TE, or both areas combined. Removal of TEO alone had no detectable effect on recognition memory, whereas removal of TE alone produced a large and significant impairment. Combined removal of both areas created no additional deficit relative to removal of TE alone. Thus, area TE is critical for rapid visual object recognition, and detailed image-level visual information can reach area TE via a route other than through TEO.

Key words: object recognition; inferior temporal cortex; ventral visual stream; aspiration lesion; *Macaca mulatta*.

Introduction

In primates, perception of object identity requires neural processing in the ventral visual stream—a set of interconnected cortical areas stretching from primary visual cortex to inferior temporal cortex (IT) (Ungerleider and Mishkin 1982; Mishkin et al. 1983; Grill-Spector and Weiner 2014; Lafer-Sousa et al. 2016). IT comprises 2 architectonically separable regions: a caudal area—TEO, and a rostral area—TE (Von Bonin and Bailey 1947). We recently demonstrated that categorization based on visual similarity is most significantly impaired by combined removals of TEO and TE; bilateral removal of either area independently resulted in only a modest deficit (Setogawa et al. 2021). Perirhinal cortex, the region rostro-medially adjacent to—and hence considered "downstream" of—TE, may play a role in some forms of object-level perception (Buckley et al. 2001; Bussey et al. 2003; Baxter 2009) but see (Suzuki 2009). However, we have demonstrated that perirhinal cortex is not necessary for categorization based on visual similarity as

measured in our tasks (Eldridge et al. 2018), and hence it will not be further considered here. It has been suggested that perceptual categorization is an early post-sensory component of object recognition (Warrington 1982). To investigate whether the impairments in categorization observed in monkeys with TEO and/or TE removals would be mirrored in tests of recognition memory, we examined visual object recognition behavior after bilateral surgical removals of TEO alone, TE alone, or both areas combined. In this study, normal visual recognition required area TE to be intact. Removal of area TEO produced no functional impairment.

Materials and methods
Subject information

Subjects were 11 adult monkeys (*M. mulatta*), ages 6 to 12. Detailed information can be found in tabular form in Supplemental Table 2-1. Three monkeys served as

unoperated controls (Monkeys P, E, and Tc). Three monkeys received bilateral aspiration removals of area TE (Monkeys K, Gt [referred to as G in Matsumoto et al. 2016, renamed Gt here to distinguish from Monkey G below], and T; 3, 3, 3.5 years prior to study, respectively), the reconstructions of these lesions have been reported previously (Matsumoto et al. 2016). Two monkeys received bilateral aspiration removals of area TEO (Monkeys Y and S; 0.5, 0.5 years prior to study, respectively). Three monkeys received bilateral aspiration removals of regions TE and TEO (Monkey G, L, and M; 1, 1, and 1.5 years prior to the study, respectively). The reconstructions for the latter 2 groups have also been reported previously (Setogawa et al. 2021).

Experimental procedures were approved by the Animal Care and Use Committee of the National Institute of Mental Health through an Animal Study Proposal and followed regulations from the *ILAR Guide for the Care and Use of Laboratory Animals*.

Surgical procedures

Surgeries were carried out in a veterinary operating facility using aseptic techniques. Animals were sedated with ketamine hydrochloride (10 mg/kg, intramuscular), and anesthesia was maintained on isoflurane to effect. Body temperature, heart rate, blood pressure, SpO_2 and expired CO2 were monitored continuously.

The intended extent of the TE and TEO lesions, respectively, have been described previously (TE, Matsumoto et al. 2016; TEO, Huxlin 2000; see Fig. 1A for illustration). TEO lesions were intended to extend rostro-caudally from an imaginary line perpendicular to the superior temporal sulcus (STS) and tangent to the inferior occipital sulcus (IOS), to a line 1 cm rostral and parallel to the first, and on the dorsal–ventral axis from the fundus of the STS to the fundus of the occipitotemporal sulcus (OTS). TE lesions extended rostrally from the rostral boundary of area TEO to an imaginary line connecting the rostral tip of STS with the rostral tip of AMTS, and were bounded dorso-medially by the fundus of the STS and ventro-medially by either the fundus of the OTS (caudally), an imaginary line extending from the rostral tip of OTS to the posterior tip of AMTS, or the medial bank of AMTS (rostrally). Overlap of actual lesions is illustrated in Fig. 2. Following surgeries, monkeys recovered for 14 days before testing resumed.

Behavioral testing

Monkeys sat in a primate chair in a darkened, sound-attenuated room. Monkeys received fluid rewards through a sipper tube attached to the primate chair. The flow of fluid was controlled by an actuator outside the testing room. Monkeys viewed images on LCD monitors (Samsung 2233RZ, 40° x 30° visual angle) 57 cm from their eyes (Wang and Nikolic, 2011). Behavioral tasks were run and timed with custom-written code (Real-time Experimentation and Control, REX) (Hays, 1982), and images were displayed using commercial software

(Presentation, Neurobehvioral Systems). Experimental cues (0.5° x 0.5° visual angle) and images (10° x 10° visual angle) were displayed on a static background of white noise (random black and white pixels). Monkeys responded by depressing and releasing a deflection-sensitive touch bar.

Pretraining

Monkeys were initially trained to depress and release the bar to receive liquid rewards with a "red-green" discrimination task (Bowman 1996). Monkeys were required to press and hold the bar through a "no-go" period (indicated by the presence of a red square, 0.5° visual angle), then release in a "go" period (green square, 0.5°). Trials were self-initiated by bar touch, and the no-go period lasted 2–3 s. Correct responses, defined as those where the monkey released the bar between 0.2 and 1 s after the appearance of the go cue, resulted in reward delivery. All other releases caused the cue to disappear and immediately ended the trial. All trials were followed by a 1 s intertrial interval regardless of outcome. Once proficient at the training task, defined as reaching 85% correct responses for 2 consecutive days, monkeys were moved on to the recognition task.

Serial recognition task

Monkeys were required to discriminate between the first and second presentations of never-before-seen or rarely seen images. On each trial, they indicated whether the image was being presented for the first or second time in the testing session by releasing the touch bar in 1 of 2 time periods (previously the "no-go" and "go" periods described above). This is equivalent to a cued 2-interval forced-choice task and is similar to those used in studies from other groups (Xiang and Brown 1998; Meyer and Rust 2018). In our version of the task, the number of intervening trials between first and second presentations varied as the experiment progressed, beginning in training with no intervening trials (i.e. AABBCC...), and increasing the number of intervening unique image trials in blocks when performance reached criterion (first day of 75% accuracy). See "Session Structure" for details on which intervals were introduced in which blocks. In the highest difficulty level of the task, we used intervals of 0, 1, 2, 4, 8, 16, 32, 64, or 128 other stimuli.

Trial and reward structure

The structure and timing of the trials was the same as the training task. We taught the monkeys to release in the red cue period for first presentations, and to release in the green period for second presentations. The trial structure is schematized in Fig. 1C.

Trials were self-initiated, and the test image appeared concurrently with the red cue. The cue remained red for 1–3 s. A release in this period, no matter whether it was a first or second presentation, caused the trial to end; the cue and image disappeared, and the intertrial interval
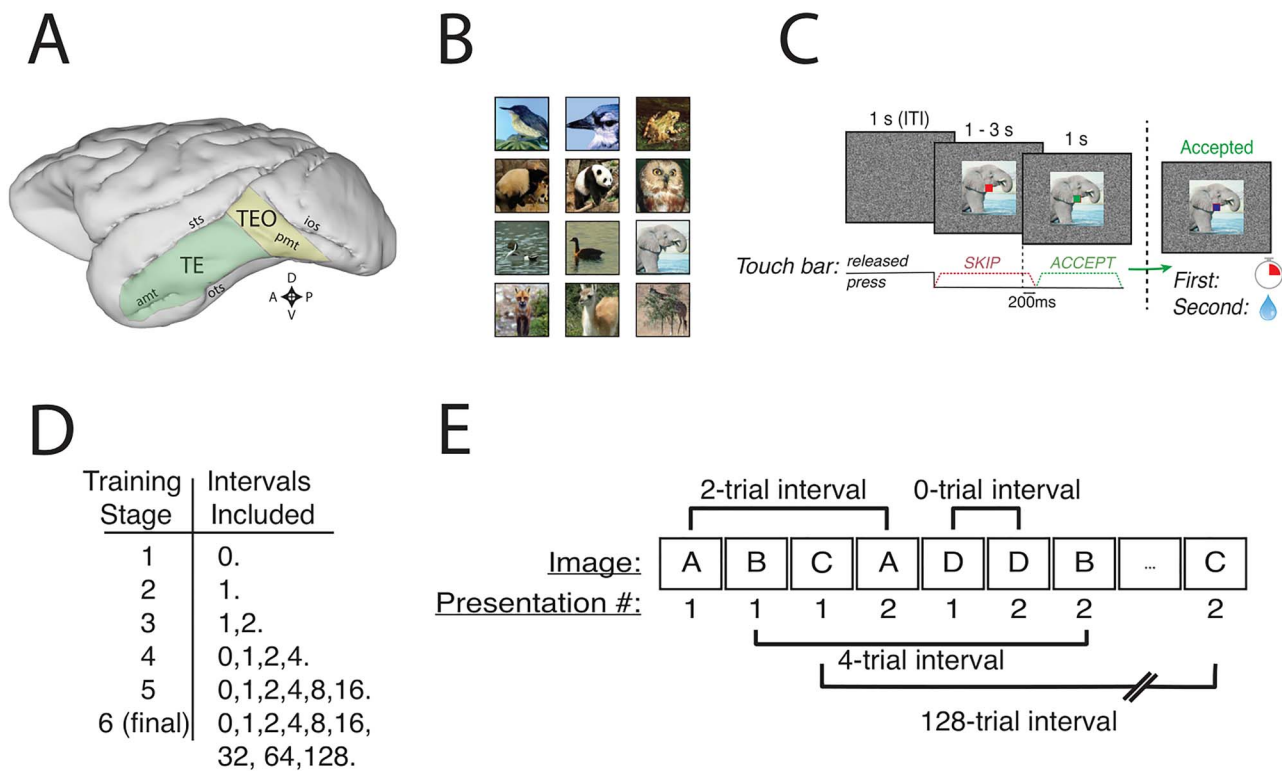
Fig. 1. Methods and serial recognition task. A) Locations of cytoarchitectonic regions TE and TEO on the lateral surface of the macaque brain. Sts, superior temporal sulcus; ios, inferior occipital sulcus; amt, anterior middle temporal sulcus; pmt, posterior middle temporal sulcus; ots, occipito-temporal sulcus. Compass rose: D, dorsal; V, ventral; A, anterior; P, posterior. B) Examples of images used (shown without an overlaid cue). C) Structure of a single trial in the serial recognition task. Trials began with a bar touch and the outcome depended on when the monkey released the bar. Release in the red period always resulted in a skipped trial, extinguishing the image and beginning the ITI. Release in the green period resulted in a timeout for first presentations and reward delivery for second presentations. Failure to release during green ended the trial and began the ITI. The reward-maximizing strategy is to release during red for first presentations and during green for second presentations. D) Blocks of intervals used in training for the serial recognition task. Monkeys progressed to the next block upon either 1) reaching 75% accuracy on the previous block or 2) going 10 sessions with approximately stable performance, indicating they were no longer learning from the training. There were minor deviations from this training regimen; see Training notes in Methods. E) Example sequence of trials at the beginning of a task session at the hardest difficulty in the serial recognition task. Distinct images are represented by letters.

began. When the monkey held the bar for the full scheduled time, the cue turned green. Releases between 0.2 and 1 s after the appearance of the green cue were scored, indicated to the monkey by the cue turning blue. If the image was a first presentation, the monkey invoked a 6–12 s error timeout (calibrated for each monkey), whereas if the image was a second presentation, the monkey earned a liquid reward. Timeouts were calibrated on a session-by-session basis to maintain the response bias to within $\sim 20\%$ of neutral. Early release in the green period or failure to release the bar caused the trial to end. All trials, regardless of outcome, were followed by a 1 s intertrial interval. Monkeys were not required to fixate on the images.

Images
Images were drawn from open-source internet repositories (examples are show in Fig. 1B and Supplemental Fig. 1-1). Images were full color, natural images of animals, such as birds, ducks, felines, deer, squirrels, horses, and so forth, predominantly in natural scenes. Some images resembled each other closely, even within a single testing session's image set. Animals were roughly centered in the images. A total library of 5,800 images was used. No attempt was made to normalize the contrast or luminance of the images.

Session structure
Each session consisted of 200 images, each presented twice in a pesudorandomized fashion (schematized in Fig. 1E), for a total of 400 trials. There was a sufficient number of images in the library such that novel stimuli could be used for the first 29 sessions without repeat. After 29 sessions, the same image sets were reused beginning with set 1, set 2, and so on, always randomizing the order of images' first presentations and an image's associated interval. For monkeys with TEO and TE + TEO removals, intervals were randomly interleaved each day to create a new sequence of correct responses. For monkeys in the control group as well as those with TE removals, a single random ordering of correct responses was created for each interval set, and it was reused every day with a new stimulus set. There was no evidence that this repetition of response order provided an advantage.

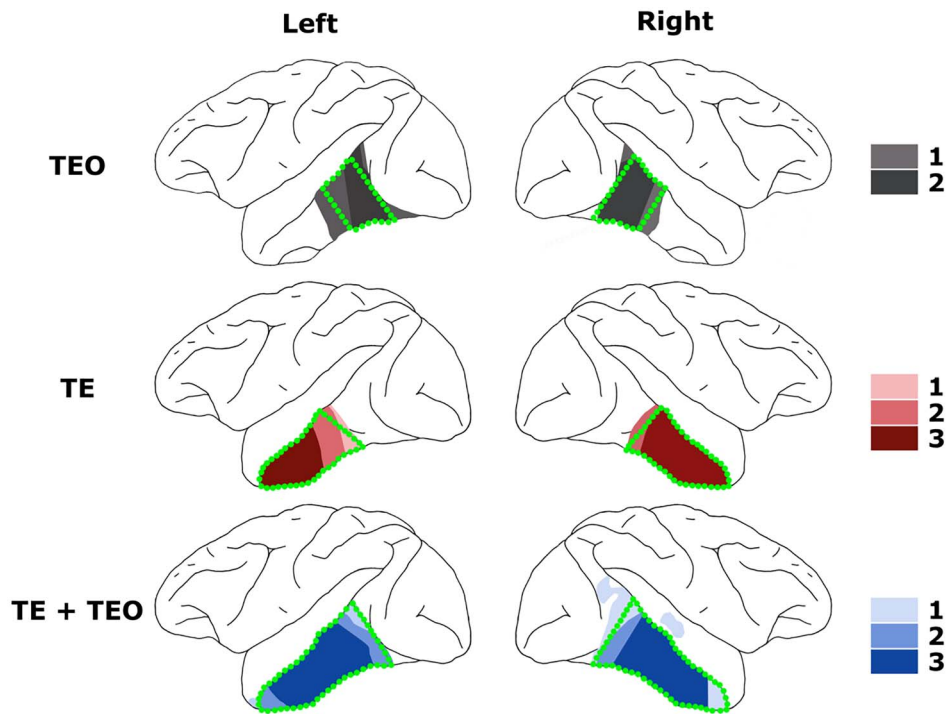Training began with an easier version of the task, in which the same image was presented 2 times in a row,

Fig. 2. Lesion reconstructions. Lesion overlay maps for the respective treatment groups on a lateral view of the rhesus brain. Darker shading indicates greater within-group lesion overlap—The number of monkeys represented by each color is indicated in the legend. Green broken lines illustrate intended lesions. Detailed lesion reconstructions for individual monkeys have been published previously (Eldridge et al. 2018; Setogawa et al. 2021).

i.e. trials followed the pattern AABBCCDD.... In this case, the interval between successive presentations of an image is always zero, so we denote this stage{0}. The next stage was {1}, i.e. ABAB-CDCD-EFEF... Subsequent stages used the interval sets {1, 2}, {0, 1, 2, 4}, {0, 1, 2, 4, 8, 16}, and {0, 1, 2, 4, 8, 16, 32, 64, 128}. Exceptions to this training sequence are noted below in "Training Notes."

Monkeys progressed to the next set of intervals after reaching 75% correct responses (averaged across all trials) within a single session. In some cases, monkeys never reached criterion, in which case they were moved on to the next interval set after 10 stable sessions. After reaching the hardest difficulty level of the task, the experiment continued until performance approached stable levels for 10 days.

Training notes

Monkey K developed a severe bias towards releasing on red twice in a row, regardless of the information received from the task. Because of this, training was restarted, again at {1}. The subsequent interval blocks were {1}, {0,1}, {0,1,2}, {0,1,2,4}, {0,1,2,4,8,16}, and {0,1,2,4,8,16,32,64,128}. Monkey K's performance exceeded 75% on the {0,1} stage, indicating it at least understood the mechanics of the task. The data presented here are from this second training, with days to criterion counted from the restart date. The monkey did not appear to gain any advantage from having performed the training steps twice. The following inconsistences in training were due to experimental error: in the group with TE + TEO removals, monkey M only tested for 2

days on the set {0,1,2,4,8,16}, reaching 70% correct, before moving on to the task's hardest difficulty level. Monkey G did not test on the {0,1,2,4,8,16} set at all. In the group with TEO removals, both monkeys did not test on the training set {1}, moving directly from {0} to {1,2}. Despite this, their performance quickly matched that of controls.

Data analysis
Generalized linear mixed-effects model
To assess task performance, we modeled the fraction of correct trials in the stable sessions as a function of experimental group using a generalized linear mixed-effects model. The model used a binomial distribution (logit link) with a fixed effect of group and a random effect intercept term for subject (monkey), which was nested within group.

Signal detection
In this serial recognition experiment, we denote first presentations as noise trials, and second presentations as signal trials. We calculated d' values for each interval in our task as $d_i' = Z(TP_i) - Z(FA)$, where:

Z is the inverse of the cumulative normal distribution function;

$$TP_i = \frac{\sum_{t=1}^{N_i} I\left(\text{Release on green in trial } t?\right)}{N_i},$$ the True Positive, or "hit" rate for second-presentation images presented i trials after their first presentation, where t indexes trials up to N_i, the total number of second presentations with

interval i, and $I(\cdot)$ is an indicator function, i.e. 1 if the monkey releases on green, 0 if on red;

and $FA = \frac{\sum_{t=1}^{N_1} I\left(\text{Release on green in trial } t?\right)}{N_1}$, the false-alarm rate for all first-presentation images, where t indexes trials up to N_1, the total number of first presentations (i.e. half of all trials), and $I(\cdot)$ is the same indicator function. Note that the FA rate is the same across all intervals for each session, since a first presentation has no interval associated with it.

Occasionally, monkeys had TP rates of 1 at short intervals, i.e. 100% correct, leading to a theoretical d' of infinity. To avoid this, we defined $Z(1) := Z[(2N - 1)/2N]$, where N is the number of total trials for the rate in question. (No monkeys achieved FA rates of 0, obviating the need to define $Z(0)$.). Changing the definition of $Z(1)$ to $Z(1) := Z[(N - 1)/N]$, a more conservative estimate, did not alter our statistical results, implying our definition of $Z(1)$ did not lead to an overestimate of d'. The range of possible d' values in the hardest difficulty level of the task, given this definition and a maximum N of 22, is $[-3.69, 3.69]$.

Bayesian multilevel modeling of group differences in d'

Using d' as an experimental measure presents statistical difficulties because it is estimated for each monkey and each testing session using a finite number of trials, and thus there is a standard error associated with each d' value (Wickens 2001). This error is given by

$$\widehat{std}\left(\hat{d}'\right) = \sqrt{\frac{\widehat{var}(f)}{\varphi^2\left(\hat{\lambda}\right)} + \frac{\widehat{var}(h)}{\varphi^2\left(\hat{d}' - \hat{\lambda}\right)}}$$

where f is the false alarm rate FA, h is the true positive (hit) rate TP, φ is the probability density function for the standard normal distribution, and

$$\lambda = -Z(f),$$
$$\widehat{var}(f) = \frac{f(1-f)}{N_n}$$
$$\widehat{var}(h) = \frac{h(1-h)}{N_s}$$

where Z is the inverse cumulative normal distribution, N_n is the number of noise trials, and N_s is the number of signal trials ((Wickens 2001) p. 202). In general, the more extreme the d' value, the larger the error. In order to incorporate this standard error into a statistical model, it was necessary to eschew typical analysis of variance (ANOVA)-type models, which assume independent and identically distributed measurements with unknown normal, independent and identically distributed error. Instead, we adopted a Bayesian multilevel (BML) framework to estimate the posterior distribution for each group's d' by decomposing the response variable (d') as follows:

$$y_{ijk(l)} = g_l + \alpha_i + \beta_{jl} + \gamma_k + \varepsilon_{ijk(l)}$$

where indices i, j, k, and l code for day, interval, monkey and group, $k(l)$ shows the nested structure in which kth monkey belongs to lth group ($i = 1, 2, \ldots, 10$; $j = 1, 2, \ldots 9$ (i.e. intervals 0, 1, ... 128); $k = 1, 2, \ldots, 11$ (i.e. individual monkeys); $l = 1, 2, 3$ (i.e. groups with TE, TEO, or TE + TEO removals). g_l represents the effect of lth group at the population level. Distributional assumptions for the parameters at the lower hierarchical levels are

$$\alpha_i \sim N\left(0, \xi^2\right), \left(\beta_{j1}, \beta_{j2}, \beta_{j3}, \beta_{j4}\right)^T \sim N\left(\mathbf{0}_{4\times 1}, \mathbf{\Sigma}_{4\times 4}\right),$$
$$\gamma_k \sim N\left(0, \tau^2\right), \varepsilon_{ijk(l)} \sim N\left(0, s_{ijk(l)}^2\right),$$

where ξ^2, τ^2, and $s_{ijk(l)}^2$ are the variances for cross-day, cross-monkey effects and measurement error, while $\mathbf{\Sigma}$ is the variance–covariance matrix among the 4 groups across the 9 intervals. The standard deviation for the measurement error, $s_{ijk(l)}$, is assigned the standard error associated with d' from above. The BML model was numerically solved through Markov chain Monte Carlo simulations in Stan with the hyperpriors for the model parameters (g_l, ξ^2, τ^2, $\mathbf{\Sigma}$) assigned through general recommendations (Albasser et al. 2011; Bürkner 2017; Carpenter et al. 2017).

The effects of interest were assessed as follows. The overall effect of each group, g_l, and the effect of each group at a particular interval, $g_l + \beta_{jl}$, were inferred through the posterior distribution based on the samples randomly drawn from the simulated chains. The resulting posterior distributions allow one to ask how likely one group's mean d' is different than another's by further assessing the posterior distribution of each group difference in d'. Specifically, we accept 2 groups as being strongly different if the strength of statistical evidence is 95% or greater, and we assign the strength of statistical evidence by quantifying the posterior probability of the group difference being positive or negative. For example, for the difference $Z = X - Y$, the posterior probabilities of $P(Z > 0) < 0.05$ or $P(Z > 0) > 0.95$ correspond to the strong differences $Y > X$ and $X > Y$, respectively.

Determination of stable performance

We determined stable performance on the hardest version of the task (i.e. the version that included a spacing of up to 128 intervening trials between first and second presentations) as the first 10 sessions in which the monkeys' smoothed daily rate of improvement was no more than 1.5% correct per day. We calculated the smoothed rates of improvement as the slope of a linear regression line fit to a moving window of 5 days of data. Monkeys Tc and Gt had 9 and 8 sessions that fit the rule, respectively, so we took 1 and 2 sessions further back, respectively. Note that we did not take into account whether monkeys had reached criterion, 75% correct, in determining stable performance.
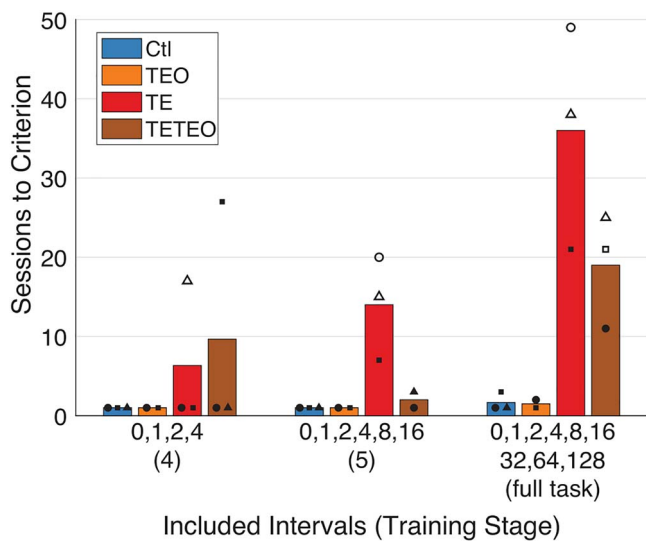
Fig. 3. TE and TE + TEO removals slow task learning. Monkeys with TE and TE + TEO removals take longer than controls and monkeys with TEO removals to reach criterion (75% correct overall) on the final 3 training sets. Within groups, each monkey's data are plotted as a different shape. A solid shape indicates the monkey reached criterion on that set, whereas a hollow shape indicates the monkey's performance plateaued below criterion. Bars represent group means, including monkeys that did not reach criterion.

Results

During the first 4 training stages (up to 4 intervening images), there were no group-level differences in the number of sessions monkeys took to reach a fixed performance threshold of 75% correct ($P > 0.05$, one-sided rank-sum test). Although the first and second stages had very simple response patterns (1st–2nd–1st–2nd . . . and 1st–1st–2nd–2nd . . . , respectively), the third and fourth stages did not. Thus, it appears that all groups were able to understand the rule for discriminating first from second presentations.

At the highest difficulty level, i.e. when the longest maximum interval between first and second presentations (128 images) was included, control monkeys and those with TEO removals reached criterion within 3 days of testing (Fig. 3). Two monkeys with TE or TE + TEO removals required more than 10 days to reach criterion, and the other 4 failed to reach criterion at all, despite 3–8 weeks of additional testing.

TE and TE + TEO removals impair recognition memory

We evaluated recognition memory at the highest difficulty level of the task using 10 sessions of stable performance from each monkey. Stable performance was defined as the first 10 days during which the smoothed rate of improvement was always less than 1.5% correct trials per day. There were significant group-level differences in task performance during these stable sessions when performance was summarized across all interpresentation intervals (Fig. 4A). Groups with TE or TE + TEO removals performed significantly less well than controls (GLM, TE $P = 0.00031$, TE + TEO $P = 0.0013$), whereas the
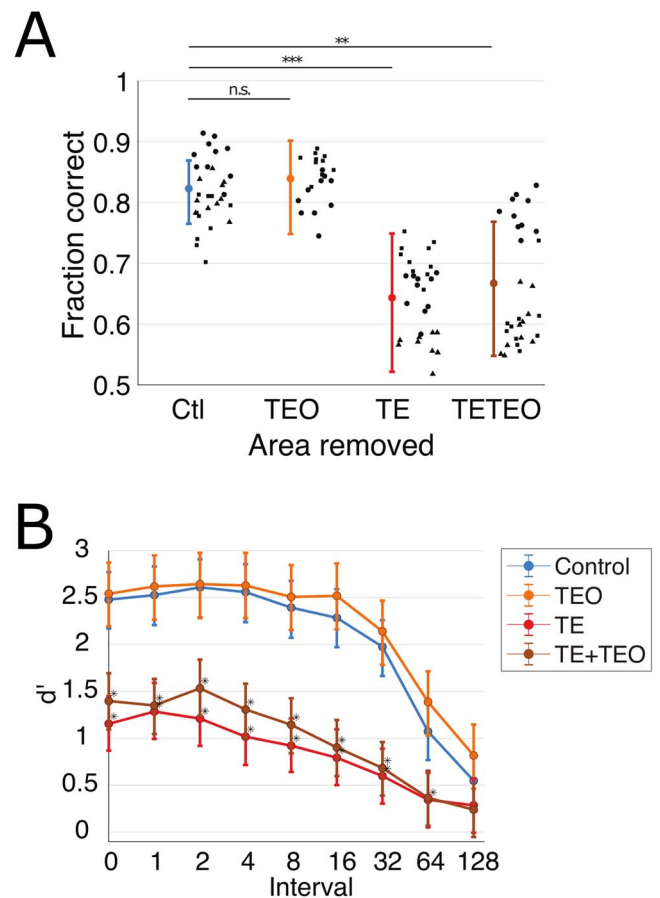

Fig. 4. Monkeys with TE and TE + TEO removals have impaired recognition relative to controls and those with TEO removals. A) Overall fraction correct trials for the 10 sessions of stable performance on the hardest difficulty level of the task. Bars indicate mean and 95% confidence intervals from GLM. Points indicate overall percent correct for individual sessions, different shapes within each group correspond to individual subjects (matched with Supplemental Fig. 4-1). B) Group-mean *d'* as a function of interval. Bars indicate median and 68% confidence intervals (analogous to mean and SEM) from the Bayesian multilevel model. Asterisks indicate strong evidence for *d'* values differing from control values (strength of evidence > 95%, see Methods).

performance of the group with TEO removals was indistinguishable from controls ($P = 0.68$).

To quantify the relationship between interpresentation interval and recognition ability, we calculated a *d'* value separately for each interval in these 10 stable sessions (see Methods). All groups had larger *d'* at shorter rather than longer intervals (Fig. 4B). The mean performance of all of the individual monkeys was above chance ($d' > 0$) for all of the intervals tested, again indicating that they learned the rules of the task (Supplemental Fig. 4-1).

To assess group-level differences of *d'*, we used a Bayesian multilevel model, which accounts for the uncertainty inherent in the *d'* measurements by incorporating their standard errors into group-level probability distributions (see Methods and raw statistical output in Supplemental File 1) (Chen et al. 2019). The posterior distributions are compared to assess the likelihood that one group's value will be larger or smaller than another group's value. There is strong evidence for a

Table 1. Statistical results from Bayesian multilevel modeling on data from the 10 days of stable performance. The Bayesian multilevel model accounts for the uncertainty associated with measuring d' by incorporating the standard error into group-level probability distributions. The posterior distributions are compared to assess the likelihood that one group's value will be larger or smaller than another group's value. There is strong evidence for a group-level difference if the posterior probability of the group difference being positive or negative is 95% or greater.

| | Bayesian multilevel modeling | | | |
| | P(group d' < control d') | | | P(TE + TEO d' < TE d') |
Interval	TEO	TE	TE + TEO	TE vs TE + TEO
0	0.45	1.00	0.99	0.27
1	0.43	0.99	0.99	0.44
2	0.48	1.00	0.99	0.21
4	0.44	1.00	1.00	0.23
8	0.40	1.00	1.00	0.28
16	0.30	1.00	1.00	0.39
32	0.36	1.00	1.00	0.41
64	0.24	0.96	0.95	0.47
128	0.28	0.74	0.77	0.53

group-level difference if the posterior probability of the group difference being positive or negative is 95% or greater. Using this metric, the recognition detection of groups with TE or TE + TEO removals was impaired relative to that of controls, averaged across all interpresentation intervals (P(CTL > TE) = 99%, P(CTL > TE + TEO) = 98%). The group with TEO removals was unimpaired relative to controls (P(CTL > TEO) = 37%). Assessing the intervals individually, the groups with TE or TE + TEO removals were impaired relative to controls at all but the longest interval (Fig. 4B, Table 1). The standard error of the d' measurements were sufficiently small that these results did not qualitatively change when we compared group d' values using an ANOVA at each interval, corrected for multiple comparisons (not shown).

Response bias

Given the asymmetric reward structure of this task, which explicitly rewards correct identifications of second presentations but not of first presentations (see Fig. 1C), monkeys might be expected to develop biases towards indicating "second." All monkeys showed this bias in the initial stages of training (high false alarm rates, Supplemental Fig. 3-1), consistent with the monkeys exploring the new task parameters. As the monkeys' performance improved, the bias diminished. Nonetheless, at the most demanding level of the task, monkeys with the largest impairments in recognition performance (the groups with TE or TE + TEO lesions) still showed strong biases towards false alarms (incorrectly responding that a novel image was a second presentation). The fact that this bias emerged only later in the training suggests that it was likely a consequence of the difficulty in recognition

for the TE and TE + TEO lesion groups. Measurements of d' are invariant across changes in observer bias (Wickens 2001), so this bias does not affect the group-level statistics.

Discussion

If the integration of information in the ventral visual stream were strictly sequential, a lesion of area TEO would de-afferent area TE, leading to equal or greater deficits in object recognition than removal of TE alone. Alternatively, if TEO and TE function as a uniform "inferior temporal cortex" computational module, lesions of either area would be expected to cause partial impairments, perhaps proportional to their size. However, removal of TEO did not impair recognition in our task, whereas removal of TE, whether alone or combined with TEO, severely impaired monkeys' recognition memory. The monkeys with TEO + TE removals were no more greatly impaired than those with TE-only removals. Thus, recognition memory depends on area TE, and enough object-level visual information reaches TE without having to pass through area TEO such that object recognition, at least as tested here, is not impaired after TEO removal. This result contrasts with that of our previous study on visual categorization based on perceptual similarity, in which removal of either TEO or TE seemed to have been compensated for by the presence of the other. A large deficit in categorization was only observed when both areas were removed (Setogawa et al. 2021). Thus, the integration of information in the ventral visual stream is not necessarily strictly sequential and, although TEO and TE can function as if they are part of a single computational module in certain contexts (Setogawa et al. 2021), our data make it clear they cannot be generally treated as such.

The absence of an impairment in the TEO-removal group in the present study could be related to the difference in the volume of tissue removed, i.e. TE is a larger architectonic region, hence the deficits observed correspond to the quantity of tissue removed. However, we believe that the most parsimonious interpretation is that TE is required for recognition of stimuli with the level of complexity and feature overlap used in the present study, i.e. there is segregation of function between areas TEO and TE. The magnitude of impairments varied to some extent within lesion groups (Supplemental Fig. 4-1). Despite this variation, all 3 of the monkeys with TE removals performed worse than all 3 controls. Of the monkeys with combined TEO + TE removals, 2 performed much worse than the controls, and the third performed better than one of the control subjects. There was no consistent relationship between total lesion extent and relative magnitude of impairment. Overall, the dissociation seen in the pattern of behavioral effects observed in the present study, compared with those in our previous study on categorization, demonstrates that areas TE and TEO have their own functional specializations.

The level of feature similarity of the images used in a given test session was not parametrically controlled, and hence recognition judgments may not occur at the "object" level (i.e. scene, or object-in-scene, memory could be drawn upon).

The deficit in recognition memory observed after bilateral TE removal is consistent with a large body of literature reporting object-level recognition deficits after TE removals (Iwai and Mishkin 1969; Gross et al. 1972; Rolls et al. 1977; Eskandar et al. 1992; Buffalo et al. 1999). However, the absence of any deficit after TEO removals shows that TEO is not necessary for normal visual recognition memory. This finding appears to conflict with the results of previous studies on visual discriminations and transforms (Weiskrantz and Saunders 1984) and object recognition in a delayed non-match to sample (DNMS) task (Spiegler and Mishkin 1981), which both described partial impairments after removals of tissue in area TEO. Conversely, it is consistent with the lack of impairment in serial and concurrent visual discrimination tasks reported after bilateral TEO removals (Kikuchi and Iwai 1980), and with the observation that TE neurons retain their selectivity—in the absence of distractors—after removals of TEO (Buffalo et al. 2005). Where conflicts with the existing literature do exist, we speculate that these conflicts with prior behavioral observations may be explained by one or more of several differences in lesion extent and/or experimental design that do not relate directly to the load on recognition memory. For example, in the study here, large numbers of visual stimuli were presented on a computer monitor, whereas in the studies referred to above, smaller sets of physical objects were presented.

Considering the results from this study, and those from our previous studies with the same cohort of monkeys (Setogawa et al. 2021) together, there is enough evidence to conclude that TE receives visual information from brain regions other than TEO, and that that information is of high enough resolution to support normal recognition memory, and above-chance perceptual categorization. We cannot determine how the information reaches TE, whether from forward bypass connections or from feedback from other cortical or subcortical sites. One strong candidate source would be the bypass from V4 to TE (Desimone et al. 1980; Ungerleider et al. 2008). Whatever the source of visual information turns out to be, the dissociation in the pattern of behavioral effects observed in these 2 studies demonstrates that each of these inferior temporal areas makes a different contribution to image processing in the brain.

Acknowledgments

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

References

Albasser MM, Amin E, Iordanova MD, Brown MW, Pearce JM, Aggleton JP. Perirhinal cortex lesions uncover subsidiary systems in the rat for the detection of novel and familiar objects. *Eur J Neurosci*. 2011:34(2):331–342.

Baxter MG. Involvement of medial temporal lobe structures in memory and perception. *Neuron*. 2009:61(5):667–677.

Bowman EM, Aigner TG, Richmond BJ. Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *Journal of Neurophysiology*. 1996:75(5):1061–1073.

Buckley MJ, Booth MCA, Rolls ET, Gaffan D. Selective perceptual impairments after perirhinal cortex ablation. *J Neurosci*. 2001:21(24):9824–9836.

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

Buffalo EA, Bertini G, Ungerleider LG, Desimone R. Impaired filtering of distracter stimuli by TE neurons following V4 and TEO lesions in macaques. *Cerebral Cortex (New York, NY: 1991)*. 2005:15:141–151.

Bürkner P-C. Brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw*. 2017:80(1):1–28.

Bussey TJ, Saksida LM, Murray EA. Impairments in visual discrimination after perirhinal cortex lesions: testing 'declarative' vs. 'perceptual-mnemonic' views of perirhinal cortex function. *Eur J Neurosci*. 2003:17(3):649–660.

Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A. Stan: a probabilistic programming language. *J Stat Softw*. 2017:76(1):1–32.

Chen G, Xiao Y, Taylor PA, Rajendra JK, Riggins T, Geng F, Redcay E, Cox RW. Handling multiplicity in neuroimaging through Bayesian lenses with multilevel modeling. *Neuroinformatics*. 2019:17(4):515–545.

Desimone R, Fleming J, Gross CG. Prestriate afferents to inferior temporal cortex: an HRP study. *Brain Res*. 1980:184(1):41–55.

Distler C, Boussaoud D, Desimone R, Ungerleider LG. Cortical connections of inferior temporal area TEO in macaque monkeys. *J Comp Neurol*. 1993:334(1):125–150.

Eldridge MAG, Matsumoto N, Wittig JHJ, Masseau EC, Saunders RC, Richmond BJ. Perceptual processing in the ventral visual stream requires area TE but not rhinal cortex. *elife*. 2018:7:e36310.

Eskandar EN, Richmond BJ, Optican LM. Role of inferior temporal neurons in visual memory. I. Temporal encoding of information

about visual images, recalled images, and behavioral context. *J Neurophysiol*. 1992:68(4):1277–1295.

Grill-Spector K, Weiner KS. The functional architecture of the ventral temporal cortex and its role in categorization. *Nat Rev Neurosci*. 2014:15(8):536–548.

Gross CG, Rocha-Miranda CE, Bender DB. Visual properties of neurons in inferotemporal cortex of the macaque. *J Neurophysiol*. 1972:35(1):96–111.

Hays AV, Richmond BJ, Optican LM. A UNIX-Based Multiple-Process System for Real-Time Data Acquisition and Control. *WESCON Conference Proceedings*. 1982:2:1–10.

Huxlin KR, Saunders RC, Marchionini D, Pham HA, Merigan WH. Perceptual deficits after lesions of inferotemporal cortex in macaques. *Cereb Cortex*. 2000:10(7):671–83.

Iwai E, Mishkin M. Further evidence on the locus of the visual area in the temporal lobe of the monkey. *Exp Neurol*. 1969:25(4):585–594.

Kar K, DiCarlo JJ. Fast recurrent processing via ventrolateral prefrontal cortex is needed by the primate ventral stream for robust core visual object recognition. *Neuron*. 2021:109(1):164–176.e165.

Kikuchi R, Iwai E. The locus of the posterior subdivision of the inferotemporal visual learning area in the monkey. *Brain Res*. 1980:198(2):347–360.

Kravitz DJ, Saleem KS, Baker CI, Mishkin M. A new neural framework for visuospatial processing. *Nat Rev Neurosci*. 2011:12(4):217–230.

Lafer-Sousa R, Conway BR, Kanwisher NG. Color-biased regions of the ventral visual pathway lie between face- and place-selective regions in humans, as in macaques. *J Neurosci*. 2016:36(5):1682–1697.

Matsumoto N, Eldridge MA, Saunders RC, Reoli R, Richmond BJ. Mild perceptual categorization deficits follow bilateral removal of anterior inferior temporal cortex in rhesus monkeys. *J Neurosci*. 2016:36(1):43–53.

Meyer T, Rust NC. Single-exposure visual memory judgments are reflected in inferotemporal cortex. *elife*. 2018:7:e32259.

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

Poggio T, Mutch J, Leibo J, Rosasco L, Tacchetti A. The computationalmagic of the ventral stream: sketch of a theory (and why some deep architectures work). In: *TechRep MIT-CSAIL-TR-2012-035*. Cambridge, MA: MIT CSAIL; 2012.

Rolls ET, Judge SJ, Sanghera MK. Activity of neurones in the inferotemporal cortex of the alert monkey. *Brain Res*. 1977:130(2):229–238.

Setogawa T, Eldridge MAG, Fomani GP, Saunders RC, Richmond BJ. Contributions of the monkey inferior temporal areas TE and TEO to visual categorization. *Cereb Cortex*. 2021:31(11):4891–4900.

Spiegler BJ, Mishkin M. Evidence for the sequential participation of inferior temporal cortex and amygdala in the acquisition of stimulus–reward associations. *Behav Brain Res*. 1981:3(3):303–17.

Suzuki WA. Perception and the medial temporal lobe: evaluating the current evidence. *Neuron*. 2009:61(5):657–666.

Ungerleider LG, Mishkin M. Two cortical visual systems. In: Goodale M, Ingle DJ, Mansfield RJW, editors. *Analysis of visual behavior*. Cambridge, MA: MIT; 1982. pp. 549–586

Ungerleider LG, Galkin TW, Desimone R, Gattass R. Cortical connections of area V4 in the macaque. *Cereb Cortex*. 2008:18(3):477–499.

Von Bonin G, Bailey P. *The neocortex of Macaca mulatta*. Urbana, IL: University of Illinois press; 1947

Wang P, Nikolić D. An LCD monitor with sufficiently precise timing for research in vision. *Frontiers in Human Neuroscience*. 2011:5:85.

Warrington EK. Neuropsychological studies of object recognition. *Philos Trans R Soc Lond Ser B Biol Sci*. 1982:298(1089):15–33. https://doi.org/10.1098/rstb.1982.0069.

Weiskrantz L, Saunders RC. Impairments of visual object transforms in monkeys. *Brain*. 1984:107(4):1033–1072.

Wickens T. *Elementary signal detection theory*; 2001. Oxford University Press.

Xiang JZ, Brown MW. Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology*. 1998:37(4–5):657–676.

Yamins DLK, DiCarlo JJ. Eight open questions in the computational modeling of higher sensory cortex. *Curr Opin Neurobiol*. 2016:37:114–120.