Behavioral/Systems/Cognitive

Activity of Inferior Temporal Cortical Neurons Predicts Recognition Choice Behavior and Recognition Time during Visual Search

Ryan E. B. Mruczek and David L. Sheinberg

Department of Neuroscience, Brown University, Providence, Rhode Island 02906

Although the selectivity for complex stimuli exhibited by neurons in inferior temporal cortex is often taken as evidence of their role in visual perception, few studies have directly tested this hypothesis. Here, we sought to create a relatively natural task with few behavioral constraints to test whether activity in inferior temporal cortex neurons predicts whether or not a monkey will recognize and respond to a complex visual object. Monkeys were trained to freely view an array of images and report the presence of one of many possible target images previously associated with a hand response. On certain trials, the identity of the target was swapped during the monkeys' targeting saccade. Furthermore, the response association of the preswap target and the postswap target differed (e.g., right-to-left target swap). Neural activity in cells selective for the preswap target was significantly higher when the monkeys' response matched the hand association of the preswap target. Furthermore, the monkeys' response time was predicted by the magnitude of the presaccadic firing rate on nonswap trials. Our results provide additional support for the role of inferior temporal cortex in object recognition during natural behavior.

Key words: inferotemporal cortex; primate; neurophysiology; object recognition; choice probability; vision

Introduction

The striking selectivity of inferior temporal cortex (IT) neurons (Perrett et al., 1982; Desimone et al., 1984; Logothetis and Sheinberg, 1996; Tanaka, 1996) and the debilitating effects of temporal lesions on recognition performance (Gross, 1973; Dean, 1974; Mishkin, 1982) are generally taken as evidence that cells in IT are responsible for visual object recognition. However, the direct link between neural activity in IT and object recognition is often assumed but rarely demonstrated. This is because most studies of IT do not correlate neural response patterns with recognition behavior on a trial-by-trial basis. Thus, although the physiological responses of these cells are well characterized, their role in overt behavior is not as clearly established.

Typically, studies of IT using overt responses use paradigms in which monkeys perform at a ceiling level. Without the inclusion of near-threshold trials, monkeys do not exhibit response variability, making it impossible to correlate the activity of IT cells with recognition performance. Sheinberg and Logothetis (1997) used a binocular-rivalry paradigm to demonstrate such a correlation. During binocular rivalry, a different image is presented to each eye resulting in a competition between the two images for perceptual dominance; only one image is perceived at a time, and one's perception can switch back and forth, similar to the perceived configuration of a Necker cube. The authors found that IT cells responded more robustly when monkeys reported seeing an image that activated those same neurons in isolation.

The study of Sheinberg and Logothetis (1997) provided strong evidence that activity in IT is correlated with perception of visual objects. Here, we sought to use a relatively natural paradigm, one that does not rely on the use of inherently ambiguous stimuli, to test this same hypothesis. We wanted to clarify the role of IT in object recognition during normal behavior with the inclusion of self-initiated, visually guided eye movements. Sheinberg and Logothetis (2001) studied the response of IT neurons during visual search and found that response times were inversely correlated with the magnitude of an eccentricity-dependant early activation around the time of a targeting saccade. However, their task did not specifically control the amplitude of targeting saccades or the exact composition of the stimulus array, which could account for their response time–neural activation correlations.

We used a more structured search array that allowed us to induce natural, but stereotypical, patterns of self-guided eye movements and compared two informative behavioral measures from the same task (response choice and response time) with neural activity in single IT neurons. We created a situation in which monkeys could be on the verge of recognizing a peripherally presented target and compared trial-to-trial variability in behavioral and neural responses. In our experiment, the "correct" response was occasionally altered by discrete changes in object identity at specific times during a visual search task. On these

"target swap" trials, the monkeys' response informed us as to the time at which recognition occurred, before or after the stimulus change. We found that IT neurons selective for a saccade target respond more robustly when the monkeys' response indicates recognition of that same object. Furthermore, on trials without a target swap, presaccadic activity in IT is predictive of monkeys' response time. Both of these observations suggest that IT neurons play a critical role in object recognition during natural behavior.

Materials and Methods

Animals, surgery, and recording techniques. Two male rhesus monkeys (*Macaca mulatta*; monkey S and monkey M), ages 6 and 10 years and weighing between 9 and 12 kg, were the subjects in this study. Before the experiment, the monkeys had been familiarized with the behavioral apparatus and had participated in unrelated studies. Both monkeys had a recording chamber implanted over the left hemisphere (Horsley–Clark coordinates: +15 anterior, +20 lateral) and a titanium head post for head restraint. All surgeries were performed using sterile technique while the animals were intubated and anesthetized using isoflurane gas. All procedures conformed to the National Research Council *Guide for the Care and Use of Laboratory Animals* as well as the Brown University Institutional Animal Care and Use Committee. Both monkeys are currently participating in other experiments.

For monkey S, during each recording session, a guide tube (25 gauge) was inserted to a level just below the dura and a single electrode was advanced through the guide tube using a micropositioner (David Kopf Instruments, Tujunga, CA). For monkey M, a single electrode was advanced through a chronic guide tube using a micropositioner (David Kopf Instruments) or between one and four electrodes were advanced through the guide tube using a five-channel mini-matrix (Thomas Recording, Giessen, Germany). For the single-electrode setup, electrodes were composed of a tungsten core with glass coating (Alpha Omega, Jerusalem, Israel). Neural signals were amplified (model A-1; BAK Electronics, Germantown, MD), filtered [monkey S only; model 3364 (Krohn-Hite, Brockton, MA); 100 Hz to 12 kHz] and digitized at 34 kHz. For the multielectrode recordings, electrodes were quartz-coated tungsten/platinum core fibers. Signals were amplified and filtered (250 Hz to 8 kHz) and digitized at 34 kHz. For all experiments, single cells were isolated on-line using a threshold and two time-amplitude discrimination windows (custom software). However, spike times used for the analyses below were based on a similar off-line analysis run on the stored analog signal.

Eye movements were recorded using an EyeLink II video eye tracking system, running at 500 Hz (SR Research, Mississauga, Ontario, Canada). Note that, except for the initial fixation, there were no explicit constraints on the monkeys' eye movements during the time that the visual display was present. However, only trials that met specific saccade pattern criterion were used in the analyses (see below). The analog outputs from the eye tracking hardware were sampled by the control system at 1 kHz (see below), and a moving average was stored to disk every 5 ms (200 Hz). Saccades were automatically extracted from off-line eye records using a velocity-based algorithm written in C, which marked the start and end time, and start and end position for every saccade on each trial. The parameters of this algorithm were set to reliably detect saccades down to ~0.4° in amplitude.

Saccade-contingent changes in target identity (see below) were triggered by the eye trace (from the on-line 1 kHz samples) entering a circular region within 5.25° of the target. The time of swap completion was signaled through dedicated Ethernet and synchronized to the vertical
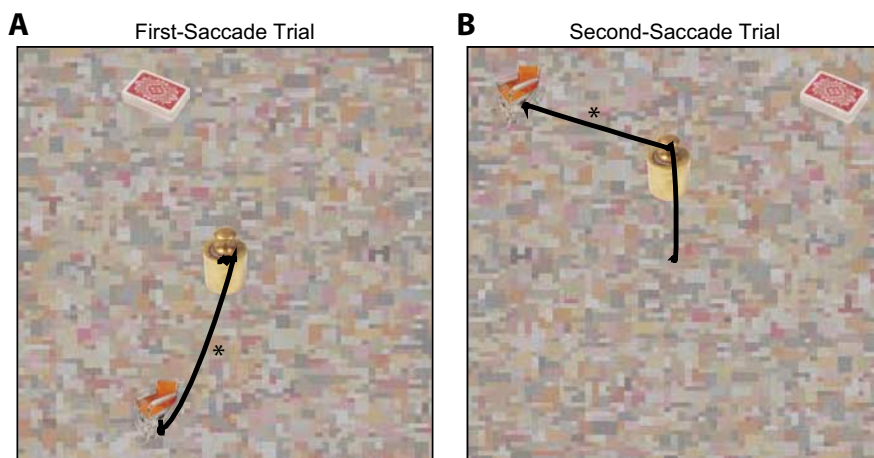

Figure 1. Example visual search displays from first-saccade (*A*) and second-saccade (*B*) trials. For both trial types, search arrays could appear in one of four directions: left (*A*), right, up (*B*), or down. In both panels, the target is the wheelchair and the weight and the cards are distractors. For trials included in the analyses, the target was always placed at one of the more eccentric positions. The black trace denotes eye position. The saccade pattern shows an initial saccade to or fixation of the closest distractor followed by a targeting saccade. This pattern was typical of all trials included in our analyses. The * indicates the targeting saccade, during which the target identity was changed on swap trials (not drawn to scale for clarity; for details, see Materials and Methods).

refresh of the display monitor. The accuracy of this signal was verified to be within 100 μs of the actual display swap in tests conducted before the experiment using a photodiode. To minimize the time between the swap request and its actual completion, the postswap display buffer was prepared and stored in a secondary video buffer just after the preswap buffer was displayed. The timing of the swap request (trigger) and the completion of the display change were stored to disc for off-line analysis. Display changes occurred, on average, within 1–11 ms (6 ms mean) of triggering the buffer swap.

Stimuli and display. Stimuli were presented on a dual-processor ×86 graphics workstation, running an OpenGL-based stimulation program under Windows XP (Microsoft, Seattle, WA). The screen resolution was 1024 × 768 with a vertical refresh rate of 100 Hz. Behavioral control for the experiments was maintained by a network of interconnected personal computers running the QNZ real-time operating system (QSSL, QNX Software Systems, Ottawa, Ontario, Canada). This system provides deterministic control and acquisition of button responses (with submillisecond precision) and eye position and communicates with the dedicated graphics machine using isolated high-speed Ethernet and direct digital input/output. Experimental control and data collection of behavioral measures was conducted using custom-written programs. All behavioral data, such as button responses and eye position signals, were available for on-line monitoring and stored to disk for off-line analysis.

The stimulus set contained 100 full color images of everyday objects (Hemera Photo Objects, Gatineau, Quebec, Canada). Objects subtended ~1.5 × 1.5°. Fifty images were used as targets, and the monkey was rewarded for pressing an associated button (left for 25 images, right for 25 images) whenever that image appeared on the display. The remaining 50 images did not have a response association and served as distractors. The monkey had extensive experience with both distractors and targets, and the response mappings for targets were consistent across day for each monkey. The images assigned as targets and distractors differed across monkey.

Monkeys were initially required to maintain fixation on a small spot for 350 ms after which the fixation spot was removed and the search array appeared. The search array appeared on a noise background that consisted of pseudorandomly colored squares (Fig. 1). Any stimulus presented at less than full contrast was blended with a homogeneous 50% gray background and not with the noise pattern. The color of the noise background was limited to the pixel colors of the four main experimental images for a given day (see below) and was displayed at 40% contrast. Individual squares were ~0.18° on a side, although neighboring squares could be the same color. The noise background made parafoveal identi-

A Normal Trials
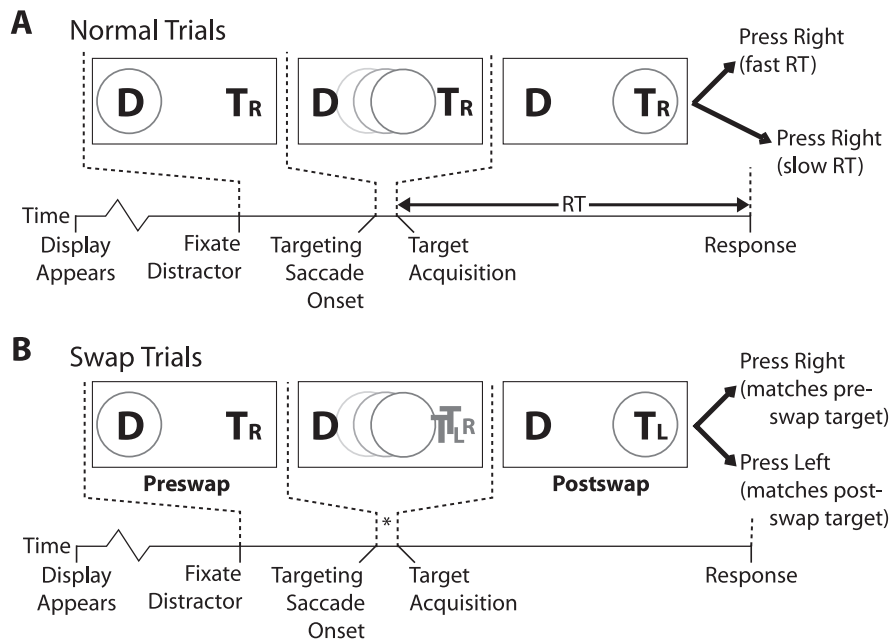
B Swap Trials

Figure 2. Diagram depicting the difference between normal trials and swap trials. Eye position is denoted by the gray circle. **A**, On normal trials, the monkey fixated a distractor image (D) with the target (T; subscript denotes button association) in the periphery. A saccade was initiated toward the target at a time referred to as the targeting saccade onset. After target acquisition, the monkey made a manual response by pressing one of two buttons (left or right). Comparisons were made across trials with different response times (RT) measured from target acquisition to button press. **B**, On swap trials, the identity of the target was changed during the targeting saccade (*). Specifically, the identity was changed such that the preswap and postswap targets had opposite hand associations (here, right target swapped to left target). Comparisons were made across trials in which the monkey's response indicated detection of the preswap target versus postswap target.

fication of objects more difficult and ensured that objects were not selected on the basis of a spatially isolated visual presence but rather on their salient features. Thus, the noise background was designed to mimic a natural scene without additional visual objects, which would limit our ability to guide the monkeys' eye movements without resorting to strict fixation requirements.

When the search array appeared, the monkey was allowed to freely view the display. Single trials of the search task always contained three images at any given moment: two distractors and one target. The monkeys' task was to locate a known target and press the corresponding button. Correct choices were immediately followed by delivery of a juice reward and the removal of the stimulus array from the screen. Feedback was also provided by auditory cues indicating correct and incorrect trials.

Behavioral experiments have uncovered differences in locating search targets for the initial saccade compared with subsequent saccades (Motter and Belky, 1998; Findlay et al., 2001). These differences suggest that laboratory experiments using sudden-onset stimuli may not precisely replicate natural vision. To further explore the differences between traditional experimental tasks and more natural behavior, we included an equal number of trials in which the target of a visual search would be located with the first saccade after display onset (first-saccade trials) or a subsequent saccade during exploration (second-saccade trials). Each image array type could appear in one of four orientations (up, down, left, or right). During first-saccade trials, one image appeared at fixation and the other two appeared at an eccentricity of 6°, with an angular separation of 135° (Fig. 1A). Again, these trials were designed such that the monkeys could locate the target image on their first saccade. During second-saccade trials, a similar image array appeared, but it was shifted by 4° up, down, left, or right (Fig. 1B). Thus, on these trials, one image was 4° from the initial fixation, whereas the other two images were 8.4° away. Second-saccade trials were designed such that the monkey's first saccade was directed to a distractor image and the target was fixated after a subsequent saccade.

In all arrays, the most eccentric images were displayed at a slightly lower contrast (75%) than the more central image (100%). The combi-

nation of proximity to fixation and higher contrast made the less-eccentric image the most salient object. On second-saccade trials, the first saccade consistently landed on the less-eccentric image. On most trials (and all experimental trials), this image was a distractor. The monkeys made subsequent saccades, most often to the target image, of their own will (for saccade pattern examples, see Fig. 1A,B). We concentrated our analysis on the time region preceding the targeting saccade onset, when the monkey is fixating a distractor image and will make a targeting saccade. All trials in which the monkey did not make this pattern of eye movements (e.g., made a saccade directly to the target on a second-saccade trial) were not included in any analyses.

We analyzed two main types of trials in the current experiment, and each type appeared in both first-saccade and second-saccade trials. During normal trials, the stimulus display was static throughout the entire trial (Fig. 2A). During swap trials, the identity of the target was changed when the monkey made a saccade toward the target (Fig. 2B). Specifically, the target identity was changed to a target with the opposite hand association (i.e., left-to-right target swap or right-to-left target swap). This is similar to the design of previous humans studies regarding the integration of visual information across saccades (Pollatsek et al., 1984, 1990). We used off-line analysis of eye position and buffer swap times to verify the timing of image swaps. Trials in which the actual display change did not occur in the middle of a saccade were removed from all analyses (mean 6.5% of trials for a given day). During swap trials, monkeys were rewarded for either response (left or right) because a target matching that hand was present at some point during the trial and we did not want to bias the monkeys to make either faster or more conservative responses. Normal trials ($n = 64$) and swap trials ($n = 64$) were interleaved along with various control trials ($n = 50$) in blocks of 178 trials.

Daily recordings sessions. Inferior temporal cortex was located based on the stereotaxic placement of the recording chamber and by counting white–gray matter transitions. Recordings were made in the ventral surface of the temporal lobe, lateral to the anterior medial temporal sulcus, in anterior inferior temporal cortex. We tested the response of every well isolated cell encountered to both target and distractors images in the monkey's stimulus set. The selectivity of the cell was analyzed by the experimenter using on-line raster plots. Any cell that showed sufficient selectivity such that there was one target image that consistently activated the cell to a greater extent than three other images (two distractors and a target with the opposite hand association) during a passive viewing task was selected for behavioral trials. For the passive viewing task, stimuli were presented centrally on a uniform gray background (50%) for 600 ms and the monkey was required to maintain fixation on the presented images throughout. Thus, a four-image set consisted of one left target, one right target, and two distractors (for an example neuron, see Fig. 4A). The limited number of images was chosen to ensure that sufficient behavioral trials could be obtained in a single session. Off-line analysis confirmed the selectivity of our neuron pool. Average response functions to effective and ineffective stimuli are shown in Figure 4B. In addition, we calculated the depth of selectivity (DOS) (Moody et al., 1998; Rainer and Miller, 2000) for each of our neurons based on the firing rate of the cell from 75 to 225 ms after stimulus onset in our passive viewing task: $DOS = [n - (\Sigma R_i/R_{max})]/(n - 1)$, where n is the number of images presented, R_i is the firing rate of the neuron to the presentation of the ith stimulus, and R_{max} is the largest firing rate across all presented images. This value takes into account the response of the cell to all four stimuli

and ranges from 0 (all images activate cell equally) to 1 (only one image activates the cell). Our neurons had an average depth of selectivity of 0.83 with a range of 0.43–1.0.

Two-thirds of the trials were composed of the four stimuli chosen based on the selectivity of the recorded unit. For these trials, we could attribute any significant modulation of neural activity to the presence of the effective target. The remaining one-third of the trials were composed of a random selection of target and distractors from the monkey's full image set and were included to ensure that the monkey maintained good performance during each block of trials and to discourage the adaptation of search strategies specific to the four stimuli chosen for that session.

We ran this experimental protocol during the recording of 44 neurons (25 in monkey S, 19 in monkey M). Two datasets were removed from all analyses because of poor behavioral performance (>10% errors) on interleaved control trials. For the analysis of swap trial data, we limited our analysis to those sets in which the monkey's behavior yielded at least five trials per condition (response matching preswap and postswap target) for first-saccade and second-saccade trials independently. This requirement, which was enforced because of the susceptibility of our analysis (see below) to error from a small number of trials (Britten et al., 1996), restricted our swap trial analysis to 32 datasets for first-saccade trials and 29 datasets for second-saccade trials. For the analysis of normal trial data, we limited our analysis to those sets in which the monkey's behavior yielded at least five trials for first-saccade and second-saccade trials independently. This requirement limited our normal trial analysis to 41 neurons for first-saccade trials and 41 neurons for second-saccade trials.

Statistical analysis. All statistical analyses were done using custom software unless otherwise noted.

To compare neuronal activity and behavioral performance, we analyzed trials in which the preswap target was the effective target for that cell. We extracted the firing rate of each neuron in a 200 ms time window centered on the initiation of the targeting saccade for every trial. This window was chosen because it represents the last time the preswap target was on the screen and thus the last time the preswap target could directly influence the activity of the cell or the monkey's behavioral report. This is not meant to imply that the recorded neural modulations are directly related to the execution of the saccade itself, but rather, this time window is appropriate for comparing visual-evoked responses to the preswap target across our population of neurons. In the following analyses, we compared the distribution of firing rates for each neuron as a function of the monkey's overt response (swap trials, Fig. 2*B*) or response time (normal trials, Fig. 2*A*).

To quantify the relationship between neuronal firing rate and overt response during swap trials, we computed empirical "receiver (or relative) operating characteristic" (ROC) curves and estimated the area under these curves (Green and Swets, 1966; Swets, 1996). The technique used here is similar to that used in previous combined behavioral and physiological studies of visual perception (e.g., Britten et al., 1996). The area under the ROC curve, also termed "choice probability" (Britten et al., 1992), gives a reliable measure for the separation of two distributions. In the current experiment, the computed area is an estimate of how well the firing rate of individual IT neurons predicted the overt behavioral response of the monkey (left or right button press). Importantly, this measure is relatively unaffected by decision criterion or bias, which is particularly difficult to control with animal subjects. Furthermore, the area under the ROC curve is a distribution-free estimate of sensitivity and does not assume that the data are normally distributed.

To quantify the relationship between firing rates and response time, we analyzed normal trials in which the target was the effective stimulus for that cell. Note that response times were defined relative to target acquisition (i.e., manual response time − time of first fixation on the target). We calculated the biweight midcorrelation coefficient (r_{bw}) (Wilcox, 1997), a robust measure of correlation in the presence of outliers, between neuronal firing rate and overt response time. We chose this robust method because our monkeys were free to respond at any time and sometimes responded well after fixating the target image (for example data, see Fig. 8*C,D*). A negative correlation indicates an inverse relationship between neural activity magnitude and the time of recognition.

We also calculated population ROC areas and population correlation

coefficient values by pooling the data from all of our cells. To do so, we normalized the firing rates by the mean and SD for each neuron individually and reran our analyses on the combined population data. This measure yielded similar results to taking a mean across the individual ROC area or correlation coefficient of each neuron, but it weighs the contribution of each neuron to the population by the number of trials supplied by that neuron.

To evaluate the influence of relative target location (i.e., upper left vs lower right) on our ROC and correlation results, we created a large number of "mini-experiments" from our data. Each mini-experiment contained a set of trials from a single neuron with matched stimulus configuration (i.e., the same relative target position within the stimulus array). For each mini-experiment, we quantified the difference in neural modulation across behavioral choice (swap trials) or the correlation between neural modulation and response time (normal trials) using the methods described above. Using a permutation test (see below), we determined whether the resulting distribution of ROC areas or correlation coefficients was statistically different from chance (0.50 for ROC analysis; 0.0 for correlation analysis).

To evaluate the influence of variability in presaccadic retinal stimulation on our ROC area and correlation results, we ran a multiple regression analysis (least trimmed squares) to asses the contribution of the following three variables in predicting the monkey's overt response (swap trials) or response time (normal trials): normalized firing rate, presaccadic target eccentricity, and presaccadic fixation duration. These analyses were run using the MASS package of R (version 2.4.0), and statistical significance of the standardized regression coefficients (β) was computed by a permutation test (see below). The influence of presaccadic target eccentricity and presaccadic fixation duration was also evaluated independently using a reiterative process. For each parameter, we first removed (without regard to behavioral choice or response time) the top and bottom 1% of the trials and retested the remaining trials for a significant difference in that parameter across behavioral choice (swap trials) or no significant correlation between that parameter and response time (normal trials). If a significant effect remained, we then removed the top and bottom 2% of the trials and retested for a significant effect of the parameter. This process continued until we obtained a subset of trials for which there was no significant effect of the tested parameter. This subset of trials was then tested for significant differences in neural modulation across behavioral choice (swap trials) or for a significant correlation between neural modulation and response time (normal trials) using the methods described above.

All of the analyses outlined above yielded qualitatively similar results when normalizing response times, fixation durations, and target eccentricities for each experimental session by their mean and SD.

Statistical significance for the area under the ROC curve, and the response time/firing rate correlation was computed using a permutation test (Efron and Tibshirani, 1993) with at least 1000 permutations. For these tests, we calculated the ROC area or the correlation coefficient for our data after randomly shuffling the assignments of firing rates and overt response (ROC analysis) or response time (correlation analysis). From many such permutations, we obtained a distribution of measures that would be expected to occur by chance if there was no relationship between the measured variables. Actual values for ROC area or correlation coefficients that lay outside the central 95% of the permuted distributions were considered significant (i.e., two-tailed test at an α level of 0.05).

Results

To explore the correlation between activity in IT neurons and overt recognition, we used a modified visual search paradigm (Figs. 1, 2). The monkeys' task was to locate a target object previously associated with a hand response (left or right button press) and to press the appropriate button to receive a juice reward. During visual search, monkeys naturally made saccadic eye movements toward the target image. Although the monkey was allowed to freely view the search array, the task was designed to obtain trials with a stereotypical pattern of natural eye move-
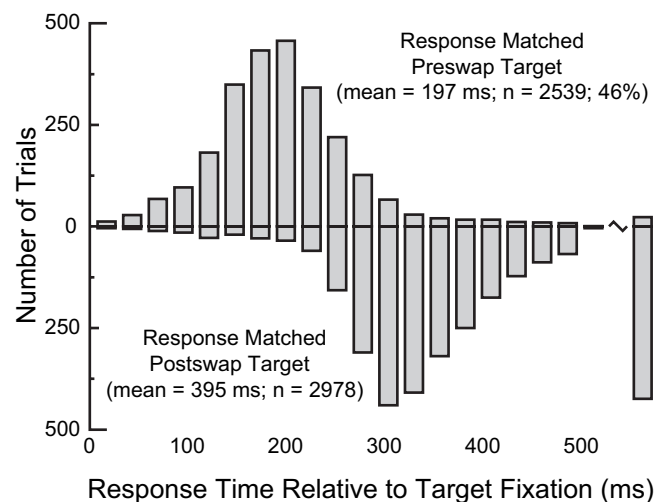
Figure 3. Response time distributions for swap trials as a function of the monkey's response collapsed across all recording sessions. For each recording session, response times were significantly faster when the monkey's response matched the preswap target ($p < 0.001$). Trials with a response time >500 ms relative to target fixation are shown in the farthest right bin ($<10\%$ of data).

ments, first to the closest distractor (which appeared at the initial fixation position on first-saccade trials) and then to the target (Fig. 1A,B). On some trials, referred to as swap trials, the identity of the target was changed mid-trial. This change occurred during the targeting saccade (Fig. 2B), at a time when the monkey was insensitive to the transient artifacts of the change itself (Matin, 1974). Furthermore, the identity of the target changed such that the response mapping of the preswap and postswap targets differed (i.e., a right preswap target was changed to a left postswap target or vice versa). Thus, on these trials, we can infer from the monkey's overt response whether or not he fully recognized the initial, preswap target.

Behavior
When attributing perceptual awareness to neural activity, it is crucial to establish that the subject's behavioral report can be trusted (Leopold et al., 2003). An analysis of the response time distributions for choices matching the preswap and postswap images confirmed that the monkeys seemed to reliably report their perception (Fig. 3). Response times were faster when responses matched the preswap target association (mean of 197 ms; must be programmed before the targeting saccade) than when responses matched the postswap target association (mean of 395 ms; must be programmed after the targeting saccade). This difference in response time distributions was significant for all 42 recording sessions (mean ROC area of 0.92; $p < 0.001$ for all sessions, permutation test).

Because the monkey was rewarded for pressing either button on swap trials, there is no measure of accuracy. However, across all interleaved nonswap trials (including control trials with random target images), error rate did not exceed 6%. Furthermore, to ensure that the swap trials were not "special" (in that the monkeys could learn to press either button if they detected a swap), both monkeys were tested in separate behavioral experiments before and after recording sessions that included consistent-swap (i.e., left-to-left target swap or right-to-right target swap) and distractor-swap (i.e., distractor-to-target swap) trials. On these trials, the monkeys were only rewarded for pressing the hand
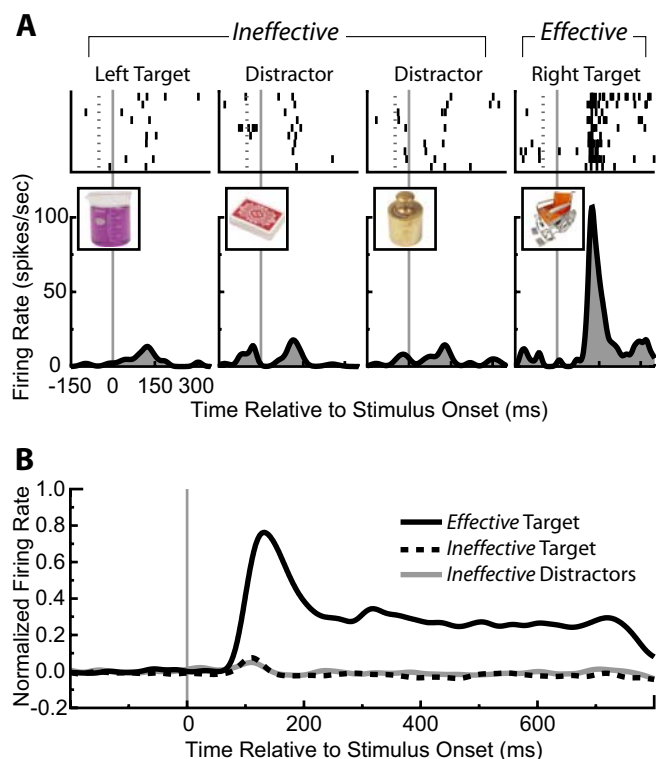

Figure 4. Neural response during a passive viewing task to assess stimulus selectivity for presented objects. **A**, Rasters and spike density functions for an example cell (same cell as in Figs. 6A–D and 8A–D). Activity is aligned to stimulus onset. On each day, one target (wheelchair here) was selected as an effective stimulus. A target with the opposite hand association (beaker here) and two distractors were chosen as ineffective stimuli. **B**, Population spike density functions showing average normalized response to effective target and ineffective target and distractors. The population of neurons was highly selective for the effective target.

associated with the presented target(s). Error rates did not exceed 2% for consistent-swap trials or 10% for distractor-swap trials.

Swap trials
To test the hypothesis that activity in IT neurons is directly involved in recognition, we analyzed swap trials (Fig. 2B) in which the preswap target was the effective stimulus for the recorded IT neuron (Fig. 4). We hypothesized that the response of an IT neuron to the preswap target should be of greater magnitude when the monkeys reported that they saw the preswap target. If the monkeys' report matched the postswap target, they presumably did not acquire sufficient peripheral information to recognize the preswap target. Thus, the activity of these same neurons should be lower.
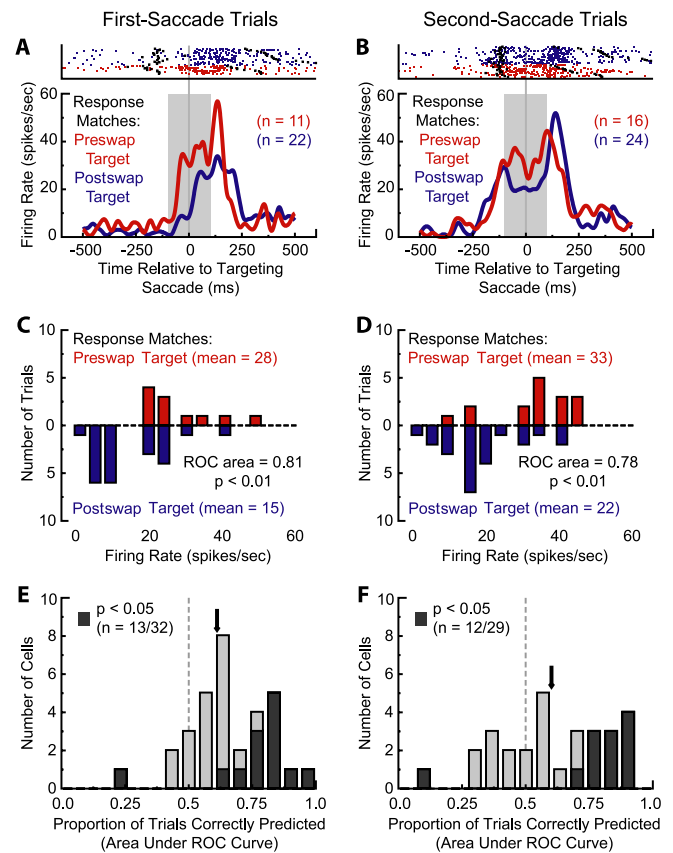
Figure 5 shows a subset of swap trial data from one experimental session. The monkey's eye trace (left) and fixation durations (top right), as well as the average firing rate of the recorded neuron (bottom right), are shown as a function of the monkey's overt response. The activity of this typical neuron collapsed across all target locations is shown in Figure 6. For both first-saccade (Fig. 6A) and second-saccade (Fig. 6B) trials, before the saccade, there is a ramp up of neural activity, which sharply decreases after the saccade, attributable to the fact that the effective stimulus (preswap target) has been swapped with an ineffective stimulus (postswap target). Furthermore, the trials are sorted by the monkey's behavioral response. The ramp up of activity before target fixation was stronger when the monkey reported seeing the preswap target.
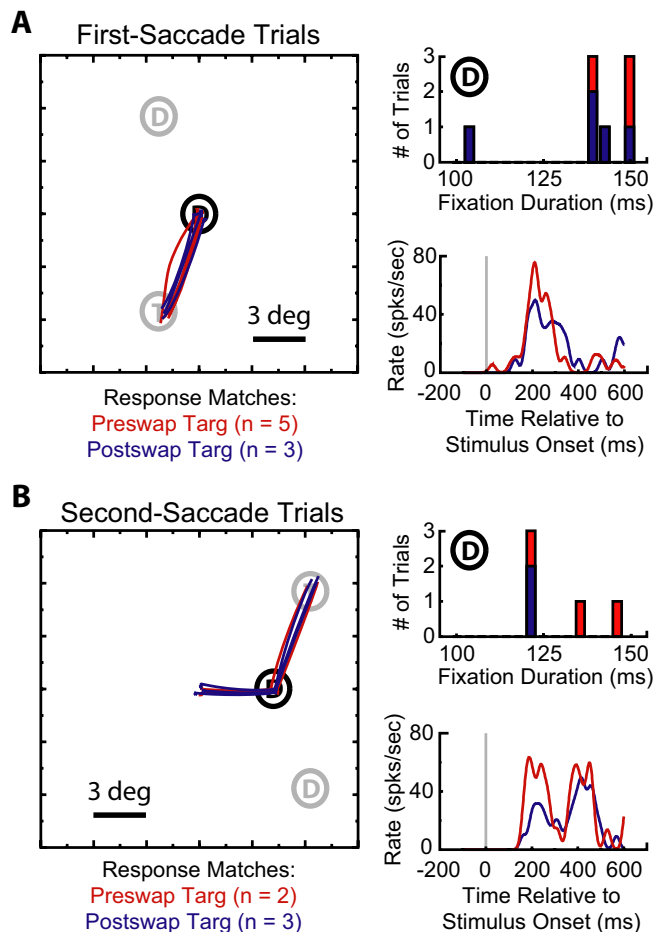
Figure 5. Example data from one experimental session. Note that, for simplification, only one of eight possible target positions is shown for both first-saccade (**A**) and second-saccade (**B**) trials. Left panels denote the monkey's eye trace during swap trials from the onset of the search array until the button press. Circled letters denote the position of distractors (D) and the target (T). Spike density functions depicting the response of a neuron selective for the preswap target (same cell in Fig. 4A) are shown in the bottom right panels. Top right panels show the fixation duration of the fixation preceding the targeting saccade (black circle; see Fig. 2B, left). Data in all panels are sorted by the monkey's response during swap trials.

Figure 6. Swap trial analysis for first-saccade trials (**A, C, E**) and second-saccade trials (**B, D, F**) for an example neuron (**A–D**) and population (**E, F**). **A, B**, Rasters and spike density functions during swap trials in which the effective stimulus was the preswap target for an example neuron (same cell in Fig. 4A). Blue and red ticks denote spikes, and black ticks denote the beginning of the presaccadic fixation (left of 0) and time of manual response (right of 0). Activity is aligned to the onset of the targeting saccade. Note the presaccadic increase in activity, which was larger for trials in which the monkey's response matched the preswap target. There was a sharp reduction in neural activity after the saccade because the effective stimulus was changed to an ineffective stimulus (postswap target). **C, D**, Distribution of firing rates in a 200 ms time window centered on the targeting saccade onset (shaded area in **A, B**). The firing rate was, in general, greater when the monkey's manual response matched the button association of the preswap target. For these distributions, the area under the ROC curve was 0.81 for first-saccade trials and 0.78 for second-saccade trials. Both of these values were significantly more than 0.50 (*p* < 0.01). **E, F**, Histogram of ROC area values obtained from the population of neurons. Dark gray bars denote neurons with ROC area values significantly different from 0.05. Black arrows denote the population ROC area measured after collapsing across trials from all neurons (0.62 for first-saccade trials; 0.61 for second-saccade trials). Both population ROC areas were significantly more than 0.50 (*p* < 0.0005 in both cases).

Because of the timing of the change in target identity, we predicted that physiological correlates of object recognition would occur close to the initiation of the targeting saccade. In other words, presaccadic information might reflect a signature of recognition of the preswap target. Presaccadic modulation of neural signals, presumably related to the obligatory shift of attention before eye movements (Kowler et al., 1995), has been observed in many regions of the macaque brain, including lateral intraparietal area (Colby et al., 1996), frontal eye field (Goldberg and Bushnell, 1981), visual cortical area V4 (Fischer and Boch, 1981; Moore et al., 1998), and IT (Sheinberg and Logothetis, 2001). Furthermore, presaccadic activity in visual areas such as V4 is correlated with saccade accuracy (Moore, 1999). Therefore, to further analyze the difference in neural activity across the monkey's behavioral response, we looked at the firing rate of the cell in a 200 ms window around the initiation of the targeting saccade (Fig. 6A,B, shaded area). This window represents the last time that the preswap target was on the screen and thus the last time the preswap target could directly affect the response of the cell.

Figure 6, C and D, shows the distributions of firing rates in the 200 ms window around the initiation of the targeting saccade as a

function of the monkey's overt response for the example cell. For both first- and second-saccade trials, the firing rate of the neuron was generally higher when the monkey's response matched the preswap target (an effective stimulus for the cell). To quantify this difference, we used ROC analysis (see Materials and Methods). Here, the area under the ROC curve, also termed choice probability (Britten et al., 1992), represents the proportion of trials for which an ideal observer could correctly predict the monkey's behavioral choice based solely on the firing rate of a single IT neuron before target fixation. For our two-alternative choice task, chance would be 0.50. For the example cell in Figure 6, C and D, this analysis revealed an ROC area of 0.81 for first-saccade trials and 0.78 for second-saccade trials. These values were significantly more than 0.50 as measured by a permutation test (see Materials and Methods; *p* < 0.01).

We performed this same analysis for all cells with at least five trials per condition (response-matched preswap or postswap target) for first-saccade and second-saccade trials independently. The results are shown in Figure 6, *E* and *F*. For first-saccade trials, 13 of 32 (41%) neurons produced an ROC area significantly different from chance, and the vast majority of those significant values were more than 0.50 (92%). The results for second-saccade trials were similar; 12 of 29 (41%) neurons produced an ROC area significantly different from chance, with all but one of those significant values being more than 0.50 (92%). These results indicate that stimulus-selective IT neurons can predict the choice behavior of monkeys in our task.

Because none of our cells were perfect at predicting the monkey's behavioral choice and many individual neurons did not yield significant ROC areas, we sought to determine how well a population of IT neurons would correlate with the monkey's overt behavior. To quantify the performance of our population of neurons, we pooled all of our data after normalizing the firing rate of each neuron by its mean and SD and computed a population ROC area (done independently for first-saccade and second-saccade trials). The population ROC area was 0.62 for first-saccade trials and 0.61 for second-saccade trials, and both of these values were significantly more than 0.50 ($p < 0.0005$). Thus, the small population of stimulus-selective IT neurons we recorded could predict the monkey's trial-to-trial choice behavior better than chance.

Swap trials: effects of stimulus configuration

To reduce overtraining on one particular stimulus configuration, we included target placements that required one of eight different targeting saccade vectors (4 orientations × 2 target positions) (for two example configurations, see Fig. 1). To verify that the inclusion of a variety of targeting saccade directions did not influence our results, we performed a separate analysis on a subset of our data. For each of the 42 neurons, independently, we selected sets of trials with the same targeting saccade direction. We were able to create 145 sets of first-saccade trials (using 66% of total trials) and 105 sets of second-saccade trials (using 54% of total trials). Each set of trials composed a mini-experiment and had the following properties: all trials came from the recording sessions of a single neuron, all trials had the same relative target position within the stimulus array (same targeting saccade direction), and each set contained at least one trial with each possible behavioral choice (response matches preswap vs postswap target).

We quantified the difference in firing rate across behavioral choice for each set of trials by calculating the ROC area. Across all of these mini-experiments, the distribution of ROC areas was significantly greater than chance (0.50) for both first-saccade (mean ROC area of 0.58; $p = 0.0008$) and second-saccade (mean ROC area of 0.59; $p = 0.004$) trials. This analysis confirms that variability in relative target position did not induce artifactual differences in firing rate across behavioral choice. Thus, even when we controlled for the exact stimulus array configuration, our population of neurons was still able to predict the monkey's choice behavior better than chance.

Swap trials: effects of presaccadic target eccentricity and fixation duration

In addition to variability in relative target location, the free-viewing nature of our task left open the possibility that variability in monkeys' pattern of eye movements on trials with opposite responses might induce a secondary correlation between neural
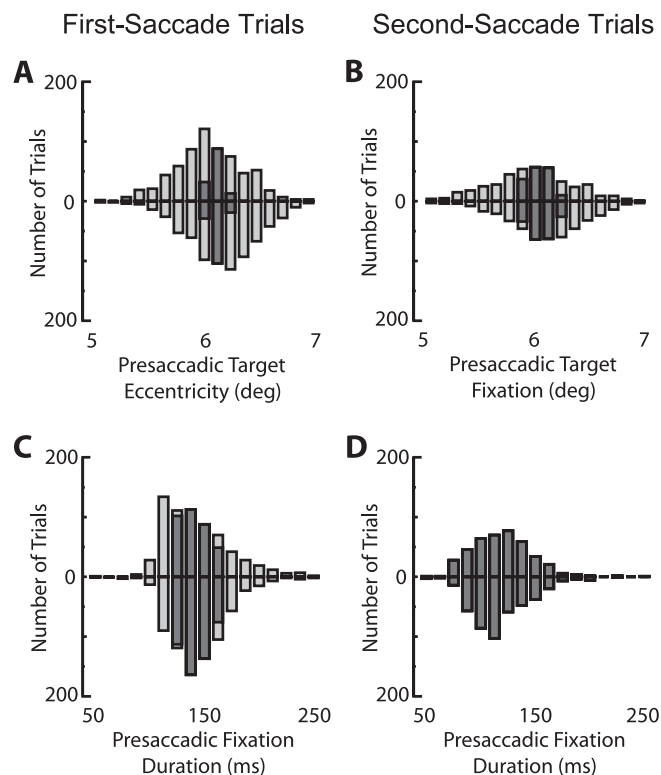
Figure 7. Distributions of presaccadic target eccentricity (*A*, *B*) and presaccadic fixation duration (*C*, *D*) across behavioral choice during swap trials. Upward and downward bars represent trials in which the monkey's response matched the preswap and postswap targets, respectively. Dark bars represent the subset of trials for which there was no significant difference across behavioral choice. For details, see Results.

activity and choice behavior. The stimulus arrays were configured such that the distance to the target was ~6° during the time of our neural measurement. Although the peripheral location of the target before the saccade was well within the range of receptive field sizes for IT neurons (Gross et al., 1972; Op De Beeck and Vogels, 2000), other reports on the sensitivity of IT neurons to small changes in the spatial location of an object (DiCarlo and Maunsell, 2003) suggest that small differences in presaccadic fixation position across individual trials could potentially bias our results. In addition to differences in presaccadic target eccentricity, variability in the duration of the fixation preceding the targeting saccade (Fig. 5) could also bias our results; longer fixation durations might lead to stronger neural responses because the neuron has more time to process the retinal stimulus, and this factor could be linked to the monkey's overt response.

We directly measured the target eccentricity and fixation duration for the fixation preceding the targeting saccade (Fig. 2*B*, left). Figure 7 shows the distributions of presaccadic target eccentricity (*A*, *B*) and presaccadic fixation duration (*C*, *D*) across behavioral choice for first- and second-saccade trials. Across the population of neurons, we found a significant difference in the presaccadic target eccentricity across choice for both first-saccade (ROC area of 0.39; $p < 0.001$) and second-saccade (ROC area of 0.39; $p < 0.001$) trials. The presaccadic eye position was significantly closer to the target on trials in which the monkey's response matched the preswap target. However, this difference was very small (~0.1°), and there is little evidence that IT neurons are sensitive to translational changes on that scale, especially for targets at an eccentricity of 6°. In the study mentioned above, Di-

Carlo and Maunsell (2003) measured the response to 1.5° changes in the position of 0.6° images.

There was a significant difference in presaccadic fixation duration across choice for first-saccade trials (ROC area of 0.46; $p = 0.008$) but not for second-saccade trials (ROC area of 0.52; $p = 0.282$). In addition, for first-saccade trials, fixation durations were shorter when the monkey's response matched the preswap target. This suggests that the duration of the fixation preceding the targeting saccade may be dependant on the monkey's choice in that the monkey released this fixation sooner on trials that he located the target. However, as with differences in the presaccadic target eccentricity, the difference in fixation duration was very small (<2 ms), and we hesitate to over-interpret the meaning of such a small effect.

To verify that small differences in presaccadic target eccentricity and fixation duration did not lead to artifactual differences in firing rate across choice, we ran a multiple regression analysis and compared the standardized regression coefficients (β) of three predictors of the monkey's behavioral choice: normalized firing rate, target eccentricity, and fixation duration. For both first-saccade ($\beta = 0.203$; $p < 0.001$) and second-saccade ($\beta = 0.186$; $p < 0.001$) trials, the neuronal firing rate was the single best predictor of the monkey's choice. Presaccadic target eccentricity was significantly correlated with the monkey's choice for second-saccade trials ($\beta = -0.120$; $p < 0.001$) but only marginally so for first-saccade trials ($\beta = -0.048$; $p = 0.078$). Presaccadic fixation duration was not significantly correlated with the monkey's choice for either first-saccade ($\beta = -0.028$; $p = 0.314$) or second-saccade ($\beta = 0.035$; $p = 0.312$) trials. The strong regression coefficients for normalized firing rate along with the inconsistent contribution of presaccadic target eccentricity and fixation duration suggest that the firing rate of our population of IT neurons is the best predictor of the monkeys' choice behavior.

We also ran a separate control analysis for presaccadic target eccentricity and fixation duration independently. For each factor, we reanalyzed the firing rate differences across choice for a subset of data after removing trials (without regard to behavioral choice; see Materials and Methods) with the most extreme values of each factor (Fig. 7, dark bars). For target eccentricity, this subset contained 20% of the total trials for first-saccade trials and 40% of the total trials for second-saccade trials. For these subsets of data, there was no significant difference in presaccadic target eccentricity across choice (first-saccade trials, ROC area of 0.44, $p = 0.096$; second-saccade trials, ROC area of 0.44, $p = 0.064$), but ROC areas for the firing rate comparison remained significant and did not change substantially from the analysis using the full dataset (first-saccade trials, ROC area of 0.58, $p = 0.028$; second-saccade trials, ROC area of 0.59, $p = 0.002$).

For presaccadic fixation duration, we only extracted a subset of trials for first-saccade trials because there was no significant difference in fixation durations across behavioral choice for second-saccade trials. For the subset of first-saccade trials (containing 60% of the total trials), there was no significant difference in fixation durations across choice (ROC area of 0.46; $p = 0.068$). However, for this same set of trials, we again found a highly significant difference across behavioral choice (ROC area of 0.58; $p < 0.001$).

Furthermore, variability in target eccentricity could not account for differences in neural activity across behavioral choice in the majority of neurons showing individually significant results (8 of 13 for first-saccade trials and 8 of 12 for second-saccade trials). Likewise, variability in fixation duration could also not account for differences in neural activity across behavioral choice in the majority of neurons showing individually significant results (10 of 13 for first-saccade trials and 11 of 12 for second-saccade trials). This includes neurons that did not show a significant difference in the given parameter across choice and those neurons for which any significant difference could be removed without altering the significance of the firing rate result.

These control analyses show that the firing rate of our population of IT neurons continues to predict the monkeys' trial-to-trial choice behavior even when presaccadic target eccentricity and fixation duration do not significantly vary across the monkeys' response.

Normal trials

Our data indicate that the firing rate of individual IT neurons is correlated with the monkey's overt response. To further explore the correlation of IT activity and recognition and to verify that our results were not an artifact of the unusual swap condition, we tested whether the firing rate of these same neurons was correlated with the time of recognition in static displays. To address this question, we looked at normal trials (Fig. 2A) in which the effective stimulus for the recorded neuron (Fig. 4) was the target. We hypothesized that the response of an IT neuron to the target would be of greater magnitude on trials when the monkey responded faster.

The activity of a typical neuron aligned to the initiation of the targeting saccade during normal trials is shown in Figure 8. Similar to swap trials (Fig. 6), before the saccade there is a ramp up of neural activity during both first-saccade (Fig. 8A) and second-saccade (Fig. 8B) trials. In contrast to swap trials, there is also a transient burst of activity after the saccade. This is because, on normal trials, the identity of the target does not change, and thus the monkey's targeting saccade lands on the effective stimulus.

The raster plots on top of Figure 8, A and B, are sorted by the monkey's response time relative to target fixation such that the fastest response trials are shown on top and the slowest response trials on the bottom. From these raster plots, it appears that this neuron tended to respond more robustly during trials in which the monkey made a fast response. To quantify the relationship between neural activity and response time, we looked at the firing rate of the cell in the same 200 ms time window around the onset of the targeting saccade that we used for the swap trial analysis (Fig. 8A,B, shaded region). We calculated the biweight midcorrelation coefficient (see Materials and Methods), a robust method in the presence of outliers, between the firing rate during this time window and the monkey's response time on a trial-by-trial basis for first-saccade and second-saccade trials independently (Fig. 8C,D). This analysis revealed an inverse relationship (negative correlation) in both cases (-0.41 for first-saccade trials; -0.55 for second-saccade trials). This value was significantly less than zero in both cases as measured by a permutation test ($p < 0.05$). Thus, the stronger the response of this cell, the faster the monkey's overt response.

We performed this same analysis in all datasets with at least five trials. The results are shown in Figure 8, E and F. For first-saccade trials, 15 of 41 (37%) neurons produced a significant correlation coefficient. The results for second-saccade trials were similar; 9 of 41 (22%) neurons produced a significantly negative correlation coefficient. For both trial types, no cells produced a correlation coefficient that was significantly greater than zero. These data suggest that the strength of activity in selective IT cells is related to the time of recognition of visual objects.

To quantify the correlation between neural activity and recognition time for our population of neurons, we pooled all of our
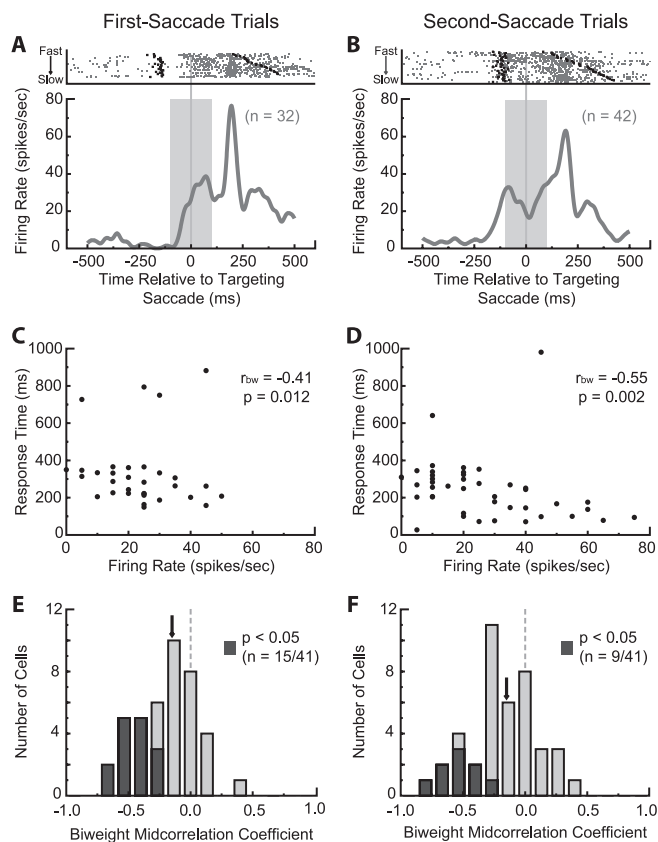
Figure 8. Normal trial analysis for first-saccade trials (**A**, **C**, **E**) and second-saccade trials (**B**, **D**, **F**) for an example neuron (**A–D**) and population (**E**, **F**). **A**, **B**, Rasters and spike density functions during normal trials in which the effective stimulus was the target for an example neuron (same cell in Fig. 4A). Gray ticks denote spikes, and black ticks denote the beginning of the presaccadic fixation (left of 0) and response time (right of 0). Activity is aligned to the onset of the targeting saccade. There is a presaccadic increase in activity, followed by a transient peak after the saccade because the monkey fixated an effective stimulus (target) at this time. Rasters are sorted by the monkey's response time with faster trials depicted in the top rows. **C**, **D**, Response times (relative to target acquisition) plotted as a function of the firing rate of the neuron in a 200 ms time window around the time of the targeting saccade (shaded area in **A**, **B**). The firing rate was, in general, greater when the monkey made faster button presses. Biweight midcorrelation coefficients (r_{bw}; see Materials and Methods) were -0.41 for first-saccade trials and -0.55 for second-saccade trials. Both values were significantly less than 0 ($p < 0.05$). **E**, **F**, Histogram of biweight midcorrelation coefficients obtained from the population of neurons. Dark gray bars denote neurons with correlation coefficients significantly different from 0. Black arrows denote the population correlation coefficient measured after collapsing across trials from all neurons (-0.15 for first-saccade trials; -0.15 for second-saccade trials). Both population correlation coefficients were significantly <0 ($p < 0.0002$ in both cases).

data after normalizing the firing rate of each neuron by its mean and SD and computed a population correlation coefficient (done independently for first-saccade and second-saccade trials). The population correlation coefficient was -0.15 for first-saccade trials and -0.15 for second-saccade trials, and both of these values were significantly less than zero ($p < 0.0002$). Thus, as a population, the stimulus-selective IT neurons we recorded were significantly correlated with monkey's trial-to-trial response times.

Normal trials: effects of stimulus configuration
As with our swap trial analysis, it is important to verify that factors leading to variability in retinal stimulation across trials cannot account for the measured correlation between neural activity and response time. We first tested the effects of stimulus configuration to ensure that differences in the relative location of the target before the targeting saccade did not influence the correla-

tion results above. For each of the 42 neurons, independently, we selected sets of trials with the same targeting saccade direction. We were able to create 193 sets of first-saccade trials (using 85% of total trials) and 126 sets of second-saccade trials (using 71% of total trials). Each set of trials composed a mini-experiment and had the following properties: all trials came from the recording session of a single neuron, all trials had the same relative target position within the stimulus array (same targeting saccade direction), and each set contained at least five trials.

We quantified the correlation between neural firing rate and behavioral response time for each set of trials. Across all of these mini-experiments, the distribution of correlation coefficients was significantly less than zero for both first-saccade (mean r_{bw} of -0.04; $p < 0.0002$) and second-saccade trials (mean r_{bw} of -0.06; $p < 0.0002$). This analysis confirms that variability in relative target position did not induce artifactual correlations between firing rates and response time. Thus, even when we controlled for the exact stimulus array configuration, our population of neurons was still able to predict the monkey's response time better than chance.

Normal trials: effects of presaccadic target eccentricity and fixation duration
We also directly measured the target eccentricity and fixation duration for the fixation preceding the targeting saccade on normal trials (Fig. 2A, left). Across the population of neurons, we found a significant correlation between the presaccadic target eccentricity and the monkey's response time in both first-saccade (0.18; $p < 0.001$) and second-saccade (0.18; $p < 0.001$) trials. The closer the eyes were to the target, the faster the monkey's response time. We also found a significant correlation between the presaccadic fixation duration and response time for both first-saccade (0.14; $p < 0.001$) and second-saccade (0.08; $p = 0.004$) trials. Fixation durations were shorter when the monkey made faster responses. Consistent with our swap trial data, this relationship is in the opposite direction that would be hypothesized if variability in fixation duration were causing the observed correlation between firing rate and response time.

To ensure that variability in presaccadic target eccentricity and fixation duration did not underlie the observed correlation between neural response and response time, we ran a multiple regression analysis and compared the standardized regression coefficients (β) of three predictors of the monkey's response time: normalized firing rate, target eccentricity, and fixation duration. For both first-saccade ($\beta = -0.061$; $p < 0.001$) and second-saccade ($\beta = -0.032$; $p < 0.001$) trials, the neuronal firing rate was the best predictor of the monkey's response time. Presaccadic target eccentricity was significantly correlated with the monkey's response time for second-saccade trials ($\beta = 0.028$; $p = 0.012$) but not for first-saccade trials ($\beta = 0.012$; $p = 0.194$). For second-saccade trials, larger target eccentricities led to longer response times. Presaccadic fixation duration was significantly correlated with the monkey's response time for first-saccade trials ($\beta = 0.056$; $p < 0.001$) but not for second-saccade trials ($\beta = 0.012$; $p = 0.15$). For first-saccade trials, longer fixation durations led to longer response times. Once again, this implies that the fixation preceding the targeting saccade was cut short on trials in which the monkey recognized the target sooner. Consistent with the results from swap trials, the highly significant regression coefficients for normalized firing rate along with the inconsistent contribution of presaccadic target eccentricity and fixation duration suggest that the firing rate of our population of IT neurons is the best predictor of the time of target recognition.

Additional analyses showed that variability in presaccadic target eccentricity and fixation duration cannot fully account for the correlation between neural activity and response time. For each factor, we reanalyzed the correlation between neural activity and response time for a subset of data after removing trials (without regard to response time) with the most extreme values of each factor until the remaining set of trials did not show a significant correlation with response time. For target eccentricity, this subset contained 54% of the total trials for first-saccade trials and 50% of the total trials for second-saccade trials. For these subsets of data, there was no significant correlation between presaccadic target eccentricity and response time (first-saccade trials, 0.06, $p = 0.068$; second-saccade trials, 0.07, $p = 0.094$), but there remained a strong inverse relationship between neural activity and response time that did not change substantially from the analysis using the full dataset (first-saccade trials, -0.13, $p < 0.001$; second-saccade trials, -0.16, $p < 0.001$).

For fixation duration, the subset of trials contained 36% of the total trials for first-saccade trials and 66% of the total trials for second-saccade trials. For these subsets of data, there was no significant correlation between presaccadic fixation duration and response time (first-saccade trials, 0.05, $p = 0.142$; second-saccade trials, 0.06, $p = 0.056$), but correlation coefficients for the firing rate comparison remained highly significant and did not change substantially from the analysis using the full dataset (first-saccade trials, -0.11, $p < 0.001$; second-saccade trials, -0.14, $p < 0.001$).

In addition, variability in target eccentricity could not account for the relationship between neural activity and response time in the majority of neurons showing individually significant results (14 of 15 for first-saccade trials and 7 of 9 for second-saccade trials). Likewise, variability in fixation duration could not account for the relationship between neural activity and response time in the majority of neurons showing individually significant results (10 of 15 for first-saccade trials and 9 of 9 for second-saccade trials). This includes neurons that did not show a significant correlation between response time and the given parameter and those neurons for which any significant correlation could be removed without altering the significance of the firing rate result.

These control analyses show that the firing rate of our population of IT neurons continues to predict the time of the monkeys' manual response even when presaccadic target eccentricity and fixation duration do not significantly correlate with the monkeys' response time.

Comparison across swap and normal trials

A total of 24 of 42 neurons showed a significant result for at least one of the ROC or correlation analyses (first- or second-saccade trial). Eight neurons with a significant result for at least one of the ROC analyses also had a significant negative firing rate-response time correlation in at least one of the two correlation analyses. Seven neurons showed a significant result for one of the ROC analyses but not the correlation analyses. Nine neurons showed a significant result for one of the correlation analyses but not for the ROC analyses. Eighteen neurons did not show a significant result in any of the analyses.

Across the neuron population, there was a significant correlation between ROC area (swap trials) and correlation coefficient (normal trials) for both first-saccade (Pearson's $r = -0.55$; $p = 0.0001$) and second-saccade (Pearson's $r = -0.51$; $p = 0.0006$) trials.

Discussion

Examples of neural correlates of decision making are abundant for cell populations in the dorsal stream of the visual system (Britten et al., 1992, 1996; Celebrini and Newsome, 1994; Dodd et al., 2001; Uka and DeAngelis, 2004; Liu and Newsome, 2005), but few studies have explored the role of the ventral visual stream in choice behavior (Uka et al., 2005), particularly under natural viewing conditions. We designed a task in which monkeys could be on the verge of detecting a target image (preswap target) to more clearly demonstrate the connection between neural activity in the ventral stream and recognition behavior. When the monkeys' report matched the preswap target, activity in IT neurons selective for that target was more robust than during trials when the monkeys' report did not match the preswap target. Importantly, this difference in neural modulation could not be explained by variability in targeting saccade direction, presaccadic target eccentricity, or presaccadic fixation duration across trials. These results indicate that object recognition, and in this case parafoveal recognition, is tightly correlated with the activity of single units in IT.

Although a significant fraction (41%) of the neurons we recorded predicted the monkey's choice behavior significantly better than chance, it is true that more than half did not. This is not surprising because we observed a relatively small number of trials for each neuron and ROC area is a relatively subtle measure (Britten et al., 1996). It is also consistent with models of object recognition that posit that object identity is encoded across large populations of neurons (Tanaka et al., 1991; Riesenhuber and Poggio, 2000; Tsunoda et al., 2001; Brincat and Connor, 2004). Indeed, our small population of neurons correctly predicted the monkey's choice on ~60% of the trials (ROC area of ~0.60).

Choice probability, which is the area under the ROC curve, has been used extensively in studies of motion (Celebrini and Newsome, 1994; Britten et al., 1996), speed (Liu and Newsome, 2005), and depth (Uka and DeAngelis, 2004; Uka et al., 2005; Nienborg and Cumming, 2006) discrimination. Our population ROC area is in the high range of reported choice probability values but is similar to those reported previously. Although it is difficult to make comparisons across brain regions and tasks (in our case a drastically different task), a relatively high choice probability may reflect the position of IT in the visual hierarchy (Nienborg and Cumming, 2006); a large choice probability suggests that IT is an area whose outputs form the basis for decisions about object identity.

To our knowledge, this study is the first to show the usefulness of quantitative measures of choice behavior for studying the role of IT in recognition-related tasks. Previously, Sheinberg and Logothetis (1997) demonstrated correlations between internal perceptual states and neuronal firing in IT in a binocular-rivalry paradigm. However, the ability of visual neurons to strongly predict one's perception of ambiguous figures may reflect the winner-take-all nature of multistable perception. This may explain why choice probabilities for multistable figure representations in MT are larger than for random dot motion displays (Dodd et al., 2001). In the current paradigm, the visual stimulus was not ambiguous at the time we measured firing rate modulations, and thus our results reflect a correlation between behavioral choice and neural response under equivalent stimulus conditions. It is true that the exact composition of the surrounding noise pattern changed on every trial, but the noise background did not alter the physical appearance of the objects themselves. Although it is possible that the surrounding noise pattern could

have modulated the responses of our IT population, we chose to use a unique noise pattern on every trial to simulate the complexity and variability of natural scenes in a controlled manner.

The results of our swap trial analysis do not seem to be an artifact of an artificial stimulus manipulation (i.e., the target swap). Our task allowed us to compare measures of behavioral choice during swap trials as well as response time during normal trials with neural activity in the same experiment. We found that, during normal trials, when no target change occurred, response times were inversely correlated with neural activity in the same population of IT neurons. The stronger the activity in IT neurons coding for the target before target fixation, the faster the monkey was to make a button response. Analogous to the results from the swap trials, this correlation between neural modulation and response time could not be explained by variability in targeting saccade direction, presaccadic target eccentricity, or presaccadic fixation duration across trials. Sheinberg and Logothetis (2001) also found a correlation between response time and neural response, but their results were confounded by stimulus configuration and the exact amplitude of targeting saccades. We used structured stimulus arrays that produced stereotypical eye movement patterns in all analyzed trials and still found a significant response time–firing rate correlation in ~40% of our cells and across our population of cells. These results emphasize the point that, as a population, IT cells respond in a graded manner and their activity is tightly linked to recognition performance.

Others have looked at the correlation between neuronal latency and response time for isolated visual stimuli. Eifuku et al. (2004) found a small number of IT neurons with significant correlations between their onset latency and the monkey's response time. Conversely, DiCarlo and Maunsell (2005) found little covariation between behavioral response time and neuronal latency in IT. Our results differ from these studies in two important ways. First, we measured neuronal response magnitude and, second, we did so at a time when neural responses were near threshold for target recognition (as shown by our swap trial data). Thus, under conditions in which target identity is uncertain, the variability in IT response magnitude is predictive of the time of recognition.

How might our results be related to the deployment of attention in our task? An increase in general arousal during some trials might lead to both increased neuronal excitability and enhanced target detection, accounting for our results. However, we did not find any correlation between behavioral choice and neural response when the preswap target was an ineffective target (first-saccade trials, population ROC area of 0.51, $p = 0.55$; second-saccade-trials, population ROC area of 0.48, $p = 0.16$). Thus, any effects of attention must be specific to those neurons encoding the current target. Of course, it is still possible that, on some trials, "more" attention was allocated to the location of the target and these were the trials in which the monkeys recognized the preswap target or recognized a static target faster. We specifically analyzed trials in which the monkeys made stereotyped saccades toward the target. Before the initiation of an eye movement, there is a complementary shift of visual attention to the saccade target (Kowler et al., 1995). Thus, on all trials, attention was internally directed toward the target during the time window we analyzed (±100 ms around the initiation of the targeting saccade). Still, there is some evidence that attentional effects in visual cortex are modulated with behavioral performance (Cook and Maunsell, 2002). Our results are consistent with the effects of attention on IT responses, namely increases in firing rate for attended stimuli (Moran and Desimone, 1985; Chelazzi et al., 1993, 2001), and

may reflect a common neural mechanism for attention- and choice-related neural modulation (Krug, 2004).

Our results are also consistent with an accumulation of evidence regarding object identity in IT as the visual search progresses. This is similar to the accumulation of evidence accounting for visual motion classification observed in parietal cortex (Huk and Shadlen, 2005). Populations of IT neurons that are selective for the target accrue visual evidence until their activity reaches a threshold level, at which point the monkey commits to a response. During swap trials, the monkey's decision is reflected in his choice of the preswap or postswap target. During normal trials, the time of recognition is observable in the monkey's response time. Information regarding target identity in populations of IT neurons may be read out by other areas of the brain involved in object recognition (e.g., prefrontal cortex) as well as areas involved in programming the manual response (e.g., parietal reach region).

Visual search proceeds differently when comparing the initial saccade after the onset of a search array and subsequent saccades (Motter and Belky, 1998; Findlay et al., 2001). Overall, we did not find large differences in the ability of IT neurons to predict the monkey's choice behavior across abrupt-onset and mid-search conditions. Individual neurons were, in general, equally good at predicting behavioral choices, and the population of neurons performed equally well in both cases. It is true that a larger percentage of neurons produced significant response time–firing rate correlation coefficients for first-saccade trials (37%) compared with second-saccade trials (22%). However, because of the fact that the manual response takes place later in the trial during second-saccade trials (by design), there is more opportunity for response variability. Furthermore, we found no difference between the population correlation coefficients across trial type (−0.15 in both cases). Thus, our results are not likely attributable to artificial laboratory conditions, but rather we are observing how IT neurons naturally develop their responses in a complex visual environment with multiple objects and unconstrained eye movements.

Our choice behavior and response time–neural activity correlation results provide converging evidence of a critical role for IT cortex in processes that support recognition and classification of complex visual objects. These results, focusing on the trial-to-trial variability in neural activity, extend this role to more natural conditions and to include parafoveal recognition. Our paradigm might also be useful for microstimulation experiments that attempt to bias behavioral choices during near-threshold recognition (Afraz et al., 2006). Such studies, along with more carefully designed paradigms, will undoubtedly uncover greater details about the role of IT in visual recognition.

References

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

Brincat SL, Connor CE (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex. Nat Neurosci 7:880–886.

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci 13:87–100.

Celebrini S, Newsome WT (1994) Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. J Neurosci 14:4109–4124.

Chelazzi L, Miller EK, Duncan J, Desimone R (1993) A neural basis for visual search in inferior temporal cortex. Nature 363:345–347.

Chelazzi L, Miller EK, Duncan J, Desimone R (2001) Responses of neurons in macaque area V4 during memory-guided visual search. Cereb Cortex 11:761–772.

Colby CL, Duhamel JR, Goldberg ME (1996) Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. J Neurophysiol 76:2841–2852.

Cook EP, Maunsell JH (2002) Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. J Neurosci 22:1994–2004.

Dean P (1974) The effect of inferotemporal lesions on memory for visual stimuli in rhesus monkeys. Brain Res 77:451–469.

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

DiCarlo JJ, Maunsell JH (2003) Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. J Neurophysiol 89:3264–3278.

DiCarlo JJ, Maunsell JH (2005) Using neuronal latency to determine sensory-motor processing pathways in reaction time tasks. J Neurophysiol 93:2974–2986.

Dodd JV, Krug K, Cumming BG, Parker AJ (2001) Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. J Neurosci 21:4809–4821.

Efron B, Tibshirani R (1993) An introduction to the bootstrap. New York: Chapman and Hall.

Eifuku S, De Souza WC, Tamura R, Nishijo H, Ono T (2004) Neuronal correlates of face identification in the monkey anterior temporal cortical areas. J Neurophysiol 91:358–371.

Findlay JM, Brown V, Gilchrist ID (2001) Saccade target selection in visual search: the effect of information from the previous fixation. Vision Res 41:87–95.

Fischer B, Boch R (1981) Selection of visual targets activates prelunate cortical cells in trained rhesus monkey. Exp Brain Res 41:431–433.

Goldberg ME, Bushnell MC (1981) Behavioral enhancement of visual responses in monkey cerebral cortex. II. Modulation in frontal eye fields specifically related to saccades. J Neurophysiol 46:773–787.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

Gross C (1973) Visual functions of inferotemporal cortex. In: Handbook of sensory physiology, Vol 7 (Autrum H, Jung R, Lowenstein WR, McKay DM, Teuber HL, eds), pp 451–482. Berlin: Springer.

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

Huk AC, Shadlen MN (2005) Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. J Neurosci 25:10420–10436.

Kowler E, Anderson E, Dosher B, Blaser E (1995) The role of attention in the programming of saccades. Vision Res 35:1897–1916.

Krug K (2004) A common neuronal code for perceptual processes in visual cortex? Comparing choice and attentional correlates in V5/MT. Philos Trans R Soc Lond B Biol Sci 359:929–941.

Leopold DA, Maier A, Logothetis NK (2003) Measuring subjective visual perception in the nonhuman primate. J Consciousness Studies 10:115–130.

Liu J, Newsome WT (2005) Correlation between speed perception and neural activity in the middle temporal visual area. J Neurosci 25:711–722.

Logothetis NK, Sheinberg DL (1996) Visual object recognition. Annu Rev Neurosci 19:577–621.

Matin E (1974) Saccadic suppression: a review and an analysis. Psychol Bull 81:899–917.

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

Moody SL, Wise SP, di Pellegrino G, Zipser D (1998) A model that accounts for activity in primate frontal cortex during a delayed matching-to-sample task. J Neurosci 18:399–410.

Moore T (1999) Shape representations and visual guidance of saccadic eye movements. Science 285:1914–1917.

Moore T, Tolias AS, Schiller PH (1998) Visual representations during saccadic eye movements. Proc Natl Acad Sci USA 95:8981–8984.

Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229:782–784.

Motter BC, Belky EJ (1998) The zone of focal attention during active visual search. Vision Res 38:1007–1022.

Nienborg H, Cumming BG (2006) Macaque v2 neurons, but not v1 neurons, show choice-related activity. J Neurosci 26:9567–9578.

Op De Beeck H, Vogels R (2000) Spatial sensitivity of macaque inferior temporal neurons. J Comp Neurol 426:505–518.

Perrett DI, Rolls ET, Caan W (1982) Visual neurones responsive to faces in the monkey temporal cortex. Exp Brain Res 47:329–342.

Pollatsek A, Rayner K, Collins WE (1984) Integrating pictorial information across eye movements. J Exp Psychol Gen 113:426–442.

Pollatsek A, Rayner K, Henderson JM (1990) Role of spatial location in integration of pictorial information across saccades. J Exp Psychol Hum Percept Perform 16:199–210.

Rainer G, Miller EK (2000) Effects of visual experience on the representation of objects in the prefrontal cortex. Neuron 27:179–189.

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

Sheinberg DL, Logothetis NK (1997) The role of temporal cortical areas in perceptual organization. Proc Natl Acad Sci USA 94:3408–3413.

Sheinberg DL, Logothetis NK (2001) Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. J Neurosci 21:1340–1350.

Swets JA (1996) Signal detection theory and ROC analysis in psychology and diagnostics. Mahwah, NJ: Erlbaum.

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

Tanaka K, Saito H, Fukada Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J Neurophysiol 66:170–189.

Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M (2001) Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. Nat Neurosci 4:832–838.

Uka T, DeAngelis GC (2004) Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. Neuron 42:297–310.

Uka T, Tanabe S, Watanabe M, Fujita I (2005) Neural correlates of fine depth discrimination in monkey inferior temporal cortex. J Neurosci 25:10796–10802.

Wilcox R (1997) Introduction to robust estimation and hypothesis testing. San Diego: Academic.