

RESEARCH ARTICLE | *Sensory Processing*

Stimulus context alters neural representations of faces in inferotemporal cortex

Behrad Noudoost,¹ Neda Nategh,² Kelsey Clark,¹ and Hossein Esteky^{3,4}

¹Department of Cell Biology and Neuroscience, Montana State University, Bozeman, Montana; ²Department of Electrical and Computer Engineering, Montana State University, Bozeman, Montana; ³School of Cognitive Sciences, Institute for Research in Fundamental Sciences, Tehran, Iran; and ⁴Research Center for Brain and Cognitive Sciences, Shaheed Beheshti University of Medical Sciences, Tehran, Iran

Submitted 22 August 2016; accepted in final form 24 October 2016

Noudoost B, Nategh N, Clark K, Esteky H. Stimulus context alters neural representations of faces in inferotemporal cortex. *J Neurophysiol* 117: 336–347, 2017. First published October 26, 2016; doi:10.1152/jn.00667.2016.—One goal of our nervous system is to form predictions about the world around us to facilitate our responses to upcoming events. One basis for such predictions could be the recently encountered visual stimuli, or the recent statistics of the visual environment. We examined the effect of recently experienced stimulus statistics on the visual representation of face stimuli by recording the responses of face-responsive neurons in the final stage of visual object recognition, the inferotemporal (IT) cortex, during blocks in which the probability of seeing a particular face was either 100% or only 12%. During the block with only face images, ~30% of IT neurons exhibit enhanced anticipatory activity before the evoked visual response. This anticipatory modulation is followed by greater activity, broader view tuning, more distributed processing, and more reliable responses of IT neurons to the face stimuli. These changes in the visual response were sufficient to improve the ability of IT neurons to represent a variable property of the predictable face images (viewing angle), as measured by the performance of a simple linear classifier. These results demonstrate that the recent statistics of the visual environment can facilitate processing of stimulus information in the population neuronal representation.

NEW & NOTEWORTHY Neurons in inferotemporal (IT) cortex anticipate the arrival of a predictable stimulus, and visual responses to an expected stimulus are more distributed throughout the population of IT neurons, providing an enhanced representation of second-order stimulus information (in this case, viewing angle). The findings reveal a potential neural basis for the behavioral benefits of contextual expectation.

object recognition; face perception; view representation; contextual effect

AS WE INTERACT WITH THE WORLD around us, we form predictions or expectations regarding stimuli we are likely to encounter (Friston and Kiebel 2009; Kersten et al. 2004; Seriès and Seitz 2013). Some of these expectations are formed over a lifetime of visual experience, but others are shaped by the immediate context or the recent statistics of the sensory environment (Berniker et al. 2010; Chalk et al. 2010; Sterzer et al. 2008). Accurate anticipation of stimuli facilitates behavioral responses (Esterman and Yantis 2010; Faulkner et al. 2002; Puri et al. 2009; Puri and Wojciulik 2008). Alongside these behav-

ioral effects, changes in visual processing have been reported at a neural level based on short-term or long-term visual experience and expectations. For example, familiar visual stimuli have sparser inferotemporal (IT) representations compared with novel stimuli (Meyer et al. 2014; Woloszyn and Sheinberg 2012). Accurate temporal cueing produces stronger visual responses in IT (Anderson and Sheinberg 2008); IT visual responses are likewise greater for behaviorally relevant stimuli relative to cue stimuli (Anderson and Sheinberg 2008). Over short timescales, image repetition generally reduces the strength of visual responses (de Gardelle et al. 2013; McMahon and Olson 2007). Expectation of a specific image, based on prolonged learning of a paired image sequence, also reduces the IT response to the expected stimulus (Meyer and Olson 2011). The findings most closely related to our own come not from neurophysiology, but from imaging studies, which have revealed that prior knowledge of the general class or category of objects to which an upcoming stimulus belongs changes both the anticipatory activity and stimulus-evoked responses in object-selective visual cortex (Egner et al. 2010; Esterman and Yantis 2010; Puri et al. 2009). Whether the previously reported influences of context, statistical predictability, or cue-based anticipation on visual processing ultimately arise from the same modulatory mechanisms is unknown. Here, we study one specific instance of this class of effects: changes in visual responses based on the predictability of recent stimulus statistics.

The goal of the present study is to determine how recent stimulus statistics change the responses of single neurons, and how these changes enhance information processing at the single-neuron and population levels. We recorded the responses of face-responsive neurons in macaque IT cortex, the final stage of visual object processing (Desimone et al. 1984; Hung et al. 2005; Logothetis and Sheinberg 1996). Neuronal responses to a stream of visual stimuli were recorded under conditions in which the probability of a specific face appearing was either 100% (face block) or ~12% (mixed block). During the face block, in which the stimuli are all different views of the same face, face-responsive IT neurons exhibit increased activity prior to the visual response, consistent with previous studies in the same area (Chelazzi et al. 1998). Repeated presentation of face stimuli also resulted in a more distributed and a more reliable representation of viewing angle information in IT neurons. These changes are accompanied by an increase in the ability of IT neurons to represent information

Address for reprint requests and other correspondence: B. Noudoost, Dept. of Cell Biology and Neuroscience, Montana State Univ., Bozeman, MT 59717 (e-mail: bnoudoost@montana.edu).

about the variable aspect of the stimulus (in this case, face viewing angle). These findings indicate that recent stimulus statistics alter the tuning and reliability of IT face responses, providing insight into the neural correlates of contextual effects on face processing, and provide a starting point for understanding how context and sensory adaptation affect coding in high-level sensory areas such as IT.

MATERIALS AND METHODS

Setup and Procedures

Two male monkeys (M1 and M2), *Macaca mulatta*, were each implanted with a recording chamber and head post. All surgical and experimental procedures were in accordance with Harvard Medical School and US National Institutes of Health (NIH) guidelines. The experimental procedures were approved by the Committee for Care and Use of Experimental Animals, Iranian Society for Physiology and Pharmacology (CCUEA-ISPP). Single-neuron electrophysiological recordings were made using tungsten microelectrodes (FHC) and a guide-tube/grid system. Monkeys passively fixated during foveal presentation of 20 face views (face condition) or 172 object images, including the 20 face views (mixed condition). Face view stimuli were 5° gray-scale images generated by Poser software (Smith Micro Software). One block (7–20 repetitions of each image; 4 images per trial) of each condition was recorded, and neurons were then used for another set of experiments (Noudoost and Esteky 2013); the order of the two conditions was randomized across experiments. Monkeys fixated on a 0.5° circular fixation spot presented at the center of the display in a 2 × 2° window. Rewards were delivered every 2 s, and stimulus presentation continued until fixation was broken. The eye position was monitored by an infrared video monitoring system (i_rec; <http://staff.aist.go.jp/k.matsuda/eye/>).

Electrophysiological Recording

Headposts and recording chambers were implanted on the skull of two adult male macaque monkeys (*Macaca mulatta*). All experimental procedures conformed to the guidelines on the care and use of laboratory animals of the NIH. The electrode was advanced with an Evert type manipulator (Narishige) from the dorsal surface of the brain through a stainless steel guide tube inserted into the brain down to 10 mm above the recording sites. The recording positions were determined stereotaxically, referring to the magnetic resonance images acquired before the surgery, and the gray and white matter transitions were determined during electrode advancement. The action potentials from a single neuron were isolated in real time by a template-matching algorithm. Further technical details about monkey surgery and recording procedure can be found elsewhere (Kiani et al. 2005).

Recording areas. Extracellular single-cell recording positions were made over anterior 13–21 mm from ear line over the lower bank of the superior temporal sulcus and the ventral convexity up to the medial bank of the anterior middle temporal sulcus, with 1-mm track intervals in both monkeys.

Experimental Conditions

One block (7–20 repetitions of each image; 4 images per trial) of each condition was recorded; the block order was assigned randomly each day.

Face condition. The stimulus set consisted of 20 views of an artificial face ranging from 0° (left profile) to 360° in 18° increments. Stimuli were 5 × 5° and were presented at the center of the monitor while the monkey was holding its fixation in a 2 × 2° window at the center of the image. Each stimulus was presented for 250 ms with a 250-ms interstimulus interval to allow enough time for a neuron to return to its baseline activity before the next stimulus was presented; four images were presented per

trial. Stimuli were presented in a pseudorandom order, with 7–20 repetitions of each stimulus within a block.

Mixed condition. Stimulus characteristics and presentation were as described for the face condition, except that the 20 face views were randomly interleaved with 152 other 5 × 5° object images (100 nonface objects and 52 animal and human faces).

Of 153 neurons recorded in these conditions (M1, $n = 89$; M2, $n = 64$), 131 neurons (M1, $n = 76$; M2, $n = 55$), which showed significant responses to at least one of the face view images, are included in the analysis. The significance of responses to individual stimuli was determined by comparing the baseline firing rate with the neuron's firing rate 70–320 ms after stimulus onset (this window was selected based on the typical latency of IT visual responses; $P < 0.05$, Kolmogorov-Smirnov test) pooled across randomized presentations of individual stimuli in the mixed-condition blocks. The baseline-firing rate was measured in the 200-ms period immediately before the onset of sequential stimulus presentation. The same method was used to calculate the proportion of the 131 neurons responding to a particular face view in Fig. 3A, inset. In investigating the effects of context on neural tuning for viewing angle, we focused on classic view-tuned neurons. Of the 131 face-responsive neurons, 82 were identified as having robust tuning for viewing angle (M1, $n = 48$; M2, $n = 34$); 14 neurons with mirror symmetric view tuning and 17 neurons with view-invariant tuning were excluded from tuning analysis. Tuning for viewing angle was quantified as the average vector in a normalized polar plot of the neuron's visual response to different viewing angles; neurons with an average vector of >0.1 (where 1 would represent a neuron which responded only to a single face view) were considered tuned for viewing angle.

Statistical Analysis

Mean-normalized response. Normalization was performed for each neuron by dividing the firing rate by the overall mean firing rate of that neuron (grand mean over time, all visual stimuli, both the face and mixed conditions). Throughout the text, the Wilcoxon signed-rank test is used for hypothesis tests, unless another test is mentioned. Data analysis and statistical tests were performed using MATLAB (MathWorks). Stimuli were present onscreen for 250 ms with a 250-ms interstimulus interval (allowing time for the visual response to decay between stimuli); stimulus-evoked responses were measured in a 250-ms window, starting 70 ms after stimulus onset, based on the previously reported visual latencies of IT neurons.

Breadth of tuning. We quantified the breadth of tuning of an individual neuron to a set of stimuli based on the methods of Rolls and Tovee (1995), which has also been used to study the sparseness of representations in a variety of brain areas (Rust and DiCarlo 2012; Vinje and Gallant 2002), using the following index:

$$\frac{\left[\sum_{s=1}^S (y_s/S) \right]^2}{\sum_{s=1}^S (y_s)^2/S}$$

where y_s is the mean firing rate of the neuron to stimulus s in the set of S stimuli (20 face views). We quantified the breadth of population tuning for each stimulus with the following index:

$$\frac{\left[\sum_{n=1}^N (y_n/N) \right]^2}{\sum_{n=1}^N (y_n)^2/N}$$

where y_n is the mean firing rate of neuron n to a particular stimulus in the set of N neurons. The neuronal and population breadth of tuning provide a quantitative measure of stimulus representation sparseness for single neurons or across the neuronal population, respectively.

To control for overall changes in firing rate, we also calculated versions of both breadth of tuning and breadth of population tuning in

which the above formulas were applied to data in which the difference between the average face and mixed firing rate was added to the mixed data.

Depth of selectivity. We also quantified neural tuning using the depth of selectivity (DOS) measure (Moody et al. 1998; Mruczek and Sheinberg 2012; Rainer and Miller 2000).

$$\text{DOS} = \frac{1}{S-1} \left[S - \left(\frac{\sum_{s=1}^S y_s}{y_{\max}} \right) \right]$$

where S is the number of stimuli, y_s is the mean firing rate of the neuron in response to stimulus s , and y_{\max} is the largest firing rate across all stimuli. DOS will range from 0 to 1, with larger values corresponding to more selective tuning, and lower values corresponding to broader tuning.

Selectivity index. We also measured stimulus tuning using the selectivity index (SI).

$$\text{SI} = \frac{y_{\max} - y_{\min}}{y_{\max} + y_{\min}}$$

where y_{\max} and y_{\min} are the mean firing rates of the neuron in response to the strongest and weakest stimulus. SI will range from 0 to 1, with larger values corresponding to more selective (less broad) tuning.

Response variability. Fano factor was computed as the variance/mean of neuronal responses to a single stimulus across multiple trials, using the mean matching methods developed by Churchland and colleagues (2010). In outline, this method involves looking at the firing rate for each neuron and time bin and discarding points until a common firing rate distribution between conditions is achieved.

Support vector machine classifier. We used a support vector machine (SVM) algorithm to classify the presented face views based on the responses of IT neurons. All of the results shown here are based on an SVM with a linear kernel; an SVM classifier with a Gaussian kernel produced similar results. For the multiclass classification across face views, we measured SVM performance for each view in a one-vs.-the-rest multiclass system. To measure classifier performance over time, we used spike counts in a sliding time bin at different latencies from stimulus onset on the test (unseen) data (see Fig. 6A). To measure classifier performance as a function of the number of neurons used (see Fig. 6B), we used a random permutation. For each group, we take a random subset of neurons (to have an equal number of neurons in each group), then for each set of selected neurons we randomly take 20% of trials out and train the classifier with the

remaining 80% and then test classifier performance with the unseen 20% of trials; the number of face stimuli responses used to train and test responses in the face and mixed blocks are the same. The SE in Fig. 6B is defined by running this algorithm 50 times.

Multidimensional scaling (MDS). We used MATLAB `cmdscale` function for multidimensional scaling. Using the normalized responses of the population of 131 IT neurons, a dissimilarity matrix was calculated. This matrix quantifies the degree to which the population response to a stimulus differs from the other 19 stimuli. A multidimensional space was generated based on this matrix. Figure 7A depicts the first two dimensions of this space (the two that best reflect the dissimilarity); the distance analyses in Fig. 7, B and C, are calculated based on Euclidean distances in the whole space (19 dimensions, based on the number of face stimuli).

RESULTS

We recorded the responses of 153 IT neurons in two macaque monkeys to 20 different views of an artificially generated face stimulus (see MATERIALS AND METHODS; M1, $n = 89$; M2, $n = 64$). Of these neurons, 131 responded to at least one of the face view images (M1, $n = 76$; M2, $n = 55$), and these face-responsive neurons were selected for further analysis. All recordings were made while awake monkeys were passively fixating a stream of foveally presented visual stimuli. IT responses were measured under two different conditions: in the face condition, stimuli were 20 different views of a single face identity (Fig. 1A); in the mixed condition, the same set of 20 face views was presented, but pseudorandomly intermixed with a variety of other faces and objects (see MATERIALS AND METHODS). In the mixed condition, only ~12% of stimuli were the face from the face condition, and ~58% of stimuli were nonface objects. Monkeys were not trained to discriminate between the two conditions; however, we found clear evidence that during the face condition IT responses are enhanced even prior to the visual response, indicating that, even in the absence of explicit behavioral training, recent stimulus statistics can influence neural responses. Figure 1B shows the effects of the face vs. mixed stimulus context on the responses of an exemplar face view selective IT neuron, averaged across the 20 face stimuli (Fig. 1C). Activity in the face condition was greater than activity in the mixed condition, even prior to the stimulus-evoked response [0–70 ms; face condition = 7.7 ± 0.75 spikes

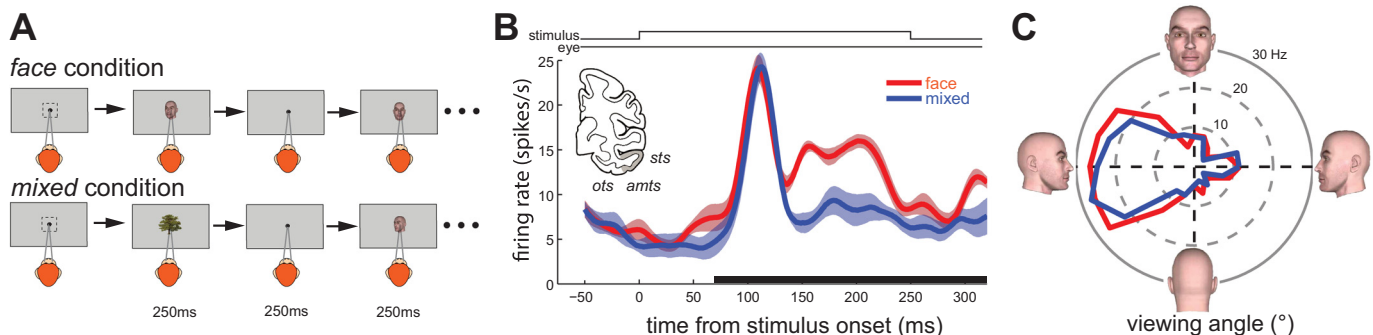


Fig. 1. Stimulus context alters visual processing in IT (example neuron). **A**: stimulus presentation paradigm. Monkeys fixated a stream of visual stimuli presented on the screen. Each stimulus was presented for 250 ms, with a 250-ms interstimulus interval. Monkeys needed to maintain fixation for 2 s to receive a reward; however, they could continue fixating to receive an additional reward every 2 s. In the face condition, 20 views of a single face were presented as stimuli; in the mixed condition stimuli were a mixture of face and nonface images. **B**: visual response of an exemplar IT neuron to stimulus presentation, averaged across all 20 face stimuli, for the face condition (red) and the mixed condition (blue). Average visual response (70–320 ms; gray bar) is enhanced in the face condition. The activity prior to stimulus-evoked response (0–70 ms) is also greater in the face condition. *Inset*: area of inferotemporal cortex from which recordings were made (sts, superior temporal sulcus; amts, anterior middle temporal sulcus; ots, occipitotemporal sulcus). **C**: polar plot showing tuning of example neuron for viewing angle in face condition (red) and mixed condition (blue).

(sp/s; mixed condition = 4.6 ± 0.56 sp/s, $P = 0.048$). We refer to this increase in firing rate before the onset of the visually evoked response as anticipatory activity; the existence of this anticipatory activity indicates that, although the animal need not monitor or report on block statistics, the stimulus statistics of the condition are reflected at the level of neural responses. This anticipatory activity was accompanied by an enhanced visual response: the stimulus-evoked response of the neuron to the face stimuli was greater in the face condition compared with the mixed condition (70–320 ms; face condition = 15.68 ± 0.84 sp/s; mixed condition = 10.92 ± 0.81 sp/s, $P = 0.0011$).

As in the exemplar neuron, the IT population showed an anticipatory enhancement in responses prior to the stimulus-evoked response in the face condition (Fig. 2A). For the population of 131 IT neurons, the average mean-normalized

response (MNR) (see MATERIALS AND METHODS) of IT neurons in the 70-ms interval prior to the stimulus-evoked response was 0.796 ± 0.0257 MNR in the mixed condition; in the face condition, this anticipatory activity significantly increased to 0.887 ± 0.0253 MNR ($P = 0.0012$; Fig. 2A). This increased anticipatory activity was significant in both monkeys [Fig. 2B; M1, change (Δ) in anticipatory response = 0.120 ± 0.0317 MNR; $P = 0.0028$; M2, Δ anticipatory response = 0.0702 ± 0.0367 MNR; $P = 0.035$]. We examined the time course of the visual response in both the face and mixed conditions. We did not find a difference between the two conditions in terms of the time from stimulus onset to the peak of the visually evoked response (time to peak, mixed condition = 117 ± 2.98 ms, face condition = 114 ± 2.78 ms, $P = 0.2370$). We found that the visually evoked activity dropped back down to the baseline level less than 400 ms after stimulus onset; there was no

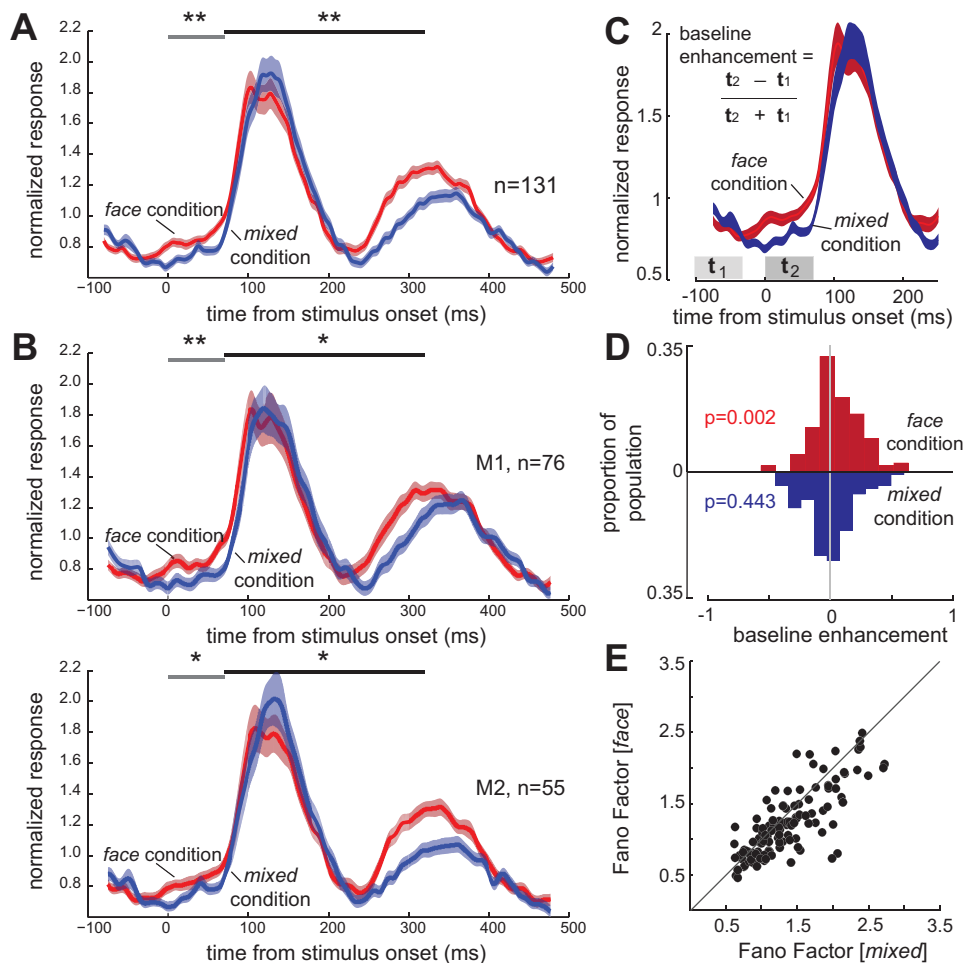


Fig. 2. Stimulus context alters population activity and reliability. *A*: mean-normalized visual response of 131 face-responsive IT neurons to stimulus presentation, averaged across all 20 face stimuli, for the face condition (red) and the mixed condition (blue). The activity prior to stimulus-evoked response (0–70 ms; gray bar) is greater in the face condition. Average visual response (70–320 ms; black bar) is enhanced in the face condition. *B*: mean-normalized visual responses of 76 face-responsive IT neurons from monkey 1 (M1; top) and 55 neurons from monkey 2 (M2; bottom), averaged across all 20 face stimuli, for the face condition (red) and the mixed condition (blue). The activity prior to stimulus-evoked response (0–70 ms; gray bar) is greater in the face condition in both monkeys. Average visual response (70–320 ms; black bar) is enhanced in the face condition in both monkeys. * $P < 0.05$. ** $P < 0.01$. *C*: baseline enhancement index is quantified as the difference between responses immediately prior to the visual response (t_2 , response during 70-ms window, starting at stimulus onset) and baseline response (t_1 , response during 70-ms window, starting 100 ms before stimulus) divided by their sum. *D*: this index was significantly greater than zero in face condition (baseline enhancement index_{face} = 0.065 ± 0.019 , $P = 0.002$), which is equal to a 13.68% increase in response. But the index was not significantly different from zero in the mixed condition (baseline enhancement_{mixed} = -0.011 ± 0.020 , $P = 0.443$), and was significantly different between the face and mixed conditions (baseline enhancement_{face-mixed} = 0.076 ± 0.017). *E*: the variability of neuronal responses decreases in the face block. The scatter plot shows the variability, measured as the Fano factor (variance/mean), of neural responses to all 20 face stimuli for the population of 131 face-responsive IT neurons. Variability was lower in the face condition (y-axis) than in the mixed condition (x-axis).

significant difference between neural activity in the two conditions 375 ms after stimulus onset ($P = 0.169$). This confirms that the enhancement observed prior to the visual response in the face condition is not due to an ongoing visual response to the previous face stimulus presentation. The average stimulus-evoked response to all face views was also significantly greater in the face condition than the mixed condition (Δ stimulus-evoked response = 0.053 ± 0.032 MNR; $P = 0.002$); this increase in stimulus-evoked response was significant in both monkeys (M1, Δ stimulus-evoked response = 0.0656 ± 0.0370 MNR; $P = 0.0166$; M2, Δ stimulus-evoked response = 0.0449 ± 0.0315 MNR; $P = 0.0241$). The magnitude of this enhancement in the stimulus-evoked response was not significantly greater than that of the anticipatory enhancement ($P = 0.558$). IT responses differed significantly between the two conditions beginning 43 ms before stimulus onset (Wilcoxon signed-rank test, $P < 0.05$). The increases in anticipatory activity and visual responses were also visible in the raw (unnormalized) population data: anticipatory activity increased from 11.6 ± 0.8 sp/s in the mixed condition to 12.7 ± 0.8 sp/s in the face condition ($P < 10^{-3}$), and the visual response increased from 16.2 ± 1.2 to 17.4 ± 1.1 sp/s ($P = 0.015$). An enhancement in anticipatory activity could also be measured by comparing activity just prior to the stimulus response to earlier in the interstimulus interval, and this measure of enhancement

was also larger for the face condition than the mixed condition (Fig. 2, *C* and *D*).

More Distributed and More Reliable Responses to Predictable Stimuli

Despite the enhanced anticipatory and visual activity, the variability of neuronal responses did not increase proportionately. We quantified the response variability in terms of the mean-matched Fano factor (defined as variance/mean of raw firing rates, for the 70- to 320-ms visually evoked response window; mean firing rate was matched across conditions, see MATERIALS AND METHODS for details). Across all face views, the mean-matched Fano factor was lower in the face condition (median change in Fano factor, face – mixed = -0.114 , $P < 10^{-3}$; Fig. 2*E*); this drop in Fano factor was consistent across monkeys (median change in Fano factor, face – mixed M1 = -0.155 ± 0.0385 , $P < 10^{-3}$; M2 = -0.0829 ± 0.0411 , $P < 10^{-3}$). Therefore, more predictable stimulus statistics resulted in a more reliable signaling of visual information.

Next we examined how recently experienced stimulus statistics affected neural tuning for face viewing angles, which varied within the face condition. We quantified IT neurons' tuning for face viewing angle using 20 face views presented in the face and mixed conditions. As shown in Fig. 3*A*, for the

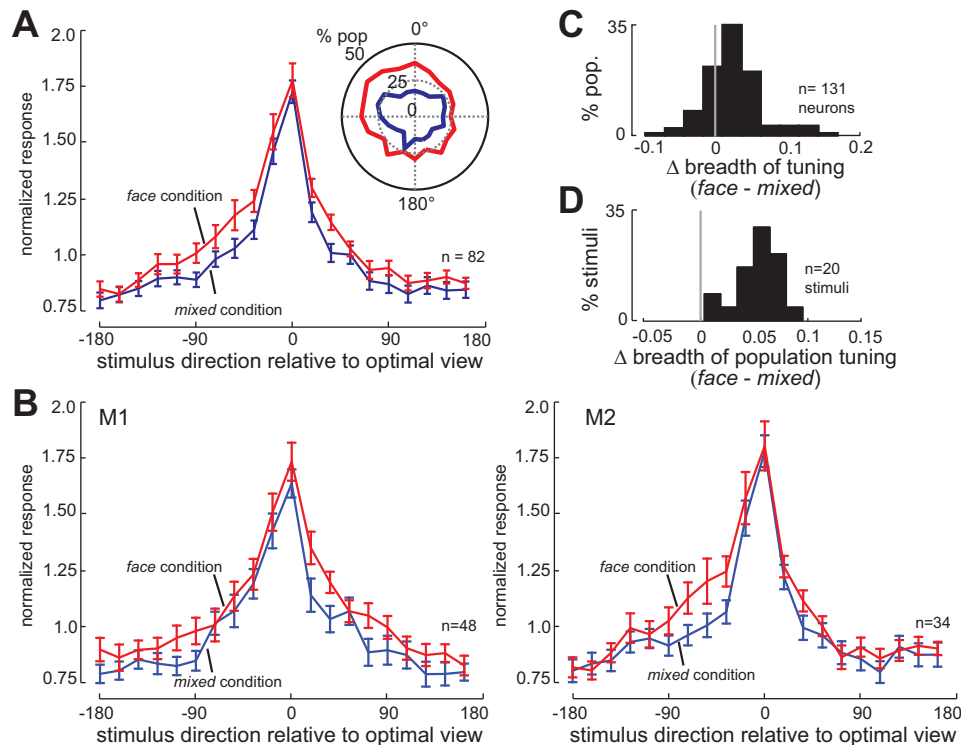


Fig. 3. Stimulus context alters population and neuronal breadth of tuning. *A*: normalized responses of the population of 82 view-tuned IT neurons to 20 views of face stimuli, ordered relative to their optimal viewing angle, for the face condition (red) and the mixed condition (blue). Neurons respond more strongly to suboptimal viewing angles in the face condition. *Inset*: polar plot of proportion of neurons (out of 131 neurons) responsive to each of 20 presented face views in face (red) and mixed (blue) conditions. A larger proportion of the IT population responds to each face view in the face condition. *B*: normalized responses of 48 view-tuned neurons from monkey 1 (M1; *left*) and 34 view-tuned neurons from monkey 2 (M2; *right*) to 20 views of face stimuli, ordered relative to their optimal viewing angle, for the face condition (red) and the mixed condition (blue). Neurons in both monkeys respond more strongly to suboptimal viewing angles in the face condition. *C*: histogram of the change in neuronal breadth of tuning for face views (face condition – mixed condition) for all 131 face-responsive IT neurons. A positive mean indicates that neurons are responding to a larger fraction of the face stimuli in the face condition. *D*: histogram of change in population breadth of tuning (face condition – mixed condition). A positive mean indicates that each stimulus is activating a larger fraction of the neuronal population in the face condition.

population of 82 neurons that exhibit a robust tuning peak (M1, $n = 48$; M2, $n = 34$; see MATERIALS AND METHODS for details), we found an increase in the visual response across all face views (repeated-measures ANOVA: viewing angle, $F = 53.12$, $P < 10^{-3}$; face vs. mixed condition, $F = 5.00$, $P = 0.0281$). This increased visual response was associated with a broadening of neuronal tuning, as quantified by the tuning bandwidth (full width at half-maximum) of neuronal responses to the 20 different viewing angles (Földiák et al. 2004). The tuning bandwidth for the population of neurons was $134.29 \pm 3.96^\circ$ during the mixed condition, which increased to $141.76 \pm 3.55^\circ$ during the face condition (bandwidth face vs. mixed, $P = 0.023$). This increased visual response across face views also resulted in an increase in the number of neurons responding to each stimulus (Fig. 3A, inset; see experimental conditions in MATERIALS AND METHODS). This increase in the tuning bandwidth was significant in M1 and showed a similar trend in M2 (Fig. 3B; Δ bandwidth face vs. mixed, M1 = 9.39 ± 4.96 , $P = 0.0321$; M2 = 6.29 ± 5.00 , $P = 0.0734$). These results demonstrate that more IT neurons respond to a given stimulus when it is more likely in the context of recent stimulus history. We also quantified the distribution of information processing using neuronal and population measures of representation sparseness (as defined in Rolls and Tovee 1995; see MATERIALS AND METHODS for details); neurons need not have a robust tuning peak to be included in this analysis. Breadth of tuning is inversely correlated with sparseness: broader tuning corresponds to a more distributed, less sparse representation. As shown in Fig. 3C, for the population of 131 neurons, the breadth of tuning significantly increased in the face condition compared with the mixed condition (Δ breadth of tuning = 0.021 ± 0.005 , $P < 10^{-3}$). Thus, when face stimuli are likely in the context of recent stimuli, IT neurons broaden their face view tuning. We also quantified the increase in the proportion of neurons responding to a given stimulus in terms of the breadth of population tuning, measured by a sparseness index (see MATERIALS AND METHODS). Larger values reflect a larger fraction of the neuronal population responding to each stimulus. As shown in Fig. 3D, the breadth of population tuning also significantly increased in the face condition compared with the mixed condition (Δ breadth of population tuning = 0.055 ± 0.002 , $P < 10^{-3}$).

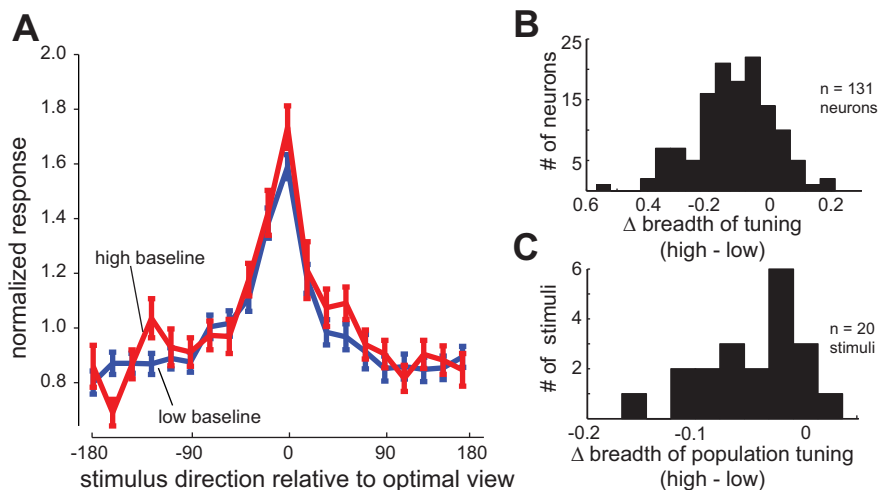
To make sure that this increase in the breadth of tuning was not attributable to changes in mean rate or the specific metric used, we verified the finding using other measures of tuning breadth. To account for changes in mean firing rate, we applied the same formula as above (Rolls and Tovee 1995), but adding the difference between condition means to the mixed responses, which should correct for the effects of baseline changes; both individual and population breadth of tuning still increased in the face condition (Δ breadth of tuning = 0.014 ± 0.006 , $P = 2.99e-04$; Δ breadth of population tuning = 0.325 ± 0.017 , $P < 10^{-3}$). We also measured the DOS (used in Moody et al. 1998; Mruczek and Sheinberg 2012; Rainer and Miller 2000), and found greater DOS in the face condition (DOS face = 0.349 ± 0.013 , DOS mixed = 0.401 ± 0.012 , $P < 10^{-3}$). An SI (Mruczek and Sheinberg 2012) also showed reduced stimulus selectivity, corresponding to broader tuning, in the face condition (SI face = 0.443 ± 0.017 , SI mixed = 0.528 ± 0.017 , $P < 10^{-3}$).

These results suggest that stimulus representation in IT becomes more distributed in a predictable context (to see how this more distributed representation affects the population's information content, see the *Population Coding* section below). Next, we tested whether spontaneous fluctuations in baseline activity were sufficient to drive the observed effects, or if instead they depend on the stimulus context.

We have shown that the firing rates just prior to the stimulus-evoked response are higher in the face condition than in the mixed condition. To ensure that the observed changes in single-neuron and population tuning were due to the effects of the block context, and not merely a secondary consequence of the changes in activity preceding the visual response, we looked at the effects of spontaneous fluctuations in activity on subsequent visual responses. Trials during the mixed condition were divided into high-baseline trials, in which the firing rate during the anticipatory response period was greater than the average rate for that neuron during the same time period as in the face condition, and low-baseline trials (at or below the face condition rate; across neurons, the percentage of trials in which the mixed condition rate exceeded the face condition average ranged from 29.3 to 92.9%; mean = $65.6 \pm 0.14\%$). We then repeated the analysis previously used to compare the face and mixed conditions on these high- and low-baseline trials. In every case, there was either no effect of the high-baseline activity, or the effect was in the direction opposite that observed during the face condition. As shown in Fig. 4A, for the population of 82 neurons that exhibit a robust tuning peak, there was no increase in the visual response in high-baseline trials (Fig. 4A; repeated-measures ANOVA: significant effect of viewing angle, $F = 27.42$, $P < 10^{-3}$; no effect of high vs. low baseline, $F = 2.51$, $P = 0.201$; no interaction, $F = 0.91$, $P = 0.566$). Thus a high-baseline firing rate does not itself produce an enhanced visual response to the following stimulus presentation. As shown in Fig. 4B, for the full population of 131 neurons, single-neuron breadth of tuning significantly decreased in the high-baseline trials compared with the low-baseline trials (Δ breadth of tuning = -0.1172 ± 0.0114 , $P < 10^{-3}$); this effect was opposite that of the face condition. Likewise, breadth of population tuning significantly decreased in the high-baseline trials compared with the low-baseline trials (Fig. 4C; Δ breadth of population tuning = -0.0461 ± 0.0042 , $P < 10^{-3}$), indicating that spontaneous fluctuations in baseline activity actually affect the proportion of IT neurons responding to a particular stimulus in a manner opposite to the effect of the block condition.

The change in firing rate prior to the stimulus-evoked response during the face condition was a predictor of changes in the subsequent visual response (see next section); however, high-baseline trials in the mixed condition did not exhibit the same changes in their subsequent visual responses. This indicates that the changes in object representation do not result merely from metabolic or "sensitizing" effects of a higher firing rate within the IT neuron. Presumably the high-baseline trials in the mixed condition are the result of fluctuations in noisy input from multiple feedforward and feedback pathways, whereas the anticipatory enhancement in the face condition is the result primarily of category-selective feedback or adaptation of the feedforward pathway, whose effects are reflected as an elevated activity prior to the evoked response and a facilitated processing during stimulus period.

Fig. 4. Trials with higher spontaneous baseline activity in the mixed condition do not mimic the effects of block statistics. **A**: normalized responses of the population of 82 face view-tuned IT neurons to 20 views of face stimuli, ordered relative to optimal viewing angle, for the high-baseline trials (red) and the low-baseline trials (blue) of the mixed condition. Neurons do not respond more strongly to suboptimal viewing angles on trials with high-baseline activity. **B**: histogram of the changes in the single-neuron breadth of tuning for face views (high baseline – low baseline), for all 131 face-responsive neurons. A negative mean indicates that neurons are responding to a smaller fraction of the face stimuli in the high-baseline trials, opposite the effect seen during the face condition. **C**: histogram of the changes in the breadth of population tuning (high baseline – low baseline) for each of the 20 views of face stimuli. A negative mean indicates that each stimulus is activating a smaller fraction of the neuronal population in the high-baseline trials, again opposite the effect seen during the face condition.



Anticipatory Response Is Tightly Linked with Enhanced Representation

We found that the face condition enhanced the anticipatory activity of face-responsive IT neurons, broadened face view tuning, increased reliability, and made the representation of face stimuli more distributed. Importantly, all of these effects are restricted to a subset of IT neurons. We found that only ~30% of the population of IT neurons (41 out of 131 neurons) exhibited the anticipatory response modulation (i.e., a significant increase in their activity measured from 0 to 70 ms after stimulus onset in the face condition compared with the mixed condition; significant effect of condition in an ANOVA, $P < 0.05$). Looking at the frequency of anticipatory activity across the subregions of IT, we found no difference in the distribution of anticipatory activity between regions; 15 out of 48 superior temporal sulcus (STS) neurons, 12 out of 42 neurons in dorsoanterior subdivision of TE (TEad), and 14 out of 41 neurons in ventroanterior subdivision of TE (TEav) showed an anticipatory response, none of which are different from the proportion in the population as a whole (Pearson's χ^2 test comparing the proportion with anticipatory activity in an area to the whole population: STS $\chi = 0$, $P = 0.995$; TEad $\chi = 0.111$, $P = 0.739$; TEav $\chi < 0.117$, $P = 0.733$).

In Fig. 5A, we show the visual response of the population of 131 neurons, with stimuli ranked based on their evoked activity. There was a significant effect of the stimulus context on visual responses (repeated-measures ANOVA: significant effect of viewing angle $F = 335$, $P < 10^{-3}$; significant effect of face vs. mixed condition $F = 4.79$, $P = 0.0304$; interaction, $F = 3.12$, $P < 10^{-3}$); the direction of this effect was consistent but not significant in the two monkeys separately (repeated-measures ANOVA: M1, significant effect of viewing angle $F = 176$, $P < 10^{-3}$; face vs. mixed condition $F = 2.01$, $P = 0.160$; interaction, $F = 1.47$, $P = 0.0885$; M2, significant effect of viewing angle $F = 174$, $P < 10^{-3}$; face vs. mixed condition $F = 3.08$, $P = 0.0851$; interaction, $F = 2.53$, $P < 10^{-3}$). Next, we examined the subsets of neurons with and without anticipatory activity. We found that those IT neurons with anticipatory enhancement exhibited enhanced responses in the face condition (repeated-measures ANOVA: significant effect of viewing angle $F = 216$, $P < 10^{-3}$; significant effect of face

vs. mixed condition $F = 21.8$, $P < 10^{-3}$; interaction, $F = 2.61$, $P < 10^{-3}$). As shown, neurons without anticipatory activity ($n = 90$; Fig. 5, right) did not show a significant effect

of stimulus context on firing rates (repeated-measures ANOVA: significant effect of viewing angle $F = 215$, $P < 10^{-3}$; no effect of face vs. mixed condition $F = 0.003$, $P = 0.958$; interaction, $F = 1.96$, $P = 0.008$). The change in responses was significantly larger for neurons with anticipatory activity than those without (mixed-model ANOVA on the change between face and mixed conditions: significant effect of anticipatory activity, $P < 10^{-3}$; no effect of view angle, $P = 0.50$; no interaction, $P = 0.58$). The breadth of tuning of individual neurons significantly increased in the face condition compared with the mixed condition for 41 neurons exhibiting an anticipatory enhancement (Δ breadth of tuning_{anticipatory} = 0.032 ± 0.004 , $P < 10^{-3}$), which was significantly greater than changes in the 90 neurons lacking anticipatory activity modulation (Δ breadth of tuning_{anticipatory} – Δ breadth of tuning_{nonanticipatory} = 0.0165 , $P = 0.010$). Similarly, the breadth of population tuning significantly increased in the face condition compared with the mixed condition for the 41 neurons exhibiting an anticipatory enhancement (Δ population tuning_{anticipatory} = 0.082 ± 0.01 , $P < 10^{-3}$), which was significantly greater than changes in the 90 neurons lacking anticipatory activity modulation (Δ population tuning_{anticipatory} – Δ population tuning_{nonanticipatory} = 0.027 ± 0.002 , $P = 0.014$). Thus neurons exhibiting anticipatory enhancement also displayed larger tuning changes associated with a more distributed stimulus representation.

Context-dependent changes in the reliability of individual neurons were also restricted to neurons displaying anticipatory activity. Figure 5B, left, illustrates the Fano factor for stimuli ranked according to the magnitude of their visual response in each condition. As shown, Fano factor is reduced in the face condition compared with the mixed condition across all stimuli (repeated-measures ANOVA; significant effect of face vs. mixed condition, $F = 6.59$, $P = 0.011$; significant effect of stimulus rank, $F = 4.60$, $P < 10^{-3}$; no interaction, $F = 0.43$, $P = 0.985$; M1, face vs. mixed condition, $F = 1.27$, $P = 0.264$, significant effect of stimulus rank, $F = 3.65$, $P < 10^{-3}$; no interaction, $F = 0.506$, $P = 0.961$; M2, significant effect of

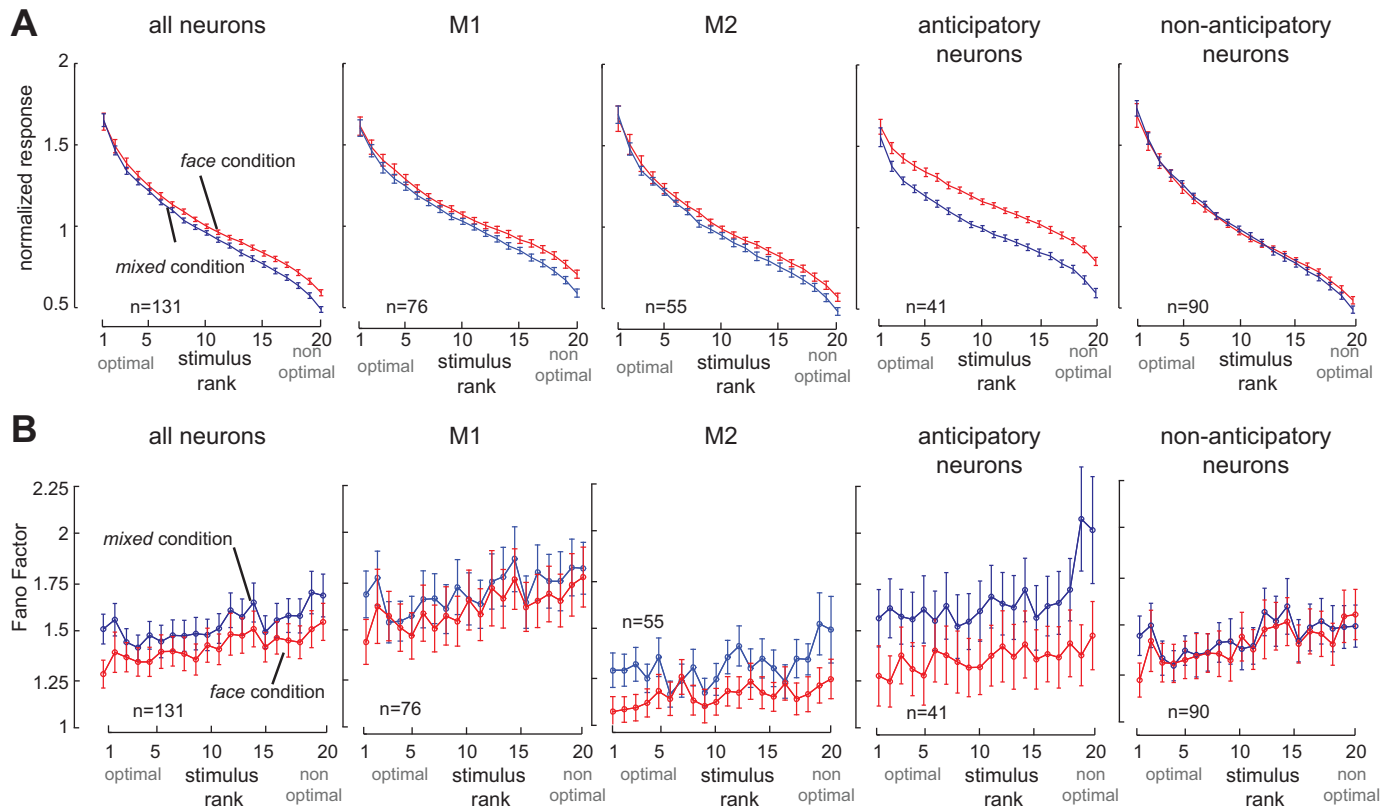


Fig. 5. Relationship between changes in baseline and stimulus-evoked activity. **A**: normalized responses of IT neurons to 20 different face views, arranged in order of stimulus rank (from highest to lowest visually evoked response for each neuron), in the face (red) and mixed (blue) conditions. From *left to right*, plots show all 131 face-responsive neurons, 76 neurons from *monkey 1* (M1), 55 neurons from *monkey 2* (M2), the 41 neurons with anticipatory activity prior to the stimulus-evoked response in the face condition (across both monkeys), and the 90 neurons lacking significant anticipatory activity. Changes in stimulus-evoked activity are most pronounced in neurons that also display anticipatory response enhancement. **B**: response variability (Fano factor) of IT responses to the 20 face views, arranged by stimulus rank as described for **A**. From *left to right*: all neurons, M1, M2, neurons with anticipatory activity, and neurons without anticipatory activity. Changes in Fano factor are most prominent in neurons with anticipatory activity.

face vs. mixed condition, $F = 11.8$, $P = 0.00120$, significant effect of stimulus rank, $F = 2.19$, $P = 0.0024$; no interaction, $F = 1.03$, $P = 0.423$). Neurons exhibiting anticipatory enhancement in their firing showed a dramatic reduction in their response variability (Fig. 5B; repeated-measures ANOVA: significant effect of face vs. mixed condition, $F = 36.5$, $P < 10^{-3}$; significant effect of stimulus rank, $F = 2.5$, $P < 10^{-3}$; no interaction, $F = 1.2$, $P = 0.26$). However, we did not find any significant change in the variability of responses in neurons lacking anticipatory activity (Fig. 5B, *right*; repeated-measures ANOVA: no effect of face vs. mixed condition, $F = 0.20$, $P = 0.65$; significant effect of stimulus rank, $F = 3.2$, $P < 10^{-3}$; no interaction, $F = 0.58$, $P = 0.92$). Averaged across stimuli, the reduction of variability in the face condition compared with the mixed condition was significant in neurons with anticipatory enhancement (Δ Fano factor_{anticipatory} = -0.294 ± 0.048 , $P < 10^{-3}$, $n = 131$), but not in neurons lacking anticipatory enhancement (Δ Fano factor_{nonanticipatory} = -0.030 ± 0.058 , $P < 0.600$, $n = 90$; Δ Fano factor for anticipatory neurons vs. nonanticipatory neurons, $P = 0.002$, Wilcoxon rank sum test). Therefore, a predictable stimulus context increases the reliability of responses in IT neurons displaying anticipatory activity. Although Fano factor itself is not measurable in the single stimulus presentation in which behavioral discrimination must occur, responses closer to the average learned response for a stimulus are presumably less likely to be incorrectly decoded

as representing other stimuli, and thus more reliable responses potentially improve the ability to discriminate between stimuli.

Population Coding

The main goal of the present study is to test whether the changes in visual responses observed during the face condition facilitate the processing of incoming visual information by IT neurons. The analysis of neural tuning and sparseness (Fig. 3) showed that the face condition produced broader tuning of IT neurons for viewing angle, and each stimulus activated a more distributed representation. Although it is often assumed that sharper tuning improves the quality of population coding, in fact broader tuning may improve the population representation, depending upon factors including the stimulus dimensionality, stimulus statistics, and interneuronal correlations (Ganguli and Simoncelli 2014; Montemurro and Panzeri 2006; Seriès et al. 2004; Zhang and Sejnowski 1999). In our data, stimuli produced broader tuning functions when presented in a series of other face stimuli. Do these changes in neuronal tuning enhance the population representation of stimulus information? To determine whether changes in visual responses were beneficial for population coding, we tested the ability of the IT population to differentiate between the different face stimuli (which varied in viewing angle) in the face and the mixed conditions.

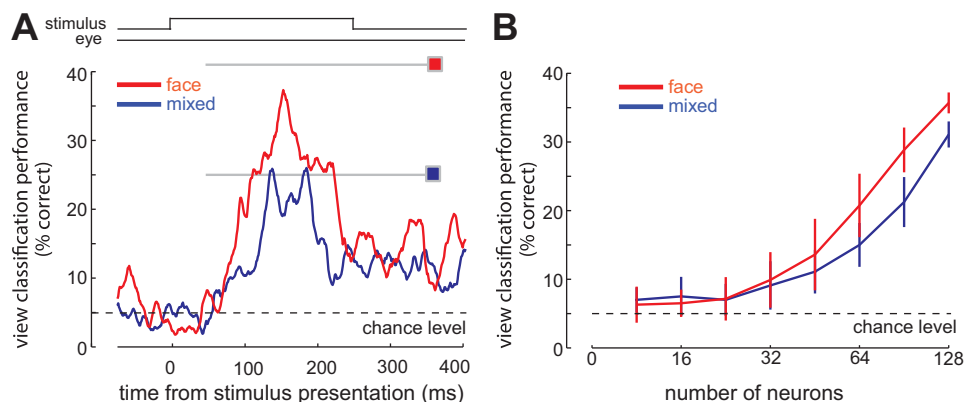
We quantified the ability of the neuronal population to discriminate between different face views by constructing a

SVM classifier and found that the population of IT neurons is better able to discriminate between face views in the face condition than in the mixed condition (Fig. 6). Figure 6A illustrates the classifier performance over the time course of visual responses in the face and mixed conditions. Performance of the classifier over the entire stimulus-evoked response (70–320 ms after stimulus onset) in the face condition was 43%, which is significantly greater than both chance level (chance performance, 5%; face vs. chance performance, $P < 10^{-3}$, Jackknife resampling) and the performance of a classifier based on responses in the mixed condition (performance_{mixed} = 25%; face vs. mixed performance, $P < 10^{-3}$, Jackknife resampling). In other words, solely based on the firing rates of IT neurons, an external observer (classifier) could better discriminate among face views during the face condition than the mixed condition. We quantified the classifier's performance for different numbers of neurons in the face and mixed conditions, shown in Fig. 6B. Performance of the classifier was significantly greater in the face condition than the mixed condition when 83 or more neurons were used for classification (performance_{face} = 23.86 ± 3.98 , performance_{mixed} = 18.63 ± 3.65 , $P = 0.045$). The higher performance of the classifier using responses from the face condition was more prominent when larger numbers of IT neurons were used (two-way ANOVA, significant effects of number of neurons, $F = 10.87$, $P < 10^{-3}$; significant effects of condition, $F = 6.87$, $P = 0.003$; interaction, $F = 4.56$, $P = 0.009$). Altogether, we found that fewer neurons would be needed to achieve the same level of view classification performance in the face condition compared with the mixed condition, indicating that a predictable stimulus context enhanced the neural population representation of face images. As with the previous effects, the SVM performance enhancement in the face condition was only observed for the subset neurons with anticipatory activity. For neurons with anticipatory activity, SVM classifier performance in the face condition was higher than in the mixed condition by 17%, which was significantly greater than the difference between conditions for nonanticipatory cells (Δ performance nonanticipatory, face – mixed = -3% ; anticipatory vs. nonanticipatory change, $P < 0.001$).

The SVM classifier just discussed can potentially utilize both changes in the mean firing rate and in the reliability of firing across trials; to determine whether changes in mean firing rate, independent of any changes in reliability, are themselves beneficial for the population stimulus representation, we used a multidimensional scaling model of the neural

space (Fig. 7). We constructed a space of neuronal responses solely based on the average firing rates of 131 IT neurons to the 20 face views. For visualization purposes, Fig. 7A depicts a two-dimensional projection of the space of neuronal responses in the face and mixed conditions using multidimensional scaling (Rolls and Tovee 1995) (see below). Each line in Fig. 7A connects the corresponding points for one of the 20 face stimuli within the multidimensional space of neural responses, from the mixed condition (bottom, blue) to the face condition (top, red). Face views move away from the center of the neural space (0,0), resulting in an expansion of the space of neural responses in the face condition compared with the mixed condition. We quantified this expansion by measuring the distance of each of 20 face views from the center of the space of neural responses in the full multidimensional space in the face and the mixed conditions (Fig. 7B). The average distance of stimuli from the center in the mixed condition was 2.563 ± 0.199 , which significantly increased to 3.166 ± 0.251 in the face condition ($P = 0.011$), implying an expansion of the perceptual space. We also found that the average distance between pairs of stimuli in neural space increased during the face condition: such an increase should correspond to the ability to more easily discriminate between stimuli based on neural responses. As shown in Fig. 7C, a predictable stimulus context significantly increased the interstimulus distance, from 3.596 ± 0.114 in the mixed condition to 4.308 ± 0.163 in the face condition ($P < 10^{-3}$). We observed that the expansion is stronger for stimuli far from each other in the neural space (distance > 3 : from 4.548 ± 0.085 in the mixed condition to 5.445 ± 0.153 in the face condition, $P < 10^{-3}$) compared with stimuli close to each other (from 1.764 ± 0.084 in the mixed condition to 2.120 ± 0.176 in the face condition, $P = 0.081$). These results verify that the context-induced changes in average firing rate of neurons expand the neural space for representing face views, making them a potential means for improving neural discrimination between face views. Both the improved SVM performance and the increased interstimulus distances in the face condition demonstrate that, even though the ability of an individual neuron's ability to differentiate between face views may decrease (due to broader face view tuning), the ability of the population to differentiate face views increases; at the population level, the benefits of including more neurons in the representation

Fig. 6. Predictable stimulus statistics improve an observer's ability to discriminate between face views based on the population representation. **A**: performance of an SVM classifier in decoding viewing angle, using the responses of all 131 IT neurons in a sliding 100-ms window. Gray lines and colored squares indicate classifier performance using the entire visual response (70–320 ms after stimulus onset); overall performance is higher in the face condition (red) than in the mixed condition (blue). **B**: performance of the classifier as a function of the number of randomly selected neurons used for decoding, for the face condition (red) and the mixed condition (blue).



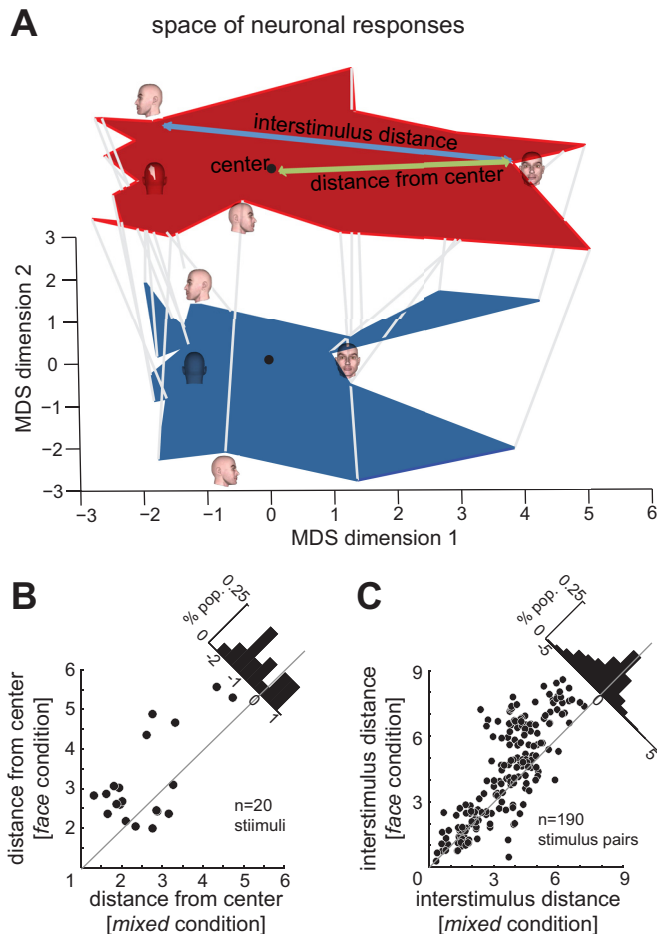


Fig. 7. Effects of stimulus context on the neuronal space representing face views. *A*: the space of neuronal responses to different face views, projected onto the two most significant MDS dimensions, in the mixed condition (blue) and the face condition (red, plotted in the same two dimensions but shifted vertically for visibility). Each gray line connects the location of the stimulus in the space generated based on responses of 131 IT neurons in the mixed condition to the location of the same stimulus in the space produced by responses of the same neurons in the face condition. The four face views show the locations of responses to those images in the neural response space in each condition. The green arrow indicates the distance of one stimulus from the center of the space (quantified for all stimuli in *B*). The blue arrow indicates the interstimulus distance for one pair of stimuli (quantified for all stimulus pairs in *C*). *B*: distance from the center of the neuronal response space for each of the 20 face view stimuli, in the mixed condition (*x*-axis) vs. the face condition (*y*-axis). Histogram in *top right*: change in distance from center of neuronal space between conditions; negative mean value indicates a shift away from the center in the face condition. *C*: distance in neuronal space between pairs of different face view stimuli, in the mixed condition (*x*-axis) vs. the face condition (*y*-axis). Histogram in *top right*: change in interstimulus distance between conditions; negative mean value indicates a greater interstimulus distance in the face condition.

outweigh the costs of reduced view-selectivity in the individual neurons.

DISCUSSION

We recorded the responses of IT neurons under conditions in which the appearance of a particular face was either certain, and therefore statistically predictable (face condition), or unpredictable (mixed condition). Although monkeys were not engaged in a behavioral task, the firing rate was greater in the face condition than in the mixed condition, demonstrating a

differentiation between the two conditions at a neural level. The increase in firing rate in the face condition compared with the mixed condition was observed even prior to the onset of the visually evoked response. In addition to this anticipatory enhancement, the face condition produced changes in visual responses: the overall average magnitude of the visual response increased, the tuning for different face views became broader, and the reliability of responses increased. These changes in tuning and reliability occurred predominantly in those neurons that exhibited enhanced anticipatory activity. Both of these changes were associated with an enhancement in the representation of information about viewing angle, reflected by the improved performance of an SVM classifier and an expansion in the space of neuronal responses. Effectively, the IT neuronal population facilitates face processing in the predictable context. Predictable stimulus onset, location, or identity have been shown to modulate visual responses (Anderson and Sheinberg 2008; Doherty et al. 2005; Meyer and Olson 2011); however, to our knowledge, this is the first study demonstrating that the statistical likelihood of the appearance of a face stimulus modulates visual processing at the level of single neurons, in such a way that the visual information can be read out more accurately from their population activity.

This paradigm does not allow us to conclude whether the effects of block statistics on face processing in IT are a result of top-down processes such as expectation, or some form of bottom-up adaptation to repetition or image statistics; likewise, we cannot determine whether the effects are forward-looking (such as anticipation, expectation or prediction) or based only on the stimulus history (and if so over what timescale). Some of the effects we see here, including changes in baseline visual response (McAdams and Maunsell 1999; Motter 1994; Treue and Martínez Trujillo 1999), increased reliability (Cohen and Maunsell 2009; Mitchell et al. 2007), and improved population encoding (Zhang et al. 2011), are similar to changes in neural activity during attention. These similarities make it tempting to suggest that these effects of block statistics are the result of top-down modulation. However, feature-based attention sharpens the tuning curves of extrastriate neurons (Martínez Trujillo and Treue 2004; Reynolds and Heeger 2009; Spitzer et al. 1988); in contrast, we found that predictable stimulus statistics broadened IT neural responses. This could be due to a difference in the specificity of the predictable stimulus properties in our paradigm compared with attention in previous studies; in feature attention studies, animals are usually cued to a particular value of the stimulus (e.g., a specific color or direction of motion). It is possible that a more restricted set of predictable stimuli (for example, repeated presentations of only one specific face view) would have a different effect on IT neural tuning.

Our results contrast with those of some previous neurophysiological studies on the effects of block statistics on single-unit responses in IT. One study used trials in which either two images appeared, or the same image appeared twice, then had blocks in which the probability of repeat vs. novel trials varied (Kaliukhovich and Vogels 2011). Repeated images evoked lower responses in IT neurons, but the block likelihood of stimulus repetition had no effect on this suppression. In this case, however, each trial was a novel image for the monkey,

and so an image-specific expectation would have to be established in the 750 ms between initial stimulus presentation and second image presentation; perhaps in this scenario it is unsurprising that they saw effects of stimulus repetition (which have often been reported for short ISI stimuli), but not a dependence on the overall statistics of the block. It may be that the mechanisms by which block statistics alter image processing are different from those underlying short-latency repetition suppression. Another factor possibly contributing to the difference in results is the use of face stimuli; it is possible that responses to faces, or any visual stimulus encountered frequently enough to have specialized processing mechanisms, are altered by statistical likelihood differently than less familiar stimulus classes (Grotheer and Kovacs 2014). Further work will be needed to establish whether the more distributed processing effects reported here are a general feature of predictable stimulus statistics or specific to the neural network specializing in face stimuli (Moeller et al. 2008; Tsao et al. 2006, 2008).

These results are the first study showing that predictable stimulus context enhances single-unit responses in IT cortex. The finding of a more distributed representation resulting in enhanced population encoding for predictable face stimuli offers a possible neural basis for the behavioral facilitation associated with contextual associations or expectation. However, before such a link can be established, several questions need to be answered. Further experiments will be necessary to determine whether the findings reported here generalize to other classes of objects or are specific to faces, or even whether they generalize across faces. Incorporating a behavioral component during future recordings could help control for the deployment of attention and begin to address whether the effects observed here are the result of top-down modulation or a form of bottom-up repetition effect.

ACKNOWLEDGMENTS

We thank M. Zirnsak for helpful comments on the manuscript.

GRANTS

B. Noudoost was supported by Whitehall 2014-5-18, National Science Foundation BCS143221, and National Eye Institute R01-EY-026924 grants.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

B.N. conceived and designed research; B.N. performed experiments; B.N. and N.N. analyzed data; B.N., N.N., K.C., and H.E. interpreted results of experiments; B.N. and K.C. prepared figures; B.N. and K.C. drafted manuscript; B.N., N.N., K.C., and H.E. edited and revised manuscript; B.N., N.N., K.C., and H.E. approved final version of manuscript.

REFERENCES

- Anderson B, Sheinberg DL.** Effects of temporal context and temporal expectancy on neural activity in inferior temporal cortex. *Neuropsychologia* 46: 947–957, 2008.
- Berniker M, Voss M, Kording K.** Learning priors for Bayesian computations in the nervous system. *PLoS One* 5: e12686, 2010.
- Chalk M, Seitz AR, Serries P.** Rapidly learned stimulus expectations alter perception of motion. *J Vis* 10: 2, 2010.
- Chawla D, Rees G, Friston K.** The physiological basis of attentional modulation in extrastriate visual areas. *Nat Neurosci* 2: 671–676, 1999.
- Chelazzi L, Duncan J, Miller EK, Desimone R.** Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol* 80: 2918–2940, 1998.
- Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI, Santhanam G, Sahani M, Shenoy KV.** Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13: 369–378, 2010.
- Cohen MR, Maunsell JH.** Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12: 1594–1600, 2009.
- Desimone R, Albright TD, Gross CG, Bruce CJ.** Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4: 2051–2062, 1984.
- Doherty JR, Rao A, Mesulam MM, Nobre AC.** Synergistic effect of combined temporal and spatial expectations on visual attention. *J Neurosci* 25: 8259–8266, 2005.
- Egner T, Monti JM, Summerfield C.** Expectation and surprise determine neural population responses in the ventral visual stream. *J Neurosci* 30: 16601–16608, 2010.
- Esterman M, Yantis S.** Perceptual expectation evokes category-selective cortical activity. *Cereb Cortex* 20: 1245–1253, 2010.
- Faulkner TF, Rhodes G, Palermo R, Pellicano E, Ferguson D.** Recognizing the un-real McCoy: priming and the modularity of face recognition. *Psychon Bull Rev* 9: 327–334, 2002.
- Földiák P, Xiao D, Keysers C, Edwards R, Perrett DI.** Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Prog Brain Res* 144: 107–116, 2004.
- Friston K, Kiebel S.** Predictive coding under the free-energy principle. *Philos Trans R Soc Lond B Biol Sci* 364: 1211–1221, 2009.
- Ganguli D, Simoncelli EP.** Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. *Neural Comput* 26: 2103–2134, 2014.
- de Gardelle V, Waszczuk M, Egner T, Summerfield C.** Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cereb Cortex* 23: 2235–2244, 2013.
- Grotheer M, Kovacs G.** Repetition probability effects depend on prior experiences. *J Neurosci* 34: 6640–6646, 2014.
- Hung CP, Kreiman G, Poggio TA, DiCarlo JJ.** Fast readout of object identity from macaque inferior temporal cortex. *Science* 310: 863–866, 2005.
- Kaliukhovich DA, Vogels R.** Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cereb Cortex* 21: 1547–1558, 2011.
- Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG.** Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22: 751–761, 1999.
- Kersten D, Mamassian P, Yuille A.** Object perception as Bayesian inference. *Annu Rev Psychol* 55: 271–304, 2004.
- Kiani R, Esteky H, Tanaka K.** Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *J Neurophysiol* 94: 1587–1596, 2005.
- Logothetis NK, Sheinberg DL.** Visual object recognition. *Annu Rev Neurosci* 19: 577–621, 1996.
- Martínez Trujillo JC, Treue S.** Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol* 14: 744–751, 2004.
- McAdams CJ, Maunsell JH.** Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19: 431–441, 1999.
- McMahon DBT, Olson CR.** Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. *J Neurophysiol* 97: 3532–3543, 2007.
- Meyer T, Olson CR.** Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc Natl Acad Sci U S A* 108: 19401–19406, 2011.
- Meyer T, Walker C, Cho RY, Olson CR.** Image familiarization sharpens response dynamics of neurons in inferotemporal cortex. *Nat Neurosci* 17: 1388–1394, 2014.
- Mitchell JF, Sundberg KA, Reynolds JH.** Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* 55: 131–141, 2007.

- Moeller S, Freiwald WA, Tsao DY.** Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science* 320: 1355–1359, 2008.
- Montemurro MA, Panzeri S.** Optimal tuning widths in population coding of periodic variables. *Neural Comput* 18: 1555–1576, 2006.
- Moody SL, Wise SP, di Pellegrino G, Zipser D.** A model that accounts for activity in primate frontal cortex during a delayed matching-to-sample task. *J Neurosci* 18: 399–410, 1998.
- Motter BC.** Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J Neurosci* 14: 2178–2189, 1994.
- Mruczek REB, Sheinberg DL.** Stimulus selectivity and response latency in putative inhibitory and excitatory neurons of the primate inferior temporal cortex. *J Neurophysiol* 108: 2725–2736, 2012.
- Noudoost B, Esteky H.** Neuronal correlates of view representation revealed by face-view aftereffect. *J Neurosci* 33: 5761–5772, 2013.
- Puri AM, Wojciulik E, Ranganath C.** Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Res* 1301: 89–99, 2009.
- Puri AM, Wojciulik E.** Expectation both helps and hinders object perception. *Vision Res* 48: 589–597, 2008.
- Rainer G, Miller EK.** Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27: 179–189, 2000.
- Reynolds JH, Heeger DJ.** The normalization model of attention. *Neuron* 61: 168–185, 2009.
- Rolls ET, Tovee MJ.** Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J Neurophysiol* 73: 713–726, 1995.
- Rust NC, DiCarlo JJ.** Balanced increases in selectivity and tolerance produce constant sparseness along the ventral visual stream. *J Neurosci* 32: 10170–10182, 2012.
- Seriès P, Latham PE, Pouget A.** Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nat Neurosci* 7: 1129–1135, 2004.
- Seriès P, Seitz AR.** Learning what to expect (in visual perception). *Front Hum Neurosci* 7: 668, 2013.
- Spitzer H, Desimone R, Moran J.** Increased attention enhances both behavioral and neuronal performance. *Science* 240: 338–340, 1988.
- Sterzer P, Frith C, Petrovic P.** Believing is seeing: expectations alter visual awareness. *Curr Biol* 18: R697–R698, 2008.
- Treue S, Martínez Trujillo JC.** Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399: 575–579, 1999.
- Tsao DY, Freiwald AW, Tootell RBH, Livingstone MS.** A cortical region consisting entirely of face-selective cells. *Science* 311: 670–674, 2006.
- Tsao DY, Schweers N, Moeller S, Freiwald WA.** Patches of face-selective cortex in the macaque frontal lobe. *Nat Neurosci* 11: 877–879, 2008.
- Vinje WE, Gallant JL.** Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *J Neurosci* 22: 2904–2915, 2002.
- Woloszyn L, Sheinberg DL.** Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. *Neuron* 74: 193–205, 2012.
- Zhang K, Sejnowski TJ.** Neuronal tuning: To sharpen or broaden? *Neural Comput* 11: 75–84, 1999.
- Zhang Y, Meyers EM, Bichot NP, Serre T, Poggio TA, Desimone R.** Object decoding with attention in inferior temporal cortex. *Proc Natl Acad Sci U S A* 108: 8850–8855, 2011.

