

Experience-Dependent Sharpening of Visual Shape Selectivity in Inferior Temporal Cortex

David J. Freedman^{1,2,3,4}, Maximilian Riesenhuber⁵, Tomaso Poggio^{3,6} and Earl K. Miller^{1,2,3}

¹The Picower Institute for Learning and Memory, ²RIKEN-MIT Neuroscience Research Center, and ³Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, ⁴Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA, ⁵Department of Neuroscience, Georgetown University Medical Center, Washington, DC 20007, USA and ⁶Center for Biological and Computational Learning and McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

Whereas much is known about the visual shape selectivity of neurons in the inferior temporal cortex (ITC), less is known about the role of visual learning in the development and refinement of ITC shape selectivity. To address this, we trained monkeys to perform a visual categorization task with a parametric set of highly familiar stimuli. During training, the stimuli were always presented at the same orientation. In this experiment, we recorded from ITC neurons while monkeys viewed the trained stimuli in addition to image-plane rotated versions of those stimuli. We found that, concomitant with the monkeys' behavioral performance, neuronal stimulus selectivity was stronger for stimuli presented at the trained orientation than for rotated versions of the same stimuli. We also recorded from ITC neurons while monkeys viewed sets of novel and familiar (but not explicitly trained) randomly chosen complex stimuli. We again found that ITC stimulus selectivity was sharper for familiar than novel stimuli, suggesting that enhanced shape tuning in ITC can arise for both passively experienced and explicitly trained stimuli.

Keywords: categorization, inferior temporal cortex, learning, object recognition, vision

Introduction

The primate inferior temporal cortex (ITC) plays an important role in visual recognition and learning. Damage to the ITC in monkeys produces deficits in visual discrimination, shape perception, perceptual learning (Kluver and Bucy 1938, 1939; Blum and others 1950; Mishkin 1954, 1966; Mishkin and Pribram 1954), as well as category-specific agnosias, deficits in the ability to recognize stimuli from a particular category (e.g., faces) (Damasio and others 1982). Correspondingly, neurophysiological recordings have revealed that ITC neurons are selectively activated by complex shapes (Gross 1973; Bruce and others 1981; Perrett and others 1982; Desimone and others 1984; Logothetis and Sheinberg 1996; Tanaka 1996; Op de Beeck and others 2001; Brincat and Connor 2004) and that their activity is correlated with monkeys' perceptual states (Leopold and Logothetis 1996; Sheinberg and Logothetis 2001).

There is a tacit assumption that ITC neuronal response properties are shaped by learning and experience. This is certainly reasonable and expected; primates become "experts" at visual discrimination and recognition with experience. Examples include orientation-specific enhancement in shape discrimination (Gauthier and others 1999), enhanced ability to discriminate between similar items (Logothetis and others 1995; Sowden and others 2000), or improved recognition of stimuli that are degraded (Gold and others 1999; Rainer and

Miller 2000). Experience-dependent plasticity is further suggested by observations that ITC neurons can respond highly selectively to familiar stimuli, such as photorealistic images of familiar natural objects and scenes (Li and others 1993; Booth and Rolls 1998; Vogels 1999; Sheinberg and Logothetis 2001; Freedman and others 2003) and familiar geometric shapes (Miyashita 1993; Logothetis and others 1995; Tanaka 1996; Kobatake and others 1998; Baker and others 2002), that stimulus repetition, familiarity, and salience can modulate ITC activity (Miller and others 1991, 1993; Fahy and others 1993; Li and others 1993; Jagadeesh and others 2001), and that visual experience results in increased clustering of neurons that respond selectively to trained stimuli (Erickson and others 2000). In addition, several previous studies have shown that ITC neuronal responses are modified by long-term visual experience or training. Logothetis and others (1995) found that ITC neurons responded selectively to recently learned views of novel objects and showed orientation-dependent responses to objects during image-plane rotation. Kobatake and Tanaka (1998) conducted ITC recordings from anesthetized monkeys and found sharpened tuning for trained compared with novel stimuli. Booth and Rolls (1998) found ITC responses to familiar objects that were view invariant, showing similar responses to an object irrespective of the viewpoint. Sigala and Logothetis (2002) found an enhanced representation of shape features that were relevant for categorizing sets of familiar stimuli. Baker and others (2002) found enhanced tuning for conjunctions of familiar stimulus feature pairs that were experienced together. Here, we add to this lexicon. In this study, we use a parametric set of visual stimuli (cat and dog morphs) and a parametric variation of those stimuli (image-plane rotation) that allows us to finely probe training-induced neuronal selectivity. This also allows for a detailed comparison of neuronal stimulus tuning with monkeys' behavioral performance during a visual categorization task with these stimuli.

Previously, we trained monkeys to categorize a set of prototypes and morphs between them into 2 categories, "cats" and "dogs" (Fig. 1). The morphing procedure resulted in a set of stimuli that smoothly changed their physical appearance from 1 prototype to another. Once the monkeys were trained, we recorded from neurons in prefrontal cortex (PFC) and ITCs during performance of a categorization task (Freedman and others 2001, 2002, 2003). These recordings revealed that neurons in PFC were often "category tuned," conveying reliable information about category membership and relatively little information about individual stimuli within each category. By contrast, the majority of neurons in ITC showed

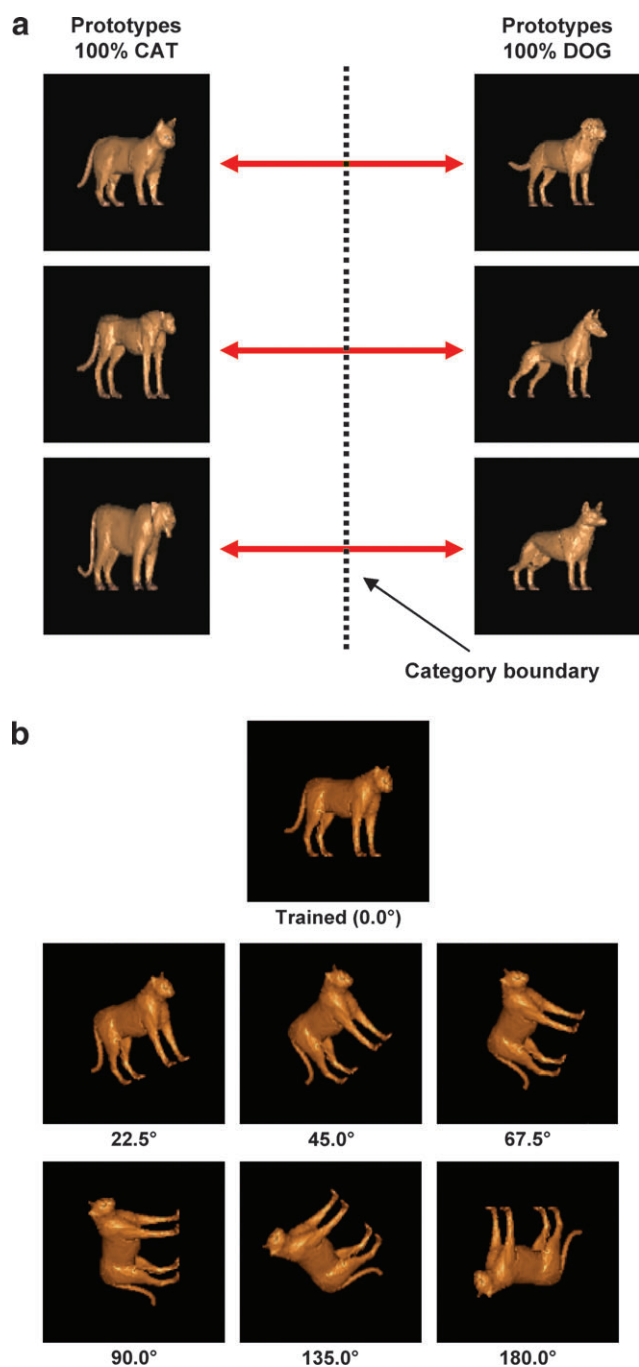


Figure 1. Organization of stimulus set. (a) The 6 prototype images and 3 morph lines (depicted by red arrows). The stimulus set was composed of 18 unique images: 6 prototypes (as shown) and 4 images evenly placed (20%, 40%, 60%, and 80%) along the 3 lines (in red) connecting pairs of prototypes. Monkeys were trained to categorize stimuli in 2 categories, cats and dogs, that were separated by a learned category boundary (depicted by the vertical dotted line). (b). Examples of a stimulus shown at each orientation used during the experiment. Each of the 18 stimuli was shown at 7 two-dimensional image-plane rotated orientations (0°, 22.5°, 45°, 67.5°, 90°, 135°, and 180°). During training, stimuli were always presented at the 0° orientation.

shape tuning; they tended to show detailed selectivity for individual stimuli and little selectivity based on category membership (Freedman and others 2003). In that study, the robust selectivity of many ITC neurons for images in this set of relatively similar stimuli raised the possibility that it was a result

of extensive experience with these stimuli during training. In this study, we test this hypothesis by comparing the strength of stimulus selectivity for these stimuli at both the trained orientation and untrained image-plane rotated orientations.

Over the course of these prior experiments, monkeys had extensive experience with the stimuli at a single fixed (trained) orientation (see Fig. 1). Here, we recorded from ITC while monkeys viewed these stimuli at the trained and 6 image-plane rotated orientations. This revealed that, although average ITC neuronal firing rates were not significantly greater at the trained orientation, the strength and sharpness of neuronal selectivity decreased as stimuli were rotated away from the trained orientation. This decrease in selectivity correlated with a deterioration of behavioral performance for rotated stimuli (observed during separate behavioral testing), suggesting a link between training-induced neural selectivity and behavioral performance. To determine if the effects of training on neural selectivity were similar for other stimuli that were passively experienced (and not explicitly trained), we examined ITC neurons' responses to randomly chosen novel and familiar stimuli and found a similar sharpening of stimulus selectivity with familiarity. Together, the results of this study suggest that experience can enhance ITC shape selectivity independent of increases in average neural firing rate.

Experimental Procedures

Subjects

Three female adult rhesus monkeys (*Macacca mulatta*) weighing 6.0, 6.4, and 6.6 kg were used in this study. Using previously described methods (Miller and others 1993), they were implanted with recording hardware. All surgeries were performed under sterile conditions while the animals were anesthetized with isoflurane. The animals received postoperative antibiotics and analgesics and were handled in accord with National Institutes of Health and the Massachusetts Institute of Technology Committee on Animal Care guidelines.

Stimuli and Behavioral Task

For an earlier experiment, monkeys were initially trained to perform a delayed match-to-category (DMC) task (Freedman and others 2001, 2002, 2003), which required grouping a continuous set of stimuli into discrete categories. A large continuous set of images was generated from 3 cat prototypes and 3 dog prototypes (Fig. 1) using a novel algorithm (Shelton 2000; Freedman and others 2001, 2002). It found corresponding points between one of the prototypes and the others and then computed their differences as vectors. Morphs were created by linear combinations of these vectors added to that prototype. For more information, see <http://www.cs.ucr.edu/~cshelton/corr/>. By morphing different amounts of the prototypes, we could generate thousands of unique images, continuously vary shape, and precisely define a category boundary. The stimuli differed along multiple feature dimensions and were smoothly morphed, that is, without sudden appearance or disappearance of any feature. They were 4.2° in diameter and had identical color, shading, and scale. Monkeys were trained (over the course of several months) to indicate (by releasing a lever) whether 2 successively presented stimuli ("sample" then "test") were from the same category.

For the neural recordings in the present experiment, monkeys "passively viewed" these stimuli one at a time; they were only rewarded for maintaining fixation (within $\pm 2.0^\circ$ of a 0.3° square fixation point at the center of the cathode ray tube for the duration of the trial. Eye movements typically were much smaller than the allowed window. Monkeys' instantaneous eye position exceeded $\pm 1.0^\circ$ of the fixation point during the early or late stimulus epochs on approximately 4.0% of the trials. Eye movements were monitored and stored using an infrared eye-tracking system (Iscan, Cambridge, MA) at a sampling rate of 120 Hz. Monkeys initiated each trial by acquiring fixation of the fixation target.

Then, after 500 ms of fixation, a stimulus was presented at the center of the gaze for 600 ms. If monkeys continued to maintain fixation, they received drops of apple juice as a reward 100 ms after stimulus offset. This was followed by a 1500- to 2500-ms intertrial interval during which fixation was not required. The sequence of stimuli was determined pseudorandomly. Monkeys typically performed more than 10 correct repetitions for each unique stimulus.

During neurophysiological recordings, we used 18 cat and dog stimuli from each of the 6 levels of cat/dog blends (cat/dog) (100:0, 80:20, 60:40, 40:60, 20:80, and 0:100) along the 3 morph lines that crossed the category boundary (the red lines in Fig. 1*a*). These stimuli were presented at the trained orientation and 6 image-plane rotations (Fig. 1*b*, counterclockwise degrees relative to trained orientation: 0°, 22.5°, 45.0°, 67.5°, 90.0°, 135.0°, and 180.0°).

For other recording sessions ($N = 298$ neurons across both monkeys over the course of 4-8 weeks of daily recording sessions), we used 20 novel and 20 familiar stimuli, colorful shapes, drawings, and photographs that were randomly chosen from a large bank of stored images (from the World Wide Web and Corel Image Library, Fig. 2). Stimuli were made familiar by giving the monkeys at least 14 sessions (2 weeks of daily sessions) of experience with them in the passive viewing task. Each “familiar” stimulus was shown approximately 50 times per training session. During each recording session, these familiar stimuli were pseudorandomly interleaved with 20 novel stimuli (chosen from the same set of random images) that, at the start of each session, had never been seen before by the monkeys. A new set of 20 novel stimuli was used for each recording session. The same set of familiar stimuli was used for all recording sessions. Because of differences between the cat/dog and novel/familiar stimuli and the amount and type of experience (explicit training vs. passive viewing) with them, only qualitative (but not quantitative) comparisons between the 2 sets of results should be made.

Following the completion of neuronal recordings, each monkey’s behavioral performance on the DMC task with the rotated cat and dog stimuli was evaluated during several behavioral testing sessions. They performed the DMC task with the 18 cat and dog stimuli as samples, presented at each of the 7 orientations. The test stimuli were 36 randomly generated morphs (at each orientation). On a given trial, the sample and test stimuli were always at the same orientation, and the test stimuli unambiguously belonged to a given category: they were always

chosen to be at a distance of at least 20% from the boundary. Trials from all 7 orientations were pseudorandomly interleaved.

Recording Methods

ITC recordings were conducted between Anterior-Posterior = 14-20 mm and Lateral = 15-21 mm. ITC recording locations, as determined by stereotaxic coordinates, magnetic resonance imaging scans, and white-gray matter transitions encountered during electrode penetrations, were in areas TEa, TE_m, TE₂, and TE₁ (Paxinos and others 2000). The locations of ITC recordings were similar to those reported in studies by several laboratories (Logothetis and others 1995; Booth and Rolls 1998; Kobatake and others 1998; Baker and others 2002; Freedman and others 2003). No attempt was made to prescreen neurons for stimulus selectivity. Instead, while advancing electrodes into ITC, we presented the monkeys with randomly chosen pictures and photographs (from the Corel Image Library) that did not resemble the cat and dog stimuli and focused our recordings on sites that were visually responsive to these images. Neuronal waveforms were amplified, digitized, and stored for offline sorting into individual neuron records using principal components analysis clustering software (Plexon Inc., Dallas, TX).

Data Analysis

Many ITC neurons show a strong, phasic discharge shortly after stimulus onset followed by a lower level of sustained activity (see Fig. 3). Thus, for many analyses, average neuronal activity was calculated in 3 time epochs: baseline, early stimulus (phasic), and late stimulus (sustained) epochs. Early stimulus period activity was averaged over a 100-ms epoch beginning 80 ms after stimulus onset. Late stimulus period activity was averaged over a 100-ms window beginning 180 ms after stimulus onset. We also calculated the average baseline firing rate over the 500 ms of fixation preceding stimulus presentation.

Neuronal sensitivity to individual stimuli and changes in stimulus orientation was determined with a two-way analysis of variance (ANOVA) with stimuli ($N = 18$) and orientation ($N = 7$) as factors, evaluated at $P < 0.01$. We classified neurons as stimulus selective if they showed a main effect of stimulus and/or a significant interaction between the stimulus and orientation factors. Neurons that had a main effect of orientation or an interaction between orientation and stimulus factors were classified as orientation selective.

Examples of Stimuli Used for Passive Viewing

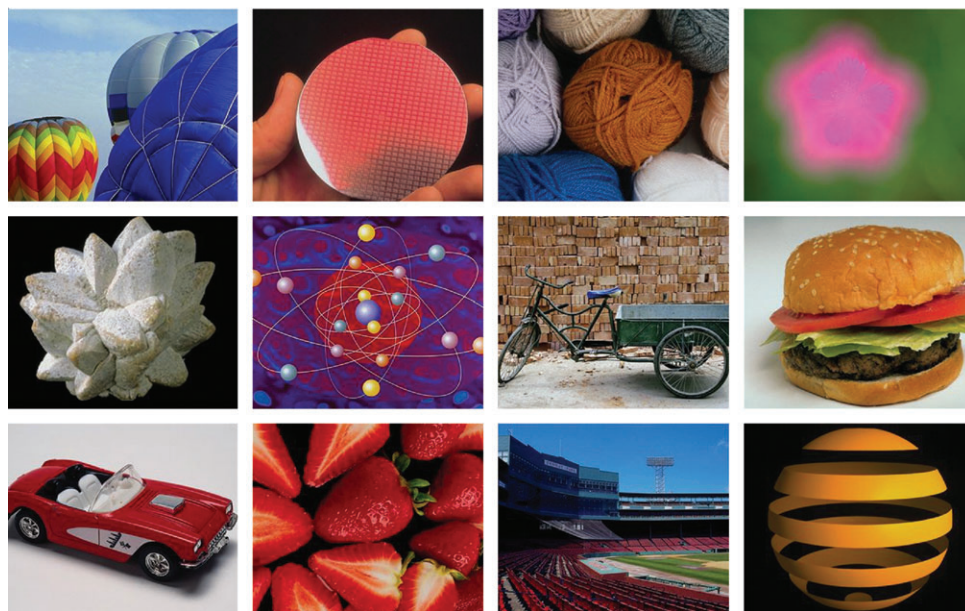


Figure 2. Examples of other stimuli used for ITC neuronal recordings ($N = 298$ neurons from 2 monkeys) during passive viewing. During each passive viewing recording session, monkeys were shown 20 novel stimuli that the monkey had not seen before that session and 20 familiar stimuli that the monkey had seen hundreds of times (during passive viewing) in the weeks prior to recordings.

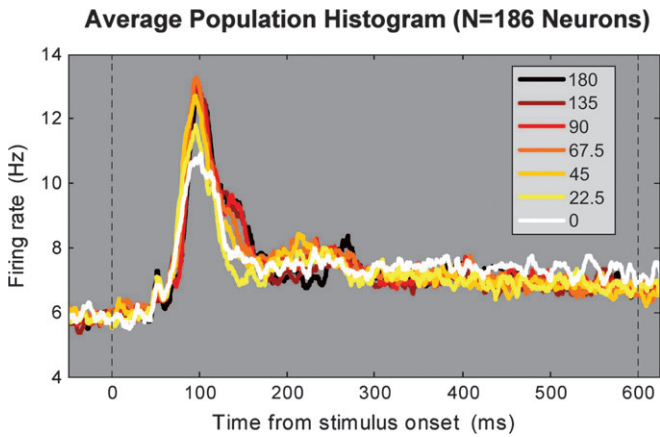


Figure 3. Population histogram of average activity across all recorded ITC neurons ($N = 186$) to stimuli at trained and 6 image-plane rotated orientations. The 7 traces correspond to average firing rate across all 18 stimuli at each orientation, respectively. The vertical dotted lines at time 0 and 600 indicate the time of stimulus onset and offset, respectively.

For each neuron, we computed the strength and sharpness of stimulus selectivity by several measures. We computed the strength of stimulus selectivity (S) at each of the 7 orientations according to the following formula where R_{\min} is the firing rate to the “worst” stimulus and R_{\max} is the firing rate to the “best” stimulus:

$$S = \frac{R_{\max} - R_{\min}}{R_{\max} + R_{\min}}$$

To measure the broadness of stimulus selectivity, we determined the number of stimuli (of 18) at each orientation that elicited a significant response greater than baseline according to a one-tailed t -test at $P < 0.01$. For this analysis, lower values indicate sharper tuning (i.e., a neuron that responds to only a small subset of the 18 stimuli is more sharply tuned), whereas higher values indicate broader tuning.

The breadth of selectivity was calculated by first normalizing the neuronal responses to the 18 stimuli on a scale from 0 to 1. We then computed the median firing rate across the 18 stimuli and subtracted this value from 1.0. By so doing, higher values of the index (i.e., smaller values of the median) indicate sharper tuning (i.e., weak responses to the majority of stimuli with strong responses to a small number of stimuli), whereas lower index values indicate broader tuning. This index gives similar results as the depth of selectivity index used by Rainer and Miller (2000). Qualitatively similar results were obtained when this index was computed using raw, instead of normalized, firing rates.

To evaluate selectivity effects for single neurons for each of these measures of stimulus selectivity, we computed linear regressions for each neuron’s selectivity index values across the 7 orientations and report the number of neurons that show a significant regression fit (at $P < 0.05$) with a slope in the same direction as the population average. This gives an indication about the number of neurons that behaved in a similar way as the population averages.

Results

Neurophysiology: General Properties

We recorded the activity of 186 ITC neurons from 2 monkeys (monkey L: $N = 134$, monkey S: $N = 52$) during passive viewing of each of the 18 cat/dog stimuli (see Experimental Procedures) at each of the 7 different orientations (Fig. 1). We recorded the activity of 298 different ITC neurons (monkey L: $N = 157$, monkey F: $N = 141$) during passive viewing of the 40 “other” stimuli (20 novel and 20 familiar complex, multicolored images, see Experimental Procedures and Fig. 2).

Table 1

Incidence of stimulus selective neurons according to two-way ANOVA at $P < 0.01$

Factors	Number of stimulus selective neurons ($N = 186$)	
	Early stimulus (80–180 ms)	Late stimulus (180–280 ms)
Stimulus	128 (69%)	127 (68%)
Orientation	4 (2%)	5 (3%)
Interaction	67 (36%)	106 (57%)
Stimulus only, no interaction	68 (37%)	37 (20%)
Orientation only, no interaction	0 (0%)	1 (1%)

The neurons were preselected for visual responsiveness (see above), and both populations had a high incidence of stimulus selectivity. We conducted a two-way ANOVA on each neuron’s average activity during presentation of the cat and dog stimuli using “stimulus” (18 levels, 1 for each cat and dog) and “orientation” (7 levels) as factors (see Table 1). By this measure, a majority of neurons was stimulus and/or orientation selective (at $P < 0.01$, $N = 135/186$ or 73%, main effect of stimulus and/or orientation or their interaction, see Table 1). Likewise, a majority (60.4% or 180/298) of neurons tested with the other (non-cat and dog) stimuli were stimulus selective for either the novel stimuli (47% or 141/298 neurons, one-way ANOVA at $P < 0.01$) or the familiar stimuli (49% or 145/298 neurons, one-way ANOVA at $P < 0.01$). Of course, one cannot make direct comparisons of incidence of stimulus selectivity between these populations; the statistical tests have different sensitivities (because of the different number of trials and stimuli), and the stimuli vary along different parameters. So, to guard against any bias in selecting neurons for further examination, the analyses presented below were conducted on all recorded neurons, not just those showing significant selectivity. However, we also conducted all population analyses across the populations of stimulus selective and visually responsive neurons and found qualitatively similar results in both cases. As neuronal properties and effects were similar between the 2 monkeys, we will collapse across them in reporting our results.

Neurophysiology: Cat and Dog Stimuli

Figure 3 shows a population histogram of average activity at the trained and 6 image-plane rotated orientations across the entire population of recorded ITC neurons ($N = 186$). The 7 traces correspond to average firing rate across all 18 stimuli at each orientation, respectively. Average activity during the early stimulus epoch (80–180 ms following stimulus onset) was significantly weaker for stimuli at the trained than 180° orientation (trained = 8.7 Hz, 180° = 9.7 Hz) according to a two-tailed paired sample t -test ($P = 0.013$). Activity during the late stimulus epoch was similar across the 7 orientations (trained = 7.4 Hz, 180° = 7.3 Hz) according to a t -test ($P = 0.69$).

Figure 4 shows examples of neurons tested with the cat and dog stimuli. Plotted on the left are histograms showing each neuron’s average activity (across all 18 stimuli) at each stimulus orientation, and the color plots to the right show the average activity to each stimulus at each orientation. The neuron in Figure 4a showed relatively similar average activity across the 7 orientations during the early, phasic, portion of its response and greater differences during the later, sustained, portion (left plot). A detailed examination of its early stimulus epoch activity as a function of both stimulus and orientation (right plot) indicates greater differences between stimuli when they were at the trained orientation than other rotations; note the greater

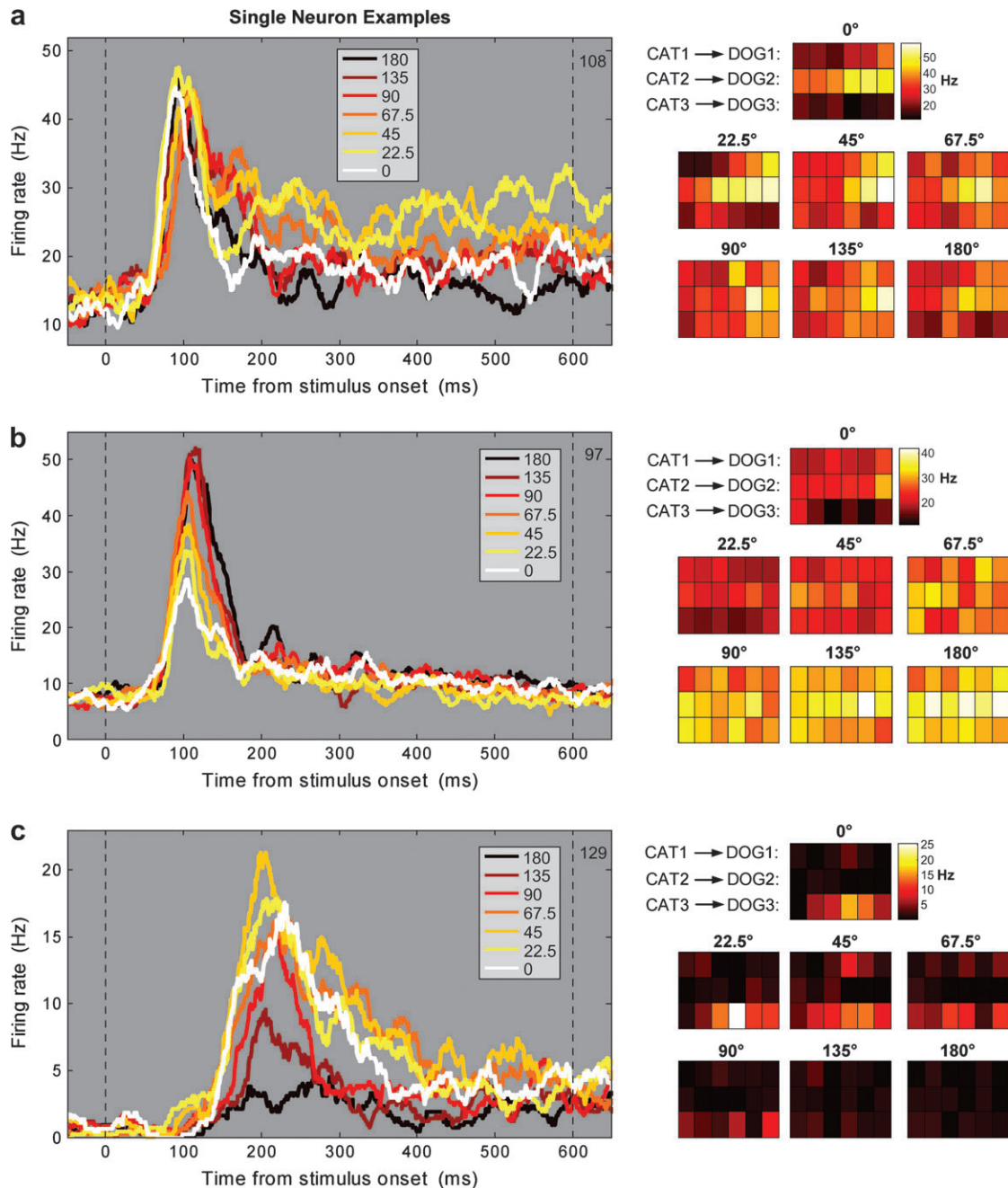


Figure 4. Single neurons examples. Figure (a–c) (left) shows the peristimulus time histograms of activity for 3 representative single neurons to the 7 stimulus orientations. The 7 traces depict the average activity (across all 18 stimuli) of each neuron to stimuli at the 7 orientations, respectively. Activity is aligned to stimulus onset, and stimuli were always presented for 600 ms. The color plots to the right of each histogram show the activity of that neuron (during the early stimulus epoch) to each of the 18 sample stimuli at each of the 7 orientations. The 3 rows correspond to the 3 morph lines, whereas the 6 columns indicate the 6 levels of morphs (from cat prototype to dog prototype) along each morph line (see Fig. 1).

differences in firing rate between the separate tiles (stimuli) at the trained orientation (0°) versus others (maximum – minimum firing rate, trained orientation = 40.0 Hz, 180° = 26.4 Hz). The neuron in Figure 4b showed the strongest average activity for stimuli at the 180° orientation (left plot). But, it showed the sharpest stimulus selectivity at the trained orientation (right plot) than at the more rotated orientations, despite weaker average firing at the trained orientation (maximum – minimum firing rate, trained orientation = 20.2 Hz, 180° = 15.3 Hz). The

neuron in Figure 4c responded, on average, most strongly to stimuli at or relatively near the trained orientation (between 0° and 67.5°) and also showed the strongest stimulus selectivity for stimuli at those orientations (maximum – minimum firing rate, trained orientation = 16.3 Hz, 180° = 2.2 Hz).

Greater selectivity at the trained orientation was even evident across all recorded neurons ($N = 186$), in the phasic portion of the neural response (early stimulus epoch). For every recorded neuron, we computed a standard index of stimulus

selectivity: the difference in early stimulus epoch activity between the preferred and nonpreferred stimuli divided by their sum (see Experimental Procedures). This yielded values ranging from 0 to 1, where larger values indicate greater differences between the stimuli (i.e., more selectivity). Figure 5a shows the average population selectivity index values at

each orientation. Figure 5b shows the difference between neurons' index values at the trained and 180° orientations. Values were highest (i.e., there was greater selectivity) at the trained orientation (0°) and became progressively smaller with increasing rotation. Index values at the trained orientation (0.765) were significantly greater than at the most rotated

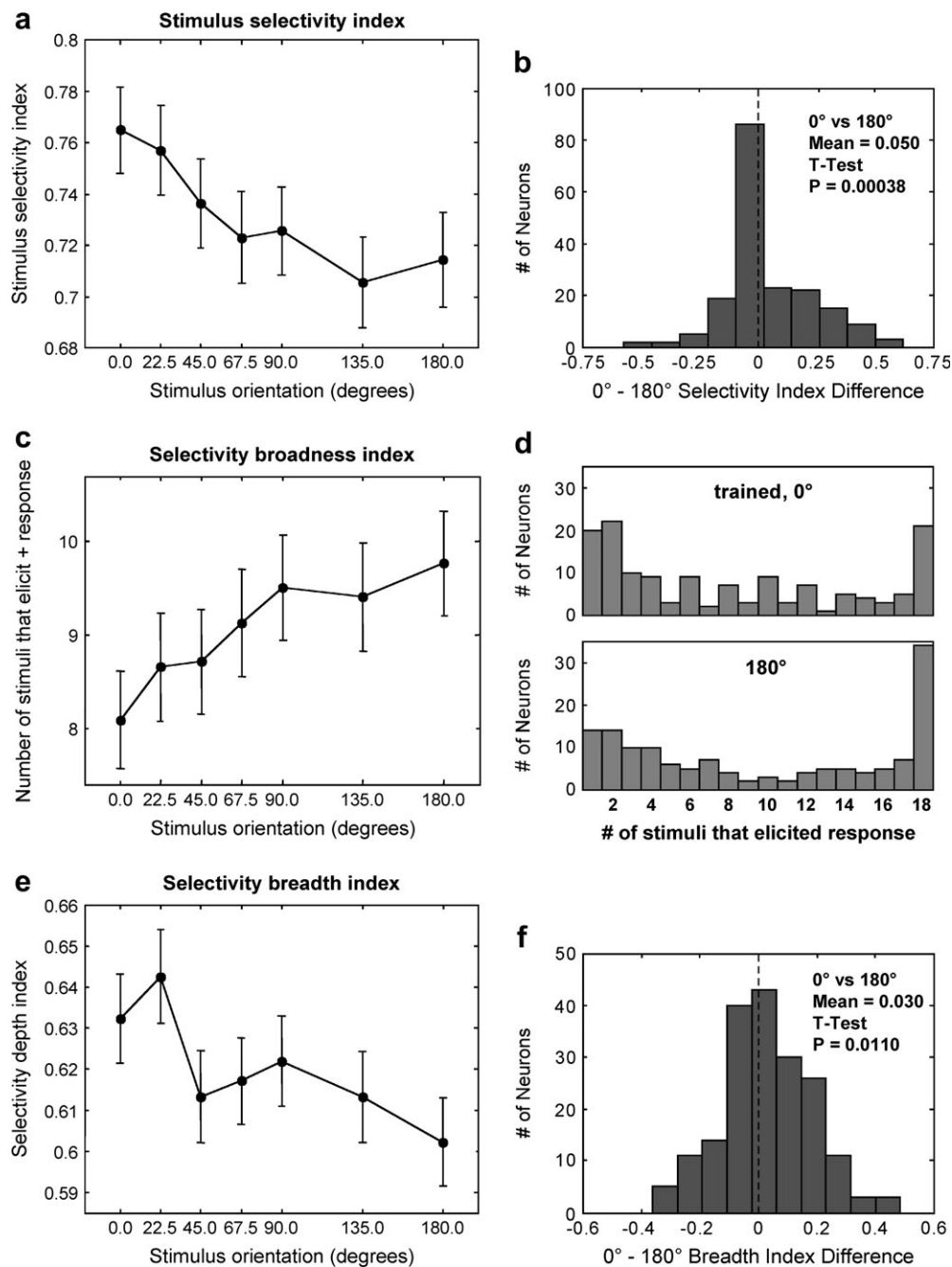


Figure 5. Analysis of stimulus selectivity, selectivity broadness, and selectivity breadth for rotated cat/dog stimuli across the entire population of 186 ITC neurons. (a) Average values of the stimulus selectivity index at each of the 7 orientations. Higher values of the index indicate greater neuronal modulation in response to the best and worst stimulus at that orientation. Lower values indicate the opposite. (b) The distribution of the difference between neurons' trained and 180° stimulus selectivity index values. The positive shift of the distribution indicates greater index values at the trained orientation. (c) Average values of the selectivity broadness index at each of the 7 orientations. Lower values indicate sharper tuning (i.e., responses to a small subset of the 18 stimuli), whereas higher values indicate broader tuning (i.e., responses to most or all stimuli). (d) The number of stimuli (1–18) that elicited a significant response above baseline at the trained (top panel) and 180° (bottom panel) orientations. For a given number of response-eliciting stimuli (on the horizontal axis), the vertical axis shows the number of neurons that responded to that number of stimuli. (e) Average values of selectivity breadth index at each of the 7 orientations. Higher values of the index indicate sharper selectivity among 18 stimuli. Lower values indicate the opposite. (f) The distribution of the difference between neurons' trained and 180° selectivity breadth index values. The positive shift of the distribution indicates greater index values at the trained orientation.

(180°) orientation (0.715, paired sample *t*-test, $P = 3.85 \times 10^{-4}$). There were no significant effects in the weaker, sustained portion of the response (the late stimulus epoch). During the early sample epoch, index values from 28 neurons showed a significant ($P < 0.05$) fit with linear regression (see Experimental Procedures). Of these, 23 had a regression slope in the same (negative) direction as the population average.

The selectivity index described above assesses neural selectivity by comparing just 2 stimuli (preferred vs. nonpreferred). So, we computed other indices that took into account each neuron's activity to all (18) stimuli. They yielded similar results. One was a "selectivity broadness" index (see Experimental Procedures). For each neuron, we determined the number of stimuli at each orientation that elicited a significant response: early stimulus activity significantly greater than baseline (one-tailed *t*-test at $P < 0.01$). So as not to bias the result, we excluded from this analysis any nonresponsive neurons (for which no stimulus elicited a significant response at the orientation under consideration). So, although every included neuron responded to at least 1 stimulus, neurons with lower index values were activated by fewer stimuli, indicating more (or sharper) selectivity. Figure 5*c* shows the average selectivity broadness index values across all neurons. Figure 5*d* gives an indication about the number of stimuli (1-18) that elicited significant responses from these neurons at the trained (top panel) and 180° orientations (bottom panel). For a given number of response-eliciting stimuli (on the horizontal axis), the vertical axis shows the number of neurons that responded to that number of stimuli. For example, a greater number of neurons ($N = 34$) responded to all eighteen 180° rotated stimuli, whereas fewer neurons ($N = 21$) responded to all 18 stimuli at the trained orientation. Neurons, on average, responded to fewer stimuli at the trained orientation (0°: 8.09) than at the most rotated orientation (e.g., 180°: 9.76) according to an unpaired (because of the small number of nonresponsive neurons excluded from the analysis at each orientation) *t*-test ($P = 0.0288$). During the early sample epoch, index values from 41 neurons showed a significant ($P < 0.05$) fit with linear regression (see Experimental Procedures). Of these, 31 had a regression slope in the same (positive) direction as the population average.

Another index, the "selectivity breadth," was computed using the median of the normalized activity of each and every recorded neuron's early stimulus activity to each of the 18 stimuli (see Experimental Procedures). Figure 5*e* shows the average selectivity breadth index at each orientation across all neurons. Figure 5*f* shows the difference between index values at the trained and 180° orientations. Selectivity breadth values were greater (indicating greater selectivity) at and near the trained orientation than at the more rotated orientations. The average index values for both the trained orientation (0°, index = 0.632) and the rotations near the trained orientation were significantly greater than those from the most rotated stimuli (180°, index = 0.602) according to a paired sample *t*-test ($P < 0.05$). During the early sample epoch, index values from 17 neurons showed a significant ($P < 0.05$) fit with linear regression (see Experimental Procedures). Of these, 11 had a regression slope in the same (negative) direction as the population average.

To ensure that these effects were not due to the monkeys' pattern of eye movements changing with experience (i.e., monkeys looked at familiar stimuli different than novel stimuli), we recomputed the analyses of stimulus selectivity, selectivity

breadth, and selectivity broadness by excluding trials in which eye position deviated from the fixation point by more than $\pm 1.0^\circ$ during the early or late stimulus epoch (4.0% of trials) and found equivalent results. In addition, there was no detectable difference in eye movement patterns across the orientations; the horizontal and vertical average eye positions for stimuli at each of the 7 orientations did not differ from one another during the early or late stimulus epochs according to one-way ANOVAs ($P > 0.2$).

For many single neurons (such as those in Fig. 4*a,b*), it seemed that sharper selectivity at the trained orientation was primarily due to weaker responses to nonpreferred stimuli at the trained orientation. To test this across the population, we averaged each recorded neuron's activity ($N = 186$) to its most preferred (best) stimulus and least preferred (worst) stimulus separately at each orientation.

The average early stimulus period activity to the best stimuli was similar across the orientations (Fig. 6*a,b*). It was equally large for the trained orientation (16.35 Hz) and 180° rotation (16.21 Hz), for example (paired sample *t*-test, $P = 0.822$, see Fig. 6*a,b*). Linear regression revealed a significant fit (of best-stimulus firing rates across the 7 orientations) for 47 neurons (32 of these neurons had a positive slope). By contrast, the average population activity to the worst stimuli did differ across orientations (Fig. 6*c,d*). It was "weakest" at the trained orientation (0°: 3.65 Hz) and became progressively stronger as stimuli were rotated away from it (180°: 4.70 Hz, paired *t*-test vs. trained orientation, $P = 9.5 \times 10^{-4}$). Linear regression of these values revealed a significant fit for 42 neurons. Of these, 29 had a regression slope in the same direction (positive) as the population average. In addition, the average activity to the worst stimuli at all orientations was significantly weaker than prestimulus baseline activity (5.83 Hz, the horizontal dotted line on Fig. 6*c*) according to paired sample *t*-tests at $P < 0.01$. During the late stimulus epoch, we did not observe significant differences in average activity across orientations to neurons' best stimuli ($P = 0.974$) or worst stimuli ($P = 0.784$, paired sample *t*-tests comparing activity with 0° and 180° stimuli).

The color plots in Figure 4 also suggest that these neurons seemed to maintain similar stimulus preferences across a wide range of orientations (the stimuli that elicited the best response at the trained orientation were also among the best stimuli at nearby orientations). This was also true across the population. According to the two-way ANOVA described above (with stimulus and orientation as factors), about half of the neurons ($N = 68/128$ or 53%) that showed a main effect ($P < 0.01$) of stimulus during the early stimulus epoch did not show an interaction with orientation (see Table 1), indicating similar selectivity across orientations. To determine the strength of this effect across the population, we determined each neuron's best and worst stimulus at the trained orientation and computed the firing rates to those 2 stimuli at each of the 7 orientations. If neurons' stimulus selectivity was unrelated between orientations, then we would expect to see strong selectivity at the trained orientation (because the best and worst stimuli were determined at that orientation) and no systematic difference between those 2 stimuli at the other 6 orientations. As shown in Figure 7, this analysis instead revealed that neurons responded more strongly (on average across the entire population, $N = 186$) to the best than worst stimuli at all 6 untrained orientations (even though the best/worst preference was determined at the trained orientation) according to a two-way ANOVA

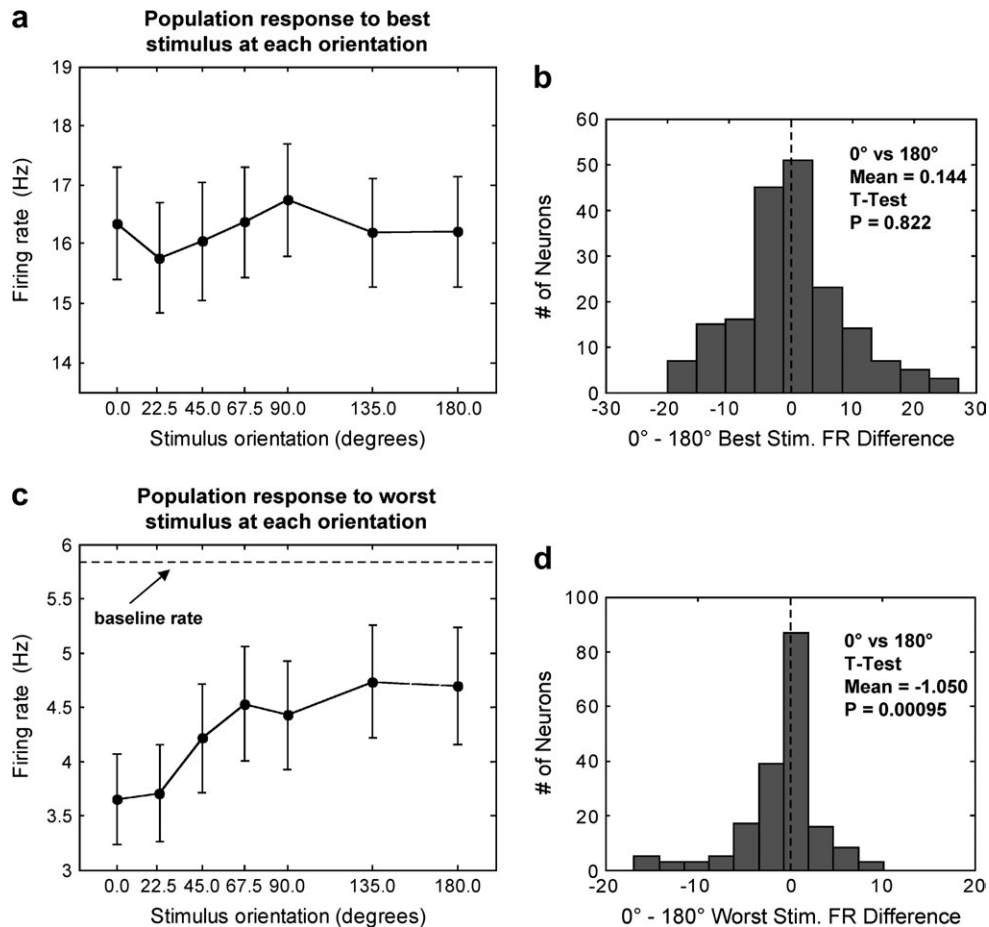


Figure 6. Average responses to each neuron's single best and worst stimulus across the entire population of neurons. (a) Average response to each neuron's single best stimulus at each orientation (computed separately at each orientation). (b) The distribution of the difference between trained and 180° responses to neurons' best stimuli. That the distribution is not significantly shifted away from a mean of 0.0 indicates similar responses at the trained and 180° orientation. (c) Average response to each neuron's single worst stimulus at each orientation. The prestimulus baseline activity level is indicated by the horizontal dotted line. (d) The distribution of the difference between trained and 180° responses to neurons' worst stimuli. The negative shift of the distribution indicates weaker responses to worst (i.e., nonpreferred) stimuli at the trained orientation.

(factor 1: best or worst stimulus as determined at the trained orientation, $P = 1.9 \times 10^{-6}$; factor 2: the 6 untrained orientations, $P = 0.888$; interaction, $P = 0.688$).

Neurophysiology: Other Stimuli

A remaining question is whether enhanced shape tuning depends on explicit training on a discrimination/categorization task or whether it could occur when stimuli are passively experienced (i.e., when the animal is not rewarded for identifying or categorizing the stimuli per se) over the course of several weeks (rather than months of explicit training for the cat/dog stimuli). Another issue is that the images used above resembled "biological" or "natural" objects (i.e., the images resembled real cats and dogs), and the monkeys were trained with the stimuli always appearing in their "canonical" orientation (upright, as animals like cats and dogs are typically observed in nature). Thus, it is critical to consider whether the effects reported above were due to training or instead due to an innate improvement in stimulus processing when natural stimuli appear in their canonical orientation. To address these issues, we recorded from an additional 298 ITC neurons from 2 monkeys (monkey L: 157 neurons, monkey F: 141 neurons) during passive viewing of a set of 20 novel stimuli and 20 highly familiar randomly chosen stimuli (Fig. 2). The passive viewing

task was identical to the passive viewing task used for the cat and dog stimuli (discussed above and see Experimental Procedures). A new set of novel images was used for each recording session, although familiar stimuli were those that the monkey had previously seen approximately 50 times per day for at least 14 behavioral sessions in the weeks immediately prior to recordings.

This yielded results similar to those observed with the cat and dog stimuli: sharper stimulus selectivity for familiar compared with novel stimuli. We also found weaker average neural activity (across all stimuli) for familiar than novel stimuli. This can be seen in average population activity (across all recorded ITC neurons, $N = 298$) shown in Figure 8. The average activity for familiar stimuli was significantly lower than for novel stimuli during both the early stimulus epoch (novel: 13.69 Hz, familiar: 11.90 Hz, t -test, $P = 1.11 \times 10^{-14}$) and late stimulus epoch (novel: 12.56 Hz, familiar: 7.67 Hz, t -test, $P = 4.2 \times 10^{-22}$).

The stimulus selectivity, selectivity breadth, and response sharpness indices were also computed for the other (non-cat and dog) stimuli across all recorded ITC neurons ($N = 298$). We once again found that familiarity seemed to enhance selectivity. The average population stimulus selectivity index was greater for the familiar than novel stimuli during the early stimulus epoch (Fig. 9a,b, familiar index = 0.78, novel index = 0.75,

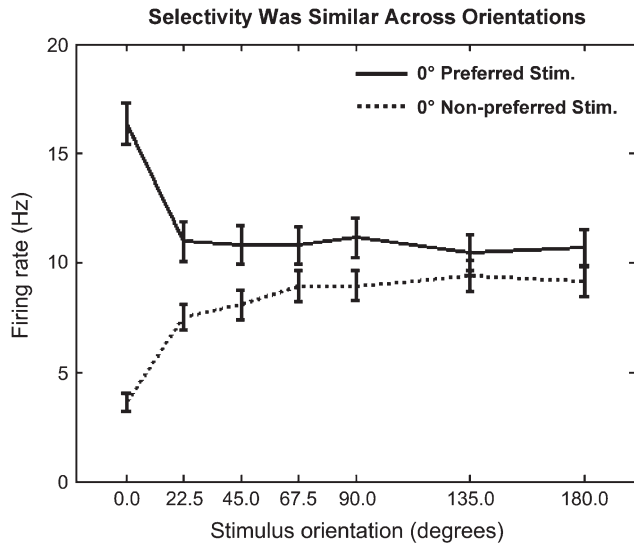


Figure 7. Rotation-invariant stimulus selectivity across the entire population of 186 ITC neurons. Each neuron's best and worst stimulus was determined at the trained (0°) orientation. Firing rates to those 2 stimuli were then computed at each of the 6 rotated orientations. The solid line shows the average firing rate to each neuron's best stimulus (as determined at the trained orientation) across the 7 orientations. The dotted line shows the average firing rate to each neuron's worst stimulus. That the solid line is above the dotted line at all 7 orientations indicated that neurons tended to respond more strongly to the preferred than nonpreferred stimulus (as defined at the trained orientation) despite changes in orientation.

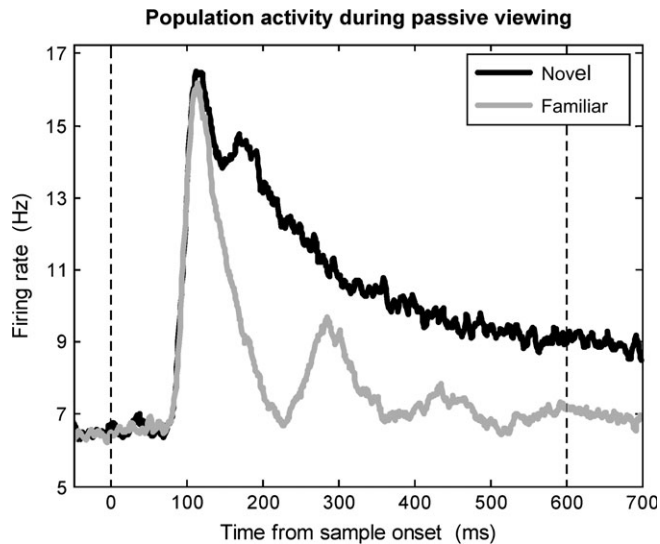


Figure 8. Average population activity across 298 ITC neurons during passive viewing of novel ($N = 20$) and familiar ($N = 20$) stimuli. The 2 traces indicate the average response to novel stimuli (black trace) and familiar stimuli (gray trace). Stimuli were presented for 600 ms.

two-tailed paired sample t -test, $P = 5.6 \times 10^{-5}$) as well as the late stimulus epoch (average index for familiar stimuli = 0.89, compared with 0.82 for novel stimuli, paired sample t -test, $P = 1.30 \times 10^{-11}$). Likewise, the selectivity broadness index (Fig. 9c,d) indicated sharper selectivity for the familiar stimuli during the early stimulus epoch (average number of familiar stimuli eliciting significant response: 9.08 stimuli) than for the novel stimuli (average = 10.32 stimuli, two-tailed t -test, $P = 1.74 \times 10^{-13}$). Figure 5d gives an indication about the number of

stimuli (1–20) that elicited significant responses from these neurons for familiar (top panel) and novel (bottom panel) stimuli. For a given number of response-eliciting stimuli (on the horizontal axis), the vertical axis shows the number of neurons that responded to that number of stimuli. For example, a greater number of neurons ($N = 50$) responded to all 20 novel stimuli, whereas fewer neurons ($N = 36$) responded to all 20 familiar stimuli. These effects were also observed during the late stimulus epoch (average number of familiar stimuli eliciting significant response = 5.51, average number of novel stimuli eliciting significant response = 9.66; t -test, $P = 1.81 \times 10^{-34}$). Finally, the selectivity breadth index (see above and Experimental Procedures) also indicated significantly greater selectivity for familiar stimuli than for novel stimuli during both the early stimulus epoch (Fig. 9e,f; familiar index = 0.69, novel index = 0.67, paired t -test, $P = 0.014$) and the late stimulus epoch (familiar index = 0.73, novel index = 0.71, paired t -test, $P = 0.041$).

We also computed the firing rate for each neuron's single most-preferred (best) and least-preferred (worst) novel and familiar stimuli, as we did for the rotated cat/dog stimuli. For the other (non-cat/dog) stimuli, neuronal activity was significantly weaker for familiar compared with novel stimuli across the entire range of best to worst stimuli. In other words, weaker activity for familiar stimuli was not only observed for non-preferred stimuli (as observed for the cat/dog stimuli at the trained orientation). Instead, both the best and worst familiar stimuli evoked weaker responses than the best and worst novel stimuli, respectively. This is somewhat different from results from the rotated cat/dog stimuli (Fig. 6), in which weaker responses were only observed for nonpreferred stimuli at or near the trained orientation and may reflect different mechanisms for explicitly trained and passively experienced stimuli.

The results of this analysis are shown in Figure 10a–d. As with the cat/dog stimuli, we compared the activity to neurons' single best and worst stimuli (separately for both the novel and familiar sets of other stimuli). This revealed a more general weakening of activity for familiar stimuli than was observed with the rotated cats and dogs. Across the entire recorded population ($N = 298$ neurons), average activity during the early stimulus epoch (80–180 ms poststimulus onset) was weaker for the worst familiar than for the worst novel stimuli (familiar: 5.029 Hz, novel: 6.0417 Hz, Fig. 10a,b) according to a paired t -test ($P = 2.198 \times 10^{-7}$). In contrast with results from the rotated cats and dogs, average activity to the best familiar stimuli was also significantly weaker than for the best novel stimuli (familiar: 23.815 Hz, novel: 25.685 Hz, Fig. 10c,d) according to a paired t -test ($P = 8.27 \times 10^{-4}$).

In addition, average activity to the worst familiar stimuli (5.029 Hz) was significantly weaker than to prestimulus fixation baseline activity (6.163 Hz, the horizontal dotted line in Fig. 10c) according to a paired t -test ($P = 0.0016$). Average activity to the worst novel stimuli (6.0417 Hz) did not differ from baseline levels (6.163 Hz) according to a paired t -test ($P = 0.817$).

Behavior

In order to determine whether training monkeys with stimuli at a fixed orientation resulted in an enhanced ability to categorize those stimuli, we tested the monkeys' abilities to perform the DMC task with the cat/dog stimuli at 6 image-plane rotated orientations (see Fig. 1) in addition to the trained orientation (the same set of images used for recordings described above).

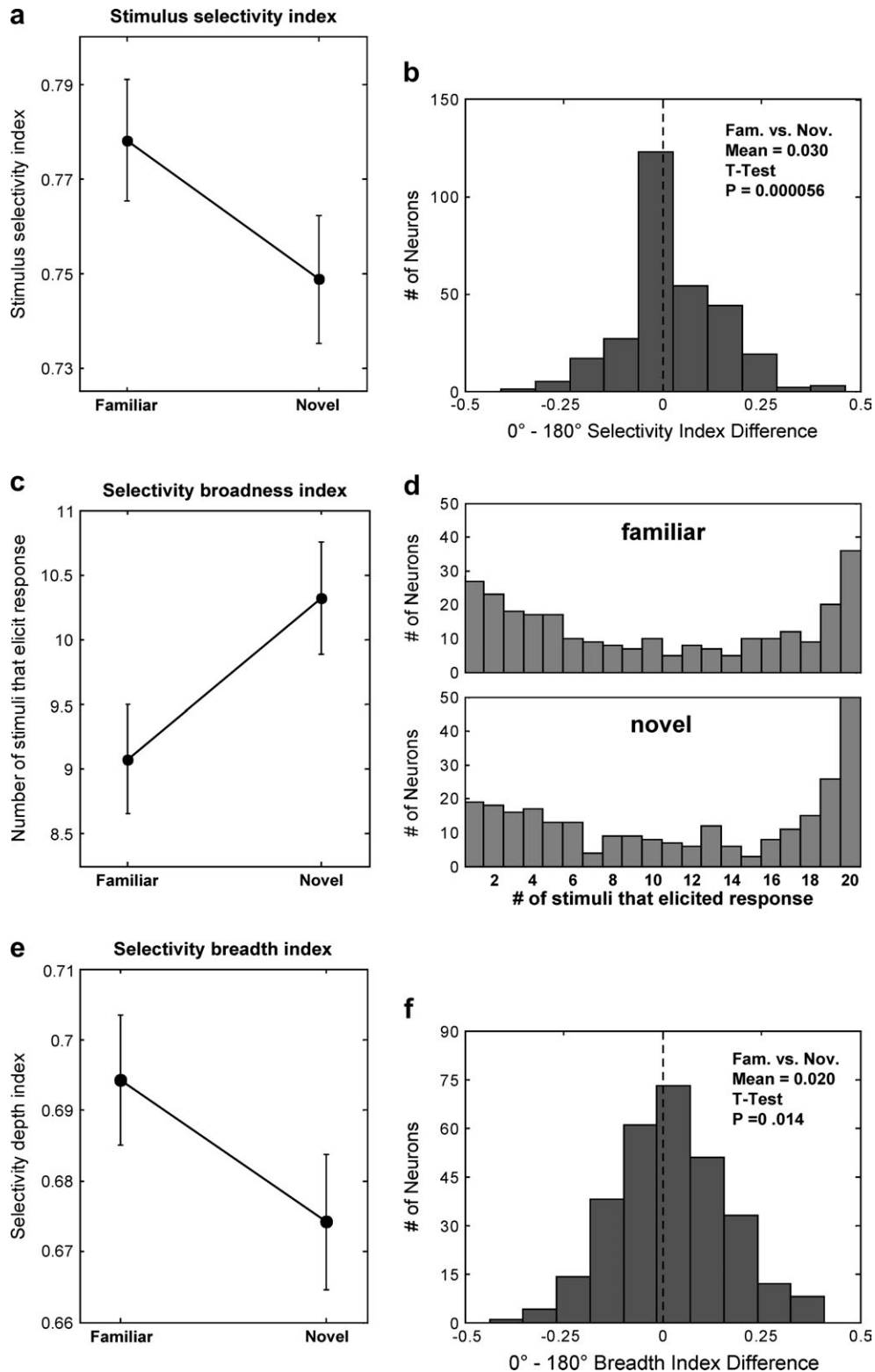


Figure 9. Analysis of stimulus selectivity, selectivity broadness, and selectivity breadth for familiar and novel (non-cat/dog) stimuli across the entire population of 298 ITC neurons. (a) Average values of the stimulus selectivity index for novel and familiar stimuli. Higher values of the index indicate greater neuronal modulation in response to the best and worst stimulus. Lower values indicate the opposite. (b) The distribution of the difference between neurons' familiar and novel stimulus selectivity index values. The positive shift of the distribution indicates greater index values for familiar stimuli. (c) Average values of the selectivity broadness index for novel and familiar stimuli. Lower values indicate sharper tuning (i.e., responses to a small subset of the 20 stimuli), whereas higher values indicate broader tuning (i.e., responses to most or all stimuli). (d) The number of stimuli (1–20) that elicited a significant response above baseline for familiar (top panel) and novel (bottom panel) stimuli. For a given number of response-eliciting stimuli (on the horizontal axis), the vertical axis shows the number of neurons that responded to that number of stimuli. (e) Average values of selectivity breadth index for familiar and novel stimuli. Higher values of the index indicate sharper selectivity. Lower values indicate the opposite. (f) The distribution of the difference between neurons' familiar and novel selectivity breadth index values. The positive shift of the distribution indicates greater index values for familiar stimuli.

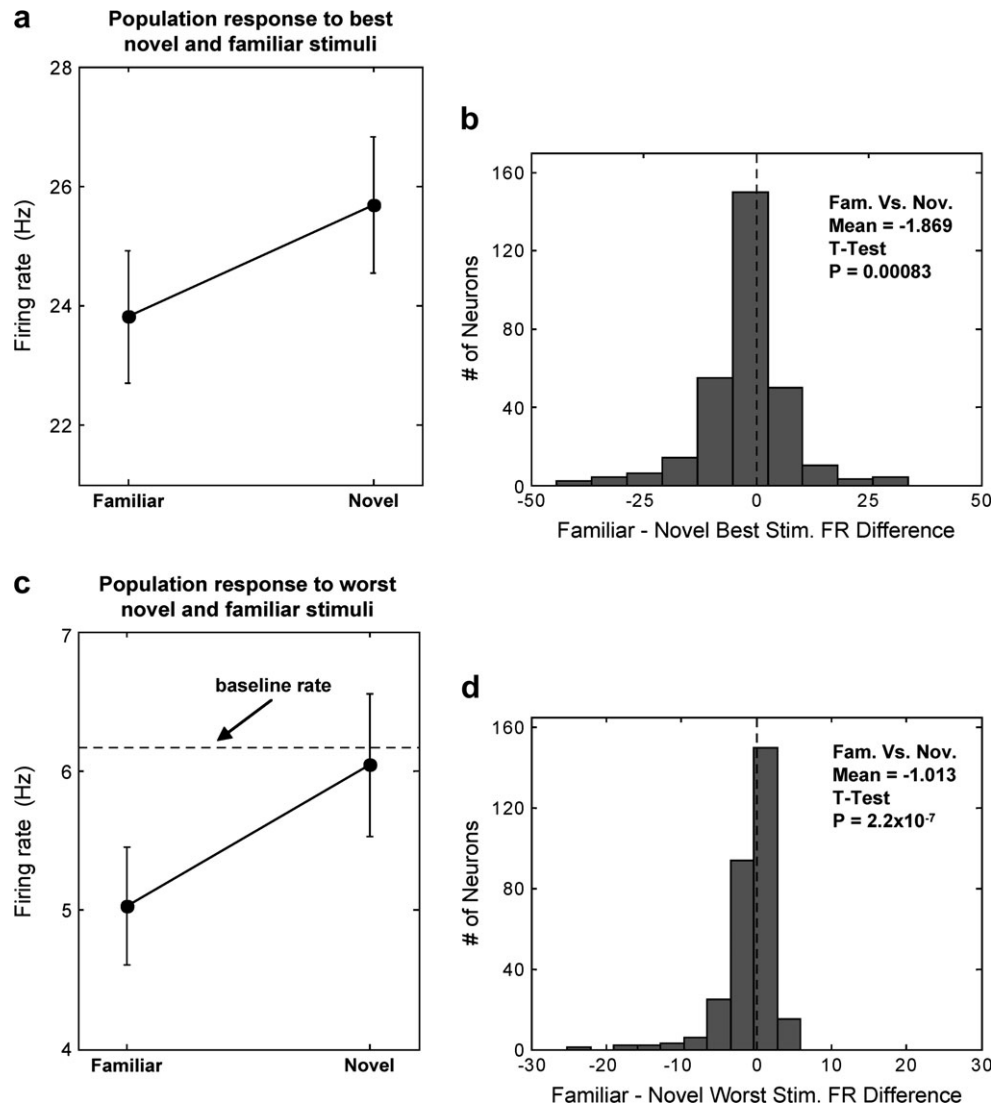


Figure 10. Average responses to each neuron’s single best and worst stimulus across the entire population of neurons. (a) Average response to each neuron’s single best stimulus (computed separately for novel and familiar stimuli). (b) The distribution of the difference between neurons’ responses to the best familiar and novel stimuli. The negative shift of the distribution indicates weaker responses to the best familiar than novel stimuli. (c) Average response to each neuron’s single worst stimulus. The prestimulus baseline activity level is indicated by the horizontal dotted line. (d) The distribution of the difference between neurons’ responses to the worst familiar and novel stimuli. The negative shift of the distribution indicates weaker responses to the worst familiar than novel stimuli.

To simplify the task design, the sample and test stimuli were always at the same orientation (see Experimental Procedures). These behavioral tests, in which monkeys received feedback about correct and incorrect behavioral responses, were conducted after all neurophysiological recordings were completed in order to avoid any neuronal effects related to training with (and rewarding behavioral responses to) judgments about the rotated stimuli. Before this, the monkeys had only been trained to categorize and match the stimuli at 1 orientation (the 0° trained orientation). Thus, their only experience with the rotated stimuli before behavioral testing on the DMC (categorization) task was passive viewing during neurophysiological recordings.

As shown in Figure 11a, DMC task performance was best (95% correct) for the trained orientation (0°) and progressively worsened as the stimuli were rotated away from it. The monkeys’ average behavioral performance for the 7 orientations correlated strongly with the average values of the stimulus

selectivity index from Figure 5a (correlation coefficient = 0.933, $P = 0.002$).

Likewise, monkeys were slower to respond to the rotated stimuli. On “category match” trials, the monkeys had to release a lever when a test stimulus was the same category as the previously seen sample stimulus. As shown in Figure 11b, monkeys’ reaction times (time between the test stimulus presentation and lever release) on correct trials were fastest at the trained orientation (0°: 300 ms) and were progressively slower for more rotated orientations (180°: 374 ms). The monkeys’ average reaction times across the 7 orientations showed a strong negative correlation with average values of the stimulus selectivity index from Figure 5a (correlation coefficient = -0.954, $P = 0.0009$), indicating that faster reaction times correspond with stronger stimulus selectivity.

Each monkey was run on several behavioral testing sessions (of approximately 500-1000 correct trials each) with the rotated stimuli across several days. Not surprisingly, their level

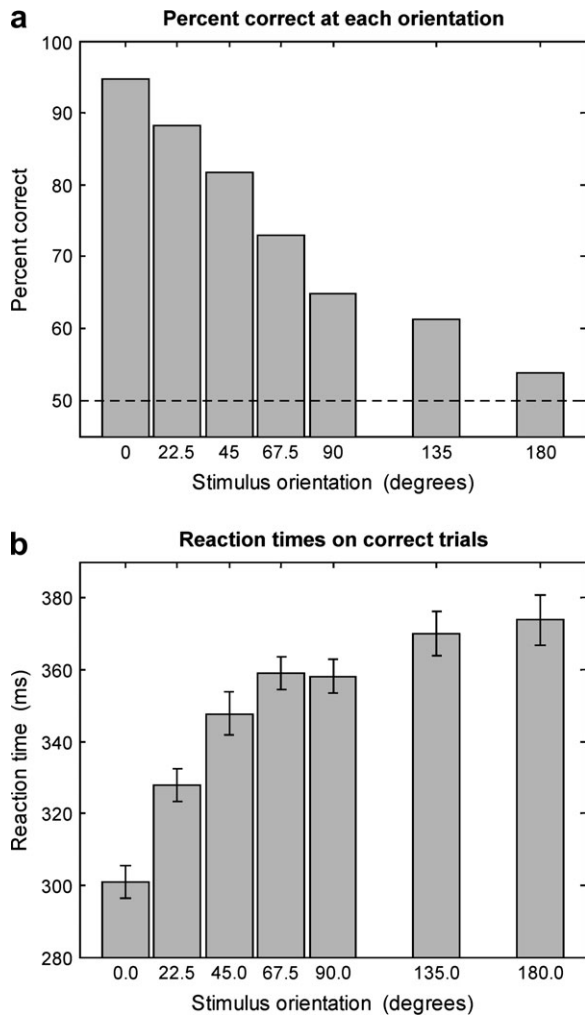


Figure 11. Behavioral performance during the cat/dog categorization task with trained (0°) and rotated stimuli. During behavioral testing, sample and test stimuli were always presented at the same orientation as one another. (a) Average categorization performance (percent correct) during DMC task with stimuli at the trained (0°) and the 6 untrained rotated orientations. (b) Average reaction time for lever release on correct “match” trials (those that required a lever release to the first test stimulus) at the trained (0°) and the 6 untrained rotated orientations.

of task performance with the rotated stimuli improved gradually across behavioral testing sessions because they received feedback about correct and incorrect classification of the rotated stimuli. Still, after several testing sessions, monkeys’ categorization performance was best near the trained orientation and became progressively worse for more rotated stimuli. The behavioral data presented in Figure 11 are from the first behavioral testing session when the monkeys had had the least amount of experience performing the DMC task with the rotated stimuli.

Discussion

We studied the activity of ITC neurons from 2 monkeys previously trained to perform a “cat and dog” categorization task. The monkeys had extensive experience with the stimuli always appearing at a fixed orientation during training. To explore whether training resulted in improved behavioral performance during the categorization task and enhanced neuronal selectivity for trained stimuli, we parametrically varied

the (two-dimensional) image-plane orientation of stimuli. Monkeys’ categorization performance became progressively worse as stimuli were rotated away from the trained orientation. We also recorded ITC neural activity while 2 monkeys passively viewed these stimuli and found strong selectivity for the trained stimuli that became progressively weaker as stimuli were rotated away from the trained orientation. Across the entire population of neurons, improved selectivity at the trained orientation seemed to be due to a sharpening of responses without increased average response rates to preferred stimuli. Interestingly, these effects were evident primarily during the early phasic burst of the neural response (80–180 ms following stimulus onset) and not during the later time epochs. Sharpening of neuronal tuning was also observed with a set of randomly selected familiar stimuli that the monkeys had seen many times during passive viewing (and were not explicitly trained). Stimulus tuning was sharper for familiar than novel stimuli though activity was, on average, weaker for familiar than novel stimuli. Together, these results suggest that both explicit training (on the cats and dogs) and passive experience (for the familiar stimuli) can result in a sharpening of neuronal shape tuning in ITC without increases in average neuronal firing rate.

Previous studies have also suggested that ITC stimulus representations are enhanced by experience. For example, Kobatake and others (1998) showed that, in anesthetized monkeys that had been trained to discriminate a set of geometric shapes, ITC neurons were more likely to respond to trained stimuli. Baker and others (2002) found sharpened selectivity for conjunctions of visual features that were experienced together during training. Sigala and Logothetis (2002) found an enhanced representation of shape features that were relevant for categorizing sets of familiar stimuli. Logothetis and others (1995) trained monkeys to discriminate between members of a set of 3-dimensional wireframe and “amoeba” objects and found many ITC neurons that responded to trained views of the familiar objects. Our findings are compatible with the results from these experiments. Although we found that a subset of ITC neurons responded most strongly to trained stimuli presented at or near the trained orientation (like the single neuron example in Fig. 4c), this subpopulation of neurons was outnumbered by neurons that responded more strongly, on average, to stimuli that were rotated away from the trained orientation. Further analysis of stimulus selectivity for the rotated cat/dog stimuli during this earliest phase of the neuronal response revealed that weaker average responses at the trained orientation were largely due to a weakening of responses to nonpreferred (worst) stimuli. This effectively resulted in sharpened stimulus tuning that could, in principle, produce an improved signal/noise (best/worst) ratio by expanding the range of spike rates evoked by trained stimuli.

For the other (non-cat/dog) stimuli, we also observed sharpened selectivity for familiar versus novel stimuli as well as weaker average responses for familiar stimuli. This suggests that experience, whether a little (passive viewing for several weeks) or a lot (months of explicit training), can result in sharpened ITC stimulus selectivity. In contrast with the rotated cat/dog results, familiarity resulted in a more general weakening of responses, not restricted only to neurons’ nonpreferred stimuli. Responses to familiar stimuli were weaker than novel stimuli across the entire range of preferred to nonpreferred stimuli. Because of this, it is unclear whether improvements in neuronal stimulus tuning for explicitly trained versus passively

experienced share the same underlying mechanisms. The different amount of experience with stimuli in these 2 experiments, and differences in the stimuli themselves, makes direct quantitative comparisons between them difficult. However, the observation that activity to the worst stimuli is suppressed below fixation baseline levels for both trained (cat/dog) and familiar stimuli does suggest that common processes may be involved. One possible explanation for the discrepancy between these effects and those with the cat/dog stimuli is that we may have observed an experience-dependent sharpening of stimulus tuning (as with the cats/dogs) in parallel with a more general suppression of activity with familiarity.

The effects described here were evident at the population level, across all recorded ITC neurons (that were preselected for general visual responsiveness, as described in Experimental Procedures). Because of this, it seems clear that visual learning and familiarity have an influence on a substantial fraction of ITC neurons. How this relates to the results of previous studies that have found clustering of ITC neurons with similar shape selectivity (Desimone and others 1984; Tanaka 1996, 2003; Erickson and others 2000) remains to be seen. One possibility is that ITC neurons are multitasking, taking on additional selectivity. Another possibility is that the tuning properties of more posterior areas in ITC are less influenced by experience than the anterior regions we examined. Experiments that compare the properties of shape selective neurons across multiple areas in ITC are needed to better understand their respective contributions.

It is also not known how these ITC shape representations are “read out” by downstream brain areas. One possibility is that object vision and behavioral improvements with learning depend critically on the population of neurons that show weakened responses with familiarity and sharpened tuning for trained stimuli. Another possibility is that the smaller population of neurons that respond more strongly for trained and/or familiar stimuli is more directly involved in visual recognition. In addition, the degree to which ITC neurons reflect monkeys’ trial to trial variability in behavioral performance during categorization or object recognition is unclear because all neuronal data in this study were collected during passive viewing of stimuli. More experiments, particularly those that allow for direct monitoring of ITC responses during learning and more complex behavioral tasks, are needed to resolve these issues.

The cat and dog stimuli in this study can be considered natural or biological because they resemble, to some extent, real animals. In addition, the trained orientation that we used in this study corresponds to their canonical orientation (the orientation at which real cats and dogs are typically observed). Because of this, it is important to consider whether this could explain the enhanced tuning for these stimuli at the trained orientation that we have observed in ITC. We feel that this is an unlikely explanation for these effects for a number of reasons. First, we also found a sharpening of selectivity for familiar compared with novel randomly chosen stimuli (that were not biological stimuli with canonical orientations). Second, because these monkeys were raised in captivity in a laboratory environment, they had very limited (if any) experience with other species of animals. Third, during several sessions of behavioral testing (to assess the monkeys’ abilities to categorize the cat/dog stimuli at untrained orientations), we found that the monkeys’ task performance initially (during the first testing session) was excellent (>90% correct) for the trained stimuli and fell toward chance (50%) as

the stimuli were rotated away from the trained orientation. Over the course of several subsequent testing sessions, task performance gradually improved for the rotated stimuli (though still remained worse than at the trained orientation). Through additional training with the rotated stimuli, it seemed very likely that the monkeys could have performed equally well at the previously trained (0°) and rotated, noncanonical, orientations. This suggests that the monkeys’ enhanced behavioral performance and, presumably, sharpened neuronal tuning at the trained orientation were due to training and not the biological nature of the stimuli.

These results fit well with other reports of decreases in average ITC activity with both short- and long-term familiarity (Fahy and others 1993; Li and others 1993; Baker and others 2002). These results are also compatible with those obtained from a study of the effects of experience on the recognition of noise-degraded stimuli (Rainer and Miller 2000). Neurons in the lateral PFC (which is directly interconnected with the ITC) also decreased their activity and sharpened selectivity in concert with improved ability to recognize the noise-degraded stimuli with experience. These results have suggested that visual familiarity simultaneously weakens and sharpens neural activity leaving behind a sparser, sharper, more energy-efficient neural representation. This may help foster the visual expertise that comes with experience. However, by contrast, neurons in area V4, which provides the ITC with visual input, show selective “increases” in activity to familiar noise-degraded stimuli (Rainer and others 2004). This signal, combined with a general weakening of activity with familiarity in the ITC (and PFC), could, in principle, account for our observations of sharpened selectivity by disproportionate weakening of responses to relatively ineffective stimuli.

In the studies mentioned above (Logothetis and others 1995; Kobatake and others 1998), neuronal responses to stimuli were averaged over many hundreds of milliseconds for the analysis of stimulus selectivity. In this study, we found that sharpening of stimulus selectivity was most pronounced in the earliest phase of the neuronal response (80–180 ms following stimulus onset). It is this initial phase that seems critical for visual recognition (Thorpe and others 1996; Sugase and others 1999; Delorme and Thorpe 2001). Given that the monkeys’ average manual reaction times (for a lever release) were approximately 300 ms when performing the categorization task with these stimuli, any information encoded by ITC beyond this early phase of the response is unlikely to be useful for guiding immediate behavioral decisions based on that stimulus. Further, it is unlikely that these effects could be due to top-down attentional modulations because the strongest learning effects were evident in the initial wave of ITC activity, which is assumed to be driven primarily via feed-forward inputs from upstream visual areas. Information encoded in the later phases of the ITC activity is likely to be of critical importance for more cognitive functions, such as short-term memory, long-term memory retrieval, memory consolidation, and spatial or feature-based attention. Clearly, further experiments are needed to understand the interaction between perceptual and cognitive signals in ITC and their role in visual learning and visually guided behavior.

Notes

We thank C. Shelton for the morphing software and T. Buschman, C. Freedman, M. Histed, N. Kanwisher, M. Machon, A. Pasupathy, J. Roy,

A. Seitz, J. Wallis, and M. Wicherski for valuable comments, help, and discussions. This work was supported by an National Institute of Mental Health grant MH22588, an National Science Foundation-Knowledge and Distributed Intelligence grant, the RIKEN-MIT Neuroscience Research Center, a McDonnell Pew grant in cognitive neuroscience (MR). Additional support was provided by the Eugene McDermott Foundation and the Whitaker Foundation (TP). DJF is supported by an individual postdoctoral National Research Service Award from the National Eye Institute.

Send reprint requests to david_freedman@hms.harvard.edu or ekmiller@mit.edu.

Address correspondence to David J. Freedman, 220 Longwood Avenue WAB 222, Harvard Medical School, Boston, MA 02115, USA. Email: david_freedman@hms.harvard.edu.

References

- Baker CI, Behrmann M, Olson CR. 2002. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat Neurosci* 5:1210-1215.
- Blum JS, Chow KL, Pribram KH. 1950. A behavioral analysis of the organization of the parieto-temporo-preoccipital cortex. *J Comp Neurol* 93:53-100.
- Booth MC, Rolls ET. 1998. View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb Cortex* 8:510-523.
- Brincat SL, Connor CE. 2004. Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat Neurosci* 7:880-886.
- Bruce C, Desimone R, Gross CG. 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus in the macaque. *J Neurophysiol* 46: 369-384.
- Damasio AR, Damasio H, Van Hoesen GW. 1982. Prosopagnosia: Anatomic basis and behavioral mechanisms. *Neurology* 32: 331-341.
- Delorme A, Thorpe SJ. 2001. Face identification using one spike per neuron: resistance to image degradations. *Neural Netw* 14:795-803.
- Desimone R, Albright TD, Gross CG, Bruce C. 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051-2062.
- Erickson CA, Jagadeesh B, Desimone R. 2000. Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. *Nat Neurosci* 3:1066-1068.
- Fahy FL, Riches IP, Brown MW. 1993. Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Exp Brain Res* 96:457-472.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2001. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291:312-316.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2002. Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. *J Neurophysiol* 88:914-928.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2003. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J Neurosci* 23:5235-5246.
- Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. 1999. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci* 2:568-573.
- Gold J, Bennett PJ, Sekuler AB. 1999. Signal but not noise changes with perceptual learning. *Nature* 402:176-178.
- Gross CG. 1973. Visual functions of inferotemporal cortex. In: Autrum H, Jung R, Lowenstein W, McKay D, Teuber H-L, editors. *Handbook of sensory physiology*. Volume VII/3B. Berlin: Springer-Verlag.
- Jagadeesh B, Chelazzi L, Mishkin M, Desimone R. 2001. Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. *J Neurophysiol* 86:290-303.
- Kluver H, Bucy L. 1938. An analysis of certain effects of bilateral temporal lobectomy in the rhesus monkey, with special reference to "psychic blindness." *J Psychol* 5:33-54.
- Kluver H, Bucy L. 1939. Preliminary analysis of functions of the temporal lobes in monkeys. *Arch Neurol Psychiatry* 42:979-1000.
- Kobatake E, Wang G, Tanaka K. 1998. Effects of shape discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J Neurophysiol* 80:324-330.
- Leopold DA, Logothetis NK. 1996. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379:549-553.
- Li L, Miller EK, Desimone R. 1993. The representation of stimulus familiarity in anterior inferior temporal cortex. *J Neurophysiol* 69:1918-1929.
- Logothetis NK, Pauls J, Poggio T. 1995. Shape representation in the inferior temporal cortex of monkeys. *Curr Biol* 5:552-563.
- Logothetis NK, Sheinberg. 1996. Visual object recognition. *Annu Rev Neurosci* 19:577-621.
- Miller EK, Li L, Desimone R. 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* 254:1377-1379.
- Miller EK, Li L, Desimone R. 1993. Activity of neurons in anterior inferior temporal cortex during a short term memory task. *J Neurosci* 13:1460-1478.
- Mishkin M. 1954. Visual discrimination performance following partial ablations of the temporal lobe. II. Ventral surface vs. hippocampus. *J Comp Physiol Psychol* 47:187-193.
- Mishkin M. 1966. Visual mechanisms beyond the striate cortex. In: Russell RW, editor. *Frontiers in physiological psychology*. New York: Academic Press p 93-119.
- Mishkin M, Pribram K. 1954. Visual discrimination performance following partial ablations of the temporal lobe. I. Ventral vs. lateral. *J Comp Physiol Psychol* 47:14-20.
- Miyashita Y. 1993. Inferior temporal cortex: where visual perception meets memory. *Annu Rev Neurosci* 16:245-263.
- Op de Beeck H, Wagemans J, Vogels R. 2001. Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat Neurosci* 4:1244-1252.
- Paxinos G, Huang XF, Toga AW. 2000. *The rhesus monkey brain in stereotaxic coordinates*. San Diego, CA: Academic Press.
- Perrett DI, Rolls ET, Caan W. 1982. Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47:329-342.
- Rainer G, Lee H, Logothetis NK. 2004. The effect of learning on the function of monkey extrastriate visual cortex. *PLoS Biol* 2:E44.
- Rainer G, Miller EK. 2000. Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27:179-189.
- 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.
- Shelton C. 2000. Morphable surface models. *Int J Comput Vis* 38:75-91.
- Sigala N, Logothetis NK. 2002. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415:318-320.
- Sowden PT, Davies IR, Roling P. 2000. Perceptual learning of the detection of features in X-ray images: a functional role for improvements in adults' visual sensitivity? *J Exp Psychol Hum Percept Perform* 26:379-390.
- Sugase Y, Yamane S, Ueno S, Kawano K. 1999. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400:869-873.
- Tanaka K. 1996. Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19:109-139.
- Tanaka K. 2003. Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cereb Cortex* 13:90-99.
- Thorpe S, Fize D, Marlot C. 1996. Speed of processing in the human visual system. *Nature* 381:520-522.
- Vogels R. 1999. Categorization of complex visual images by rhesus monkeys. *Eur J Neurosci* 11:1223-1238.